Chapter- 6

Factors Determining Educational Deprivation in Rural Uttar Dinajpur District: An Empirical Exercise- I

6. Introduction

Chapter 5 illustrated the educational scenario of Uttar Dinajpur district and identified critical gaps in literacy achievement and educational attainment with a focus on elementary education. It has been found that educational deprivation in the district is mostly associated with Islampur sub-division. On the other hand, Raiganj subdivision is better placed, its outcomes being comparable with the state average. It is seen that the literacy rate of each of the CD Blocks in Islampur subdivision is much lower than the literacy rate of the blocks in Raiganj subdivision. Out of 10 least literate blocks of West Bengal (2001), 5 were found in Uttar Dinajpur district and all these 5 Blocks are in Islampur subdivision of the district. It has also been found that out of 207 deprived mouzas of the district, 182 mouzas (88% of total deprived mouzas) have been fallen at the Islampur subdivision (Table-6.1). To overcome the educational backwardness of Islampur subdivision thus appears to be more challenging in the district. Another unique feature of this district is that there is a large variation in rural-urban literacy rate. Literacy rate of the rural areas lie much below the state average, while the urban literacy rate is close to the state average of urban literacy rate. Moreover, it has been an experience of 2001 Census that the low literate areas of 1991 achieved higher literacy jump than the high literate regions/district. However, this has not been the experience of this district even at the sub-district level. Thus it appears that certain distinctive underlying socioeconomic factors could be responsible for the educational backwardness of this district, and especially in the region of rural Islampur subdivision. The understanding of the socioeconomic factors is better captured through intensive field survey which the present study has undertaken in the rural areas of Islampur sub-division of this district in order to identify the nature and causes of such backwardness in educational scenario.

Income deprivation and fallout of economic process is a popular way of addressing the problem of poverty. However, it is inadequate in answering the central question that underlies the proposition- 'Poverty of what'. (Nagaraja, 2004). It is commented that poverty can be explained better if preceded by some suitable socio-economic or any other variable. In order to interpret the multi-dimensional aspects of poverty, Nagaraja, (2004) used the term 'literacy poverty' which encompasses illiteracy and low levels of achievements in education. The World Development Report (2000-01) has also recognized this multi-dimensional concept of poverty, rather than using it as mere income deprivation. According to NSS 1999-2000, 27 percent of the rural households and 12 percent of the urban households did not have any literate adult (15+) (GOWB, 2004). Reviewing the progress of elementary education, the present Indian Prime Minister Dr. Manmohan Singh said, "only 47 out of 100 children enrolled in class-I reach class-VIII, putting the dropout rate at 52.79 per cent", which in his opinion is painful and unexpectedly high (The Hindu, New Delhi, February 22, 2005, cited in Ramachandran, 2005). This Chapter along with the next one is designed to analyse and represent the extent of educational poverty in terms of the quantum of deprivation in the educational attainment stressing on literacy attainment and elementary education. The whole analysis is carried out on the basis of primary field survey in an educationally deprived district of West Bengal.

It has been stated in Chapter 4 that there are 37,956 inhabited villages in the 17 districts of West Bengal (as per census 2001). As is evident from Census 2001, 843 of these villages have literacy rates below 25 percent (Table- 4.6 of Chapter-IV) and the identified villages

have been named as educationally deprived villages in the state. This one-fourth literacy rate norm is assumed to be the cut-off mark for educational poverty and may be attributed as the 'educational poverty line' [EPL]. The villages that could not maintain at least 25% literacy rate are said to lie below the EPL. The distribution of the deprived villages across the districts also highlights a significant point. Out of these 843 villages, around 46 percent of these villages are concentrated in three districts namely, Uttar Dinajpur, Malda and Murshidabad. These three districts alone account for 90 per cent of the aggregate population of the 843 EPL. A closer look at the disaggregated demographic structure of the population in these three districts reveal that they have very high concentration of Muslim population compared to the remaining districts of the state. Another notable feature is that the educationally deprived villages constitute a significant proportion of Schedule Tribe population (15 per cent). Thus, the Muslims and Schedule Tribe segments of the population constitute a significant segment of the total population of these EPL villages.

In order to substantiate the above, proportion of such deprived villages has been regressed on the proportions of Muslims, Scheduled Caste and Schedule Tribe (ST) population at the district level. The parameter estimates of Muslim and ST population are positive and statistically significant, while the same for SCs is insignificant with negative magnitude. This suggests that educational deprivation is more prominent among the Muslims and STs, while it is not so significant for the SCs.

	Coefficients	Std. Error	't' values
(Constant)	-2.66	3.52	-0.76
MUSLIM	0.71***	0.06	2.76
SC	-0.08	0.09	-0.37
ST	0.44*	0.15	1.72

Source: - Calculated from 'Census View, Directorate of Census Operation, GOWB, 2004

The above findings indicate that in West Bengal, educational deprivation is more likely to be associated with two distinct sections of the population - the Schedule Tribes and the Muslims. It is also evident from secondary data that it is the Muslim and ST females who are more deprived when compared to male counterparts. More than 50 per cent of the deprivation is observed in three districts of the state located in its northern part, which is popularly known as Uttar Banga or North Bengal. Thus, North Bengal merits special attention as well as additional intervention for its educational upliftment.

It is from this finding that an attempt has been made to carry out a village level survey for identification of the factors underlying the educational deprivation in West Bengal. The district of Uttar Dinajpur has been purposefully chosen for the analysis. It is so because the district is least literate in the state so far as total literacy and female literacy rate of the state is concerned and at the same time, it includes highest number of deprived villages that have been identified in the state. Moreover, it is the district with a substantial proportion of Muslim Population (47.36 per cent) and a certain proportion of ST Population (5.1 percent) too. These two sections of the population account for more than 50 percent of the total population in the district and therefore educational backwardness of the district also implies educational backwardness of these two sections of the population.

6.1 Background of the Analysis

Literature on economics of education has established that there is a positive association between educational backwardness and level of poverty. The explanation offered is that the opportunity cost of sending the children to school, instead of using them as household help or wage earner, is not an economically feasible option (Bhatty, 1998). This positive association

is emerged in different studies (Chakraborty, 2006; Duraisamy, 2004; Dholakia 2003; Reddy and Rao, 2003; Nambissan and Sedwal, 2002; Devi, 2001; Krishanji 2001, etc).

On the other hand, Nidhi Mehrotra (1995) on the basis of field- based information from Kerala, Uttar Pradesh and Himachal Pradesh, notes that evidence of child labour does not by itself establish that poverty is the prime reason for their not attending school (cited in Bhatty, 1998). Santha Sinha (2000) in her article noted that "—what is found is that not only are literacy rates similar between groups having dissimilar income levels but also vary widely between groups with same income levels. In other words, situations where better off families have engaged their children in work while parents with lower incomes have retained their children in school are not uncommon." Sinha also observed that there are factors other than the purely economic compulsions arising out of poverty, which dictate whether a child is sent to work or to school.

A common proposition that poverty alleviation is a prerequisite for achieving the goal of UEE has been falsified by several country experiences. Many countries have successfully made primary education compulsory and universal when per-capita income in those countries was low and poverty was wide spread. Japan introduced compulsory education in 1872, North and South Korea, Taiwan and People Republic of China, all of which made education compulsory shortly after the Second World War. In the West too, many countries have introduced the same before the industrial revolution. These countries have successfully ensured the universalisation of primary and elementary education in their country and they have regarded mass education as an instrument for the reduction of poverty (Weiner, 1996), justifying the need for education for poverty reduction. While studying the determinants of schooling for boys and girls in Nigeria under a policy of free primary education it has been found (Lincove, 2009) that controlling for costs, household wealth bears a positive relation with primary school attendance. Interestingly, it has greater income elasticity for girls than boys. Girls' attendance also depends on opportunity costs generated by providing child care for younger siblings and living on a family farm. Studies from other countries also suggest more or less similar results. Most studies analyzing the determinants of enrollment (especially girls' enrollment) have found the association between household income and enrollment in school to be positive and statistically significant, whether income is measured directly using a household consumption module or indirectly through some household asset index (Federal Bureau of Statistics, 1998; Hazarika, 2001; Sathar & Lloyd, 1994; World Bank, 2002). Both, the size and significance of income effects are typically larger for girls than boys when results for boys and girls are compared.

The impact of Female Labour Force Participation Rate (FLFPR) on child schooling is still a matter of debate. From the studies of Pandey (1990), Jeejeebhoy (1993), and Mukhopadhaya (1994) it is found that in general, FLFPR has a depressing effect on child schooling. This is partly because the daughters have to shoulder the responsibilities of household chores and sibling care and partly because the lack of maternal attention and supervision discourages children's, particularly girls' schooling. An important result has been found in the village level study of Sengupta et al (2002) for West Bengal. While they find mothers' work participation has a significant negative effect on daughters' school enrolment, it also has negative but not significant impact on grade completion. However, the factor does not appear to have a significant impact on the probability of dropout or retention in school. This has also been confirmed by Jaychandran (2001). The positive relationship between FLFPR and schooling of children, especially of girls, is however not found in Andhra Pradesh where high rates of FLFPR coexist with a high incidence of child labour (highest in India) and relatively low level of school attendance rates (Jaychandran, 2001 cited in Dreze and Sen, 2002; Rao and Reddy, 2003). A recent study (Reddy and Rao, 2003) in this area also does not find any significant impact of female work participation on the enrolment ratio of both male and female. They have concluded the result by using household level data for 12 villages in three districts of Andhra Pradesh. The studies of Psacharpoulos et al (1989) and Tansel (1997) also note that the positive effect of addition to resources from mothers' earnings can overshadow

the negative impact of mothers' absence from home. Similarly, Dreze and Sen (2002), while discussing the schooling revolution in Himachal Pradesh, opined that a high level of female labour force participation raises the economic returns to female education and it is also revealed that status of women, including their educational status, will improve as a consequence of their increasing participation in labour market and development process (Rekha Wazir, 2000).

Parental education emerges as a significant determinant in household education decisions. All the field studies done under the UNDP programme confirm this result (Bhatty, 1998). It is found in rural Punjab that illiteracy of the decision-making members in the family is an important reason for continued perpetuation of illiteracy among women (Thind and Jaswal, 2004). They also noted 'the resistance was very much based on the out-dated beliefs that a woman's place was inside the house and education was of no use to her'. Jeemol Unni (1998) in her study in rural Gujarat found that the education of both parents is positively associated with the schooling of the child. However, the gender differential that she observed was very interesting. While the fathers' education positively influences boys' schooling, the education of the mother has a strong positive influence on the education of the girl child only.

In a village level survey-based study in Orissa, Sailabala Devi (2001) observed that both father and mothers' education have a positive significant influence on the probability of enrolment in primary and upper-primary levels for boys and girls. But mothers' education has a strong influence than that of fathers on girls' enrolment. Parental level of education was also found to be significant in lowering the dropout rates. Sengupta et al. (2002) found similar result in their study in West Bengal. Malathy Duraisamy (2000) in her micro level study in two selected districts of Tamil Nadu found that a 10 percent increase in fathers' education leads to a 1 to 3 per cent increase in the probability of being enrolled and 0.1 to 0.2 percent increase in the educational attainment. A similar trend was also found in case mothers' level of education.

The study of Anuradha Pande (2000) in the rural hill areas of Uttar Pradesh revealed that the literacy level of the community as a whole has a significant impact on children's education-higher the literacy level, lower is the number of dropouts and non-enrolled children. The author also found it remarkable that fathers' education had a much greater influence on a child's chances of enrolment in school, especially of a girl child. Interestingly, it was found that the relationship between mothers' literacy level and educational status of a child is not significant unlike most other studies. An interesting result in this respect has also been emerged in a recent study (Vaid, 2004) where it has been found that gender and parental level of education is only significant for father's literacy at the stage of child's transition from illiteracy to some primary school. She thus concluded that unlike the other study, a more educated mother would lead to a higher education for the daughter does not hold. It should be noted that she used the National Election Study data set (1996) of the Centre for the Development Societies, New Delhi.

However, Thomas (2000) in a study on selected backward villages in Kerala also found that across the villages, the proportion of never-enrolled children in the school going age-group (5-14 years) did not bear any systematic relation to overall literacy level. The study of Llyod and Brandon (1994) in Ghana has emphasized that mothers favour the education of sons over daughters because of their greater dependence on children in their old age and their expectation of greater monetary returns from investment in sons.

Kiran Bhatty (1998) in the article 'Educational Deprivation in India - A Survey of Field Investigations', has concluded that parental motivation is generally high (PROBE Survey also supported this proposition) particularly for male children but, for female children, however, it is still an obstacle. Job aspiration and improvement status are the main determinants of parental motivation for male education; in the case of female education, these motives have less influence. Jabbi and Rajyalakshmi (1997) found in Bihar that the reasons for non-

enrolment of children were more economic and home related in the case of girls and more school related in case of boys.

School enrolment and attainment also bears a significant relationship with the size and composition of the family. It has been noted by several studies that there exists a negative relation between the number of children and child schooling, because the additional burden of children may put a restriction on family resources hampering child schooling. In this respect, there are many studies relating to Indias. Jejeebhoy (1993) in rural Maharastra found that an older girl child with many younger siblings has a corresponding lower chance of her schooling. The same results were found in the study of Psacharpolous et al, 1989 and Pandey, 1990. The lower chance of schooling of a girl child is particularly true, if there are younger male children in the family. Studies from other countries also support the result. Knodel and Wongsith (1990) in their study in Thailand found the similar negative impact of a girl child belonging to a larger family. Debi's study (2001) in rural Orissa observed that larger the number of infants and old persons, lower would be the enrolment rate and grade attainment of female children. By using the state level Indian data Reddy (1995) found the similar negative impact of the variable at the state level for the year 1991.

Krishanji (2001), by using the secondary level data for the inter district analysis in Andhra Pradesh, has used child-woman ratio as an explanatory variable to predict child enrolment at primary and elementary level. It was found that the variable has an adverse effect on enrolment of female children in the age group of 5-9 years. But the same was not found significant in explaining the enrolment ratios of both male and female children belonging to the age group of 9-14 years. This analysis thus, suggests that enrolment at primary level is more responsive to the number of siblings in the family. On the other hand, Jeemol Unni (1998) found somewhat different result in rural Gujarat. Unni, focusing on the schooling decision, observes the determinants of schooling and estimates the least square equations separately for boys, girls and all children. On estimation, it was found that the number of children per household did not show any significant result upon any of the three categories of children. Similarly, Duraisamy (2001) by using the village level data in Tamil Nadu found that the number of children in the household did not exert a statistically significant effect on school enrolment and grade attainment of both boys and girls at the primary level. But the variable was found to be significant while exerting a negative influence on enrolment and grade attainment at the secondary level.

6.2 Methodology and Study Area

It has been stated earlier that the prime objective of this chapter is to find out the nature and causes of educational poverty in a particular the state of West Bengal. It has also been stated that West Bengal has been purposively chosen because the state has been remaining closer to the national average literacy rate over the last 50 years (1951-2001) and is yet to move beyond this level. After selecting the state, the educational scenario of the state on the basis of literacy rate and elementary education has been assessed. From the assessment (discussed in Chapter-IV) it is observed that the district of Uttar Dinajpur in the state is one of the most deprived districts so far as its literacy trend and development of elementary education is concerned. As such the district has again been selected purposively to measure the educational deprivation at a more disaggregated level. Administrative frame of Uttar Dinajpur district comprises of two subdivisions - Islampur subdivision with five CD blocks and Raiganj subdivision with four CD blocks. In order to locate the educational deprivation within the district, literacy rates of the district at various disaggregated level has been calculated. Subdivision wise literacy comparison reflects (discussed in Chapter-V) that Islampur subdivision (38.5%) lie well below the Raiganj subdivision (58.1%). At the same time, the literacy rates of all the five CD blocks in Islampur subdivision are below the four CD blocks of Raiganj subdivision. The distribution of the 207 villages that lie below the EPL across the CD Blocks in each subdivision has been shown in Table-6.1. Out of these 207 villages, 182 (88%) villages are concentrated in Islampur subdivision and only 12% are in the other

subdivision. Along with this, the extent of out of school children is also seen to be concentrated in the Islampur sub-division too.

The primary survey was thus decidedly carried out in Islampur subdivision in order to have a better understanding of the problem of educational poverty in the district. Again, in Islampur Subdivision, Chopra block is found to have the least number of deprived villages (11), while Goalpokhar - I & II (until recently, a single block) has the highest number of such deprived villages and together also have the highest number of out of school children. Moreover, from the block level literacy rate, it is found that Goalpukur-I with 31.6% literacy rate and Chopra with 43.29% literacy rate may be ranked as the least and highest literate blocks of the Islampur subdivision. So it is expected that the selection of the above mentioned two blocks would rightly represent the diverse level of educational development. Thus, in view of the objective of identifuing the nature and causes of educational variation, these two blocks have been selected for the final survey.

Table-6.1: Block-wise distribution of educationally deprived villages and out of school children

CD Block/ Subdivision/ District	No of Villages in with a Literacy rate <= 25%	Out of School Children Age 6 – 14 (Total)		
		Boys	Girls	Total
Chopra	11	8382	8447	16829
Islampur	17	9803	9228	19031
Goalpokhar - I	51	13400	13841	27241
Goalpokhar - II	55	7273	6850	14123
Karandighi	48	19136	18078	37214
Islampur Subdivision	***182***	***57994***	***56444***	***114438***
Raiganj	9	10242	9594	19836
Hemtabad	0	1708	1761	3469
Kaliaganj	0	3375	3228	6603
Itahar	16	8322	7992	16314
Raiganj Subdivision	***25***	**23647**	**22575**	**46222**
District Total	**207**	**81641**	**79019**	**160660**

Source: - Calculated from 'Census View, Directorate of Census Operation, GOWB, 2004 ; DISE-2003-04.

After the selection of two blocks, two villages (mouzas) from each of the blocks have been chosen. For this, the villages as per the literacy rate of 2001 have been taken into account, of which one low literate village lying below EPL and another comparatively literate village has been purposively selected from each block with a view to capture the regional variation in educational development. The stratified, purposive sampling design of the survey is sketched in the next page.

A priori information on socio-demographic characteristics of the study villages are analysed from the secondary data [Table-6.2] for a better understanding of the nature of the problem. Kantigach and Juropani are the two lowly literate villages from Goalpokhar-I (low literate block) and Chopra block (high literate block) with a dominant Muslim and Scheduled Tribe population respectively. Dangipara (from Goalpokhar-I block) and Uttar Bhagalpur (Chopra block) villages are the two comparatively highly literate villages from the two blocks respectively. The population distribution of these two highly literate villages reveals that the Dangipara is a Scheduled Caste concentrated village while the second has a concentration of general caste population. Thus the four villages altogether represent the different segments of population distribution of the district.

The literacy rate of Kantigach was found to be only 2.2% (Census 2001), while the same was as high as 72.5 % in Dangipara village followed by Uttar Bhagalpur and Juropani. As per

Census 2001, it is observed that the composition of workers in the four villages is of a diverse nature. In Kantigach, there is an absolute dependency on agriculture, while in Juropani and Uttar Bhagalpur, three-fourth of the workers work as Other Workers (OW) (Table-6.3). The proliferation of tea gardens in Chopra block has been the reason behind this concentration of OW in this block. So far the proportion of non-worker is concerned, Kantigach has the lowest proportion of population as non-worker. Only 38.4% of total population is non-worker out of which 26.5% is child population (Table-6.2).

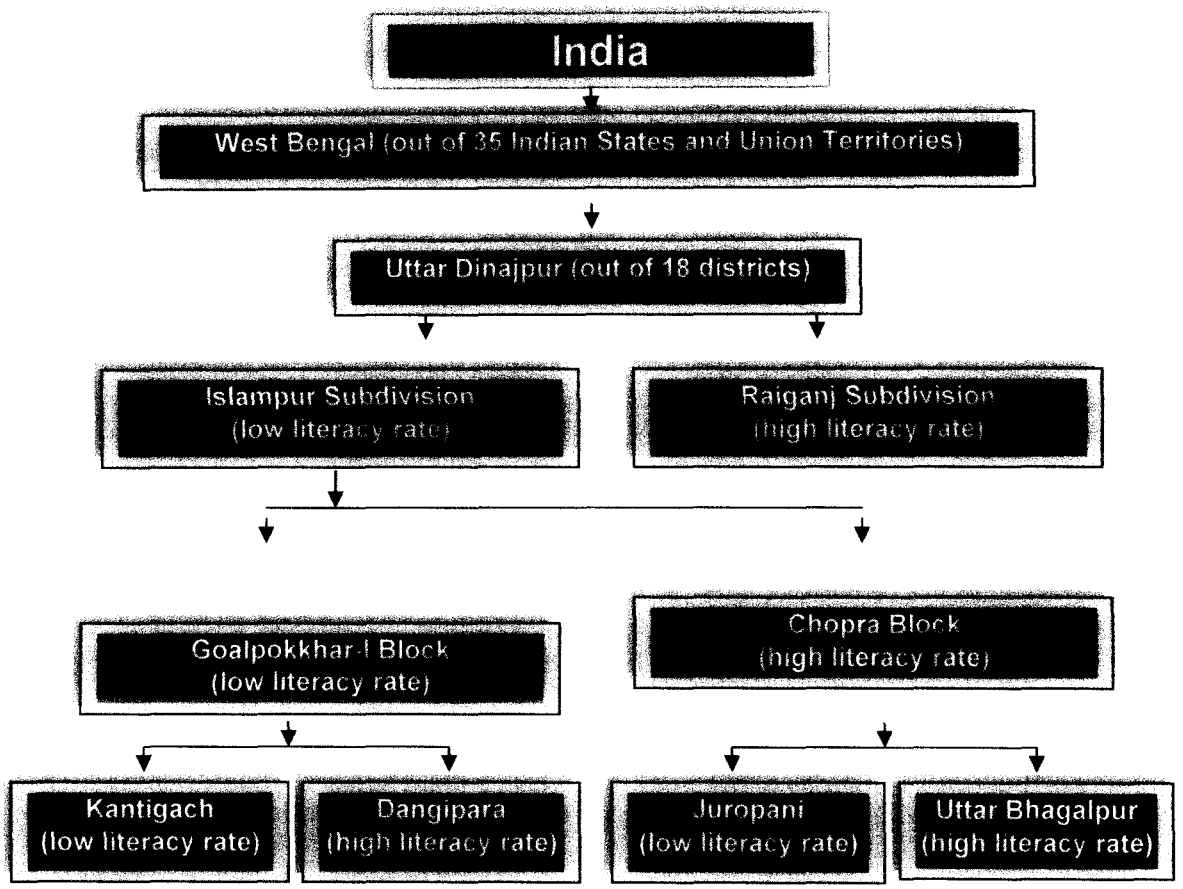


Table - 6.2: Socio-Economic Composition of the Study Villages

Block	Goalpukhar- I		Chopra	
Village/Indicators	Kantigach [LL]	Dangipara [HL]	Juropani[LH]	Uttar Bhagalpur [HH]
Population (Person)	555	721	598	826
No. of households	111	121	122	143
% of child (0-6) Population	26.5	18.3	27.3	15.6
Literacy rate, male	3.5	83.9	41.4	78.6
Literacy rate, female	1.0	58.8	17.8	54.2
Literacy rate, person	2.2	72.5	29.2	67.3
% of SC	0.0	99.6	1.2	0.6
% of ST	0.0	0.0	98.8	25.4
% of Non-worker	38.4	64.8	57.9	54.6
% of Agricultural Worker	10.8	57.5	0.4	3.5
% of cultivators	89.2	25.6	24.6	19.7
% of other worker	0.0	16.9	75.0	76.3

Source: - Calculated from 'Census View, Directorate of Census Operation, GOWB, 2004

If the old-age population is added with the child population, then it appears that most of the family members, apart from children and old-age members, have to work for a living. On the other hand, Dangipara has the lowest proportion of population in the workforce (notably a lower proportion of female work participation rate). The high work participation rate here does not necessarily mean a better economic condition. More often than not, one has to work for mere survival. Needless to say that educational attainment under such circumstances will not be affected by high work participation rate (as it was found in our earlier mouza level regression analysis too). However, a better understanding will emerge in the latter section from the analysis of field based data. A notable difference of the two type of villages is that in the high literate villages (in both the blocks) it is seen that there is a small proportion of child population (Table-6.2) apparently indicating a lower birth rate, while establishing the inverse relationship between literacy and fertility rate.

For the detailed survey, a complete enumeration (household census) of each and every household in the villages was first conducted. From the household census, the data on school-going age (5-14) children was primarily collected. The procedure facilitated in distinguishing the households categorically into two groups - households with school-going age (5-14Yrs) children and households that do not having any children belonging to this category. As such 135 households in Kantigach, 81 in Juropani, 122 in Dangipara and 139 in Uttar Bhagalpur were found to have children in the schooling age group (Table-6.2). From the listing of these households, 32 households from each village have been chosen by random sampling method for a detail household survey. It may however be noted here that 30 households from each village have actually been surveyed keeping 02 households as additional if at the time of survey some of the chosen households were not available. It has been calculated that out of total 1081 children in 477 households of 4 villages, 250 were covered under the detail survey thereby covering 23.1% of the total children in this respect. Again, out of total 477 households of 4 villages with some schooling age children, 120 households have been surveyed and as such a 25.2% of the total households comprise the actual sample size in the survey process. A detail picture is being depicted in Table 6.3.

Table - 6.3: Household and Children in the Survey

INDICATORS	Kantigach	Juropani	Dangipara	Uttar Bhagalpur	TOTAL
Total No. of Households	169	106	149	172	596
No of Households with some Schooling age (5-14Yrs) Children	135	81	122	139	477
No of Households with no Schooling age (5-14Yrs) Children	34	25	47	33	139
Size of the Sample Households	30	30	30	30	120
% of the Households Surveyed	22.2	37.0	24.6	21.6	25.2
Total No. of 5-14 years age group children in the villages	282	211	289	299	1081
Children covered in the Survey	59	72	58	61	250
% of children Surveyed	20.9	34.1	20.1	20.4	23.1

Source: - Field Survey

6.2.1 Survey Questionnaire

After the selection of two villages in Uttar Dinajpur district, they were primarily assessed on the basis of information taken from some informed villagers/panchayat level leadership and teachers. An open discussion was organized in some informal way by accumulating the villagers at a common place. The discussion was carried out with a view to identifying the household level socio-economic factors that may have their effect in connection with the educational development. This has primarily helped to ascertain the plausible set of explanatory variables. On the basis of this, a questionnaire for the households was prepared (APPENDIX-6.I).

6.2.2 Field Work

The survey of the villages took place in between November, 2006 and December, 2007. In collecting the data some language problem was faced in Kantigach and Juropani villages because of their mother tongue. For this, some local educated youths and teachers assisted the investigators for better understanding of the response to the questionnaire. This facilitated in getting the data very close to the actual answers.

6.2.3 Statistical Technique

The main objective of this research is to examine the relationship between the educational attainment and a set of socio-economic demand side variables at household level. With the help of least square procedures (Myers, 1990), multiple regression analysis has been attempted which has become one of the most common statistical techniques for investigating and modeling relationships among variables. In behavioral science an outcome is attributable to several factors. In such a case, the true effects of a particular independent variable or potential predictor variable (X) are assessed by 'controlling' out the effects of other X variables. For this simple reason, performing multiple regression has become a popular technique to analyse the set of data.

Depending upon the purpose of the estimation of the model, multiple regression is generally used for prediction and explanation. In optimizing the prediction objective, the task is to eliminate the superfluous variables, not to test theoretically based hypotheses. The parameter estimates here are of little importance. But, prediction does not always suffice the purpose of all analyses. Explanation is the essence of behavioral research. Multiple regression through explanation allows to separate causal factors, analyzing each one's influence on the explained variable. Most of the regression applications in social science are for explanatory purposes, and this approach has been introduced in this the present analysis of data.

Two main indicators of educational development have been decided for regression analysis – literacy attainment and child schooling. Child schooling has separately been discussed and analysed in the next chapter. This chapter exclusively concentrates on literacy deprivation and variation in the study area. Literacy is a widely used concept and has unanimous measurement criteria in the literature of education. The literacy rate is quantitative in nature and in order to deal with such a response variable multiple regression technique using the ordinary least square method may be applied.

The fundamental model underlying multiple regression analysis posits that a continuous outcome variable is, in theory, a linear combination of a set of predictors and error. Thus, for an outcome variable, Y, and a set of k predictor variables, X1,...,Xk, the Multiple regression model takes the form

$$Y_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots\ldots + \beta_k X_{ik} + u_i$$

Where α is the Y-intercept (i.e., the expected value of Y when all X's are set to 0), βjs are multiple (partial) regression coefficients and u_i is the error of prediction.

As it is seen in the above equation, there are 'k' regressors (explanatory variables) and a constant which are fundamentally unobservable. The errors that are also unobservable accommodate all other influences on Y not accounted for in the above equation. As such this unobservable error may be assumed to be normally distributed with the assumption E(ui)= 0 for all i, E(ui)= δ^2 for all i, E(ui uj)= 0 for all $i \neq j$ and E(ui Xi) = 0 for any i and j. Since the error term is mean independent of the Xs, varying the X's does not have an impact on the error term. Hence, there are k regressors (explanatory Variables) whose coefficients i.e., βs and a constant or intercept term α (i.e. k+1 parameters) that have to be estimated.

The problem of multi-collinearity (i.e., collinearity among the regressors) is a serious concern in dealing with the socio-economic variables. This is because the unbiased assumption of Least Squares estimation does not remain valid if there is a relationship between the error term and the independent variables and/or between the independent variables themselves. This in turn can seriously bias the estimated coefficients, especially when there are unobserved or omitted variables

(Levine and Renelt, 1992; McMahon, 1998; Petrakis and Stamatakis, 2002). This would definitely weaken the validity of the findings, since it would make them sensitive to the specification of the model. But, multicollinearity is frequently both a theoretical problem and problem with a particular sample of data. Economic variables will almost always have some correlation with one another making it necessary to compute correlation coefficients of independent variables. But high correlation coefficients do not necessarily imply multicollinearity.

The central question is whether the multicollinearity is so severe as to alter the estimation of the coefficients. A general rule of detection of multicollinearity is that a high R squared with all low t-score and high simple correlation coefficients among the regressors. Sometimes unexpected sign of the coefficient estimated is also found. It is sometime suggested that do nothing or drop some variables to get rid off from the severe multicollinearity. However several measures have also been suggested by the statisticians to test the severity of multicollinearity. One can make a judgment by checking related statistics, such as tolerance value or variance inflation factor (VIF), Eigenvalue, and condition number.

A very common measure as a rule of thumb has been suggested by Kleinbaum et al. In this rule the severity of multicollinearity is tested by calculating the variance inflation factor (VIF) of each of the regressors. It may be noted here that VIF is defined as $VIF=1/(1-r^2ij)$ where rij is the simple correlation coefficient between two regressors. VIF shows how the variance of an estimator is inflated by the presence of multicollinearity. It is clear from the definition of VIF that as rij approaches 1 (i.e. regressors become collinear), the value of VIF approaches to infinity and if there is no collinearity between the regressors, the value of VIF will be 1. In this vast range, as rule of thumb, it is suggested that if the value of VIF of a regressor exceeds 10, then that variable is said to be highly collinear (Kleinbaum et al 1988). This rule of thumb is followed in the regression analysis to identify the problem of multicollinearity in particular.

Sometimes eigenvalues, condition indices and the condition number is referred to when examining multicollinearity. The condition number (κ) is defined as the largest eigenvalue (λmax) divided by the smallest eigenvalue (λmin), i.e.

$$\kappa = \lambda max / \lambda min.$$

The condition index (CI) equals the square root of the largest eigenvalue (λmax) divided by the smallest eigen value (λmin).

When there is no collinearity at all, the eigen values, condition indices and condition number will all equal one. An informal rule of thumb is that if the condition number is between 100 and 1000 there is moderate to strong multicollinearity and if it exceeds 1000 there is severe multicollinearity. But again, these are just informal rules of thumb. With SPSS, one can compute all these by adding the COLLIN parameter to the Regression command.

Another estimating problem in connection with the Multiple Regression Analysis has been taken care of in the analysis that describes the proportion of the variance of the dependent variable that is explained by the model. In regression terminology it is called goodness of fit. Multiple correlation coefficient is normally used for testing such variability. The multiple correlation coefficient is the correlation between the dependent variable and the linear combination of predictors which minimizes the sum of the squared residuals. Its square ($R^2 = \frac{\text{explained variation}}{\text{total variation}}$) is a very common instrument for measuring the degree of explanatory power in an equation which shows the amount of variance in the dependent variable accounted for by the independent variables. R^2, also called the coefficient of determination, is often described as the proportion of variance "explained", or "described" by regression. The R^2 value for a regression can be made arbitrarily high simply by including more and more predictors in the model. However a better statistical measure is suggested by an alternative coefficient of determination adjusted for the number of independent variables in the regression model. Unlike the coefficient of determination, R^2-adjusted may decrease if variables are entered in the model that does not add significantly to the model fit. Adjusted R^2 is defined as –

$$R^2_{adj} = [1 - \frac{\text{unexplained variation}/(n-k-1)}{\text{total variation}/(n-1)}]$$

Thus the adjusted R^2 is one of several statistics that attempts to compensate for the artificial increase in accuracy. This statistic is applied to measure the goodness of fit in the model.

The coefficients in the regression model have the following simple interpretation:

$$\beta_j = \frac{\partial Y_i}{\partial X_{ij}}$$

This implies that each coefficient measures the impact of the corresponding X on Y keeping all other factors (Xs and u) constant (i.e., the expected change in Y resulting out of per unit change in Xj, assuming all other X's are held constant). In other words, the coefficient of any regressor estimates that the average value of y changes by βj units for each 1 unit increase in Xi holding all other variables constant. Apart from analysis of regression coefficients in the light of statistical significance an attempt has been made to compare the relative magnitude of regression coefficients by analyzing the Partial and Semi-partial Correlation Coefficients of the explanatory variables with the Literacy characters (dependent variable). The relevance of such analysis, beyond the simple interpretation of regression coefficients, is discussed in the result section of the regression analysis.

6.3 Analytical Framework

In order to identify the underlying factors that are responsible for educational deprivation or educational attainment, certain possible exogenous factors may be considered. Researchers have identified different enabling attributes that have impacts on the achievement attributes. In the process of educational development, the enabling attributes may mainly be divided into two categories - household related socio-economic demand side factors and school related supply side factors such as the state of educational infrastructure, staffing and amenities. As such a simple supply-demand framework is adopted to examine the phenomenon of educational deprivation (Venkatanarayana, 2004). Actually the phenomenon of educational deprivation arises owing to the inadequate demand for and/or inadequate supply of schooling. The supply of schooling is a state subject and is not a sufficient condition for increase in the levels of educational attainment. Although the supply side factors are necessary condition for child schooling, it is the socio-economic conditions at the household level that are more crucial in raising the demand for child schooling (Krishnaji, 2000) as such may be considered as an enhancing attribute for literacy attainment too. In a recent paper (Chakraborty, 2006), the schooling variation of children has been well explained by the household characteristics only.

Considering the above, among these two broad categories of explanatory variables, it was decided to proceed with a narrower model. School level supply side variables are not included in the present analysis. In the study, only four villages are surveyed and those are almost equally provided with one pre-primary school, one primary school and one SSK (Table-6.4). There is very little variation in this respect. Moreover, non-participation of school is very common within a village being accessed by equal educational facilities and it is also common that within a family, schooling of children is not homogeneous.

6.4 Socio-economic Characteristics of the Villages under Survey

6.4.1 Basic Amenities

Kantigach is a village (Mouza bearing JL No. 98) under Goalpokhar-I Block of Lodhan Gram Panchayat in the district Uttar Dinajpur. The village is roughly 7 kms from both Bangladesh and Bihar. The language commonly used is a jumble of Hindi and Urdu. Some Bengali words are also habitually used. It is a small village with an area of 74.46 hectares where 169 households were identified out of which 166 households belonged to Muslim community and 03 households were found to be ST section. This establishes the dominance of Muslims in the said village. The land is

mainly used for the agricultural purpose other than residential use. But the land per household ratio is very low. So both the farmers and the agricultural labourers have very limited scope to expand on this primary activity. A tradition of working out side the village for earning money has been a common working pattern that is seen in almost all the households. The infrastructural facilities in the village are very poor. There is a primary school in the adjacent village Uttar Talbari with no upper primary school. One SSK was launched in this village in the year 2004 with two Sahayikas without having any specified area for running it. Recently, a building grant of Rs. 2, 00,000/- (Two Lakhs) has been sanctioned and the land for the purpose has been transferred. As there is no primary school within the village, the need of the SSK is very necessary. Even then the villagers could not come to a unanimous decision for utilizing the fund. As such the fund has been sanctioned to another SSK. This explains the level of un-awareness of the villagers. The whole fact was brought in to the notice of the researchers by the Panchayat member.

Juropani on the other hand is a ST dominated village under Chopra Block of Uttar Dinajpur having 106 households (100% ST) with an area of 201.93 hectares. Among the Schedule Tribes in the village two sub communities are seen. They are Santal and Oraon. The languages they speak are not similar. Sadgi and Oraon are the two main languages used for their intra-society communication. They can also interact in Bengali language with the outer world. Most of the villagers responded in Bengali. Translators also helped in interpreting the responses of others. The mouza has two sub-villages - Juropani and Tulsivita. The first one is at the border line of Bangladesh. It has a semi pucca road and the houses are systematically arranged on both sides of the road. The second one is half a kilometer away from Bangladesh. The village has two medium sized tea gardens and almost all the land has been used for this purpose. The main occupation of the villagers is to work in the tea gardens. It has been found that three households who, apart from working at the tea garden, are engaged in producing a alcoholic liquor which they call Hanria. The village has one government run Primary School with a SSK and ICDS center in the village being connected by electricity for domestic use. As a whole, the village is better equipped with basic amenities compared to the earlier village. The detail of other amenities in the villages is depicted in Table-6 .4 below.

Table-6.4: Basic Amenities to the Study Villages

Amenities	Kantigach	Juropani	Dangipara	Uttar Bhagalpur
Distance from nearest Primary Health Centre (in Kms)	2 Kms	6 Kms	4	7
Distance from nearest bus stop (in Kms)	1.5 Kms	1.5 Kms	0.5 Km	0.5 Km
Distance from nearest Bank/ Credit Society /others (in Kms)	1.5 Kms	2 Kms	1 Km	1 Km
Pre-primary School Facility	One ICDS	One ICDS	One ICDS	One ICDS
Primary School Facility	Nil	Govt= 1, Private =1	Govt= 1, Private =1	Govt= 1
Upper primary School Facility	Nil	Nil	1 High School	Nil

Source: - Field Survey

Dangipara is a village under Goalpokhar-I block of Uttar Dinajpur district with an area of 84.57 hectares and is dominated by the Scheduled Caste population. It is located on the international border with Bangladesh. Though it is a remote village of the district, it has substantial education facilities within it. The village is communicated by a pucca road having a distance of about 10 km from the National Highway No. 34. The nearest town is Kishanganj, a district town of the state Bihar, which is 18 km way from the village. Islampur, a subdivisional town of West Bengal, is at the same distance which is usually used by the villagers. Most of the affluent families have their alternative residence at Islampur, mainly because of better and higher educational access of their

children. The village is moderately literate having 72.5% of its population as literate (Census, 2001). The dependency on agriculture is sufficiently high (83.1% of total worker) in the village (Census, 2001).

Uttar Bhagalpur is another village of the district under Chopra block having its area 138.5 acres which is marginally high from the earlier one. The village is located at just a walking distance from the National Highway No. 34 and as such it has a better communication facility. Work diversification is seen in this village. Kuti or procuring rice from paddy is a major source of income of the households. The villagers bring the paddy from nearest paddy producing area (Kanki, Suryakamal of Goalpokhar-II block) and process the same into rice. In this process of work, the female members are also engaged. But their remuneration is not given to them as the households treat it like other household works. Some households are seen to be engaged in self employed small business. In a marginal portion of land, tea gardens in small scale (within 2 to 5 acre) are also seen in the village.

The infrastructural and educational facilities that are available to the study villages is presented in a tabular form in Table- 6.4. It is seen that so far as access to education is concerned, all the villages are equally provided with one pre-primary and one primary school within them except Kantigach. In Kuntigach there is no primary school but one unrecognized Madrasah has been found. There is one private run primary school in addition to one government primary school in Juropani and Dangipara. It is seen from Table-6.4 that the bus connectivity with the villages are also very similar along with another important amenity - banking facility. However, as a whole, the Dangipara village may be said to have better educational and other infrastructural facilities compared to the other villages.

6.4.2 Population Profile

As it has earlier been stated that 30 households from each village have been surveyed in detail and as such a total of 120 households in 4 villages have been covered. The detail population profile of the villages is shown in Table-6.5.

The average household size of the sample villages is 6.1 and it ranges from 5.4 in Uttar Bhagalpur to 6.6 in Dangipara. Child population in the sample villages is found to be 13.5% of the total population. Village-wise, it is highest in Kantigach (19.3%) followed by Juropani (14.5), Uttar Bhagalpur and Dangipara. The child population seems to have some correlation with the literacy rate of the villages. The proportion of aged members (aged 60 years and above) is found to be lowest in Dangipara (6.0%) followed by Kantigach (7.7%), Uttar Bhagalpur (9.3%) and Juropani (10.8%). The noticeable thing is that the sex ratio of the sample households is noticeably low irrespective of their educational level.

Table-6.5: Population Distribution of the sample Villages

Name of the Village	Population		Sex Ratio	Household Size	No of Siblings (0 – 5 Years age)		Old-age (> 60 yrs) Members	
	M	F			M	F	M	F
Kantigach	102	79	775	6.0	17	18	5	9
Juropani	94	92	979	6.2	15	12	10	10
Dangipara	106	93	877	6.6	7	9	7	5
Uttar Bhagalpur	87	75	862	5.4	11	9	9	6
Total of 4 Villages	389	339	871	6.1	50	48	31	30

Source: Field Survey

6.4.3 Literacy Profile of the Sample Population in the Concerned Villages

Literacy rate can be considered as a crucial measure of human capital on the ground that literate people can be trained less expensively than illiterate people and the literate people also generally have a higher socio-economic status and enjoy better health and employment prospects. Policy

makers also argue that literacy increases job opportunities and access to higher education. The UN Millennium Development Goals (MDGs) use literacy data as an indicator in the goal to achieve universal primary education and for promoting gender equality as it has been seen to improve literacy which has high correlation with poverty reduction. UNESCO has declared this coming decade as the UN Literacy Decade in order to stress the importance of literacy. It is also claimed that literacy, besides being a fundamental human right, is a foundation not only for achieving EFA but, more broadly, for reaching the overarching goal of reducing human poverty (UNESCO, 2005 UNESCO (2005); 'EFA Global Monitoring Report 2006' United Nations Educational, Scientific and Cultural Organization, 7, Place de Fontenoy, 75352 Paris 07 SP, France). It is now widely being recognised that "literacy skills are fundamental to informed decision-making, personal empowerment, active and passive participation in local and global social community" (Stromquist, 2005, p. 12)

Such correlation, however, depends upon the definition of literacy or the dimension from which literacy is being defined. The traditional definition considers literacy as an ability to "read, write, spell, listen, and speak. A basic literacy standard in many societies is the ability to read the newspaper. Some have argued that the definition of literacy should be expanded. For example, in Scotland, literacy has been defined as: "The ability to read and write and use numeracy, to handle information, to express ideas and opinions, to make decisions and solve problems, as family members, workers, citizens and lifelong learners." In June 2003 the UNESCO organized a meeting in Paris. Literacy experts from different parts of the world were invited to formulate a definition of literacy that would guide developments in literacy assessment and ensure that it addressed appropriate country issues. The following definition was proposed:

"Literacy is the ability to identify, understand, interpret, create, communicate and compute, using printed and written materials associated with varying contexts. Literacy involves a continuum of learning in enabling individuals to achieve his or her goals, develop his or her knowledge and potentials, and participate fully in the community and wider society."

It is especially after the 1990s, when the internet came into wide use, it has been asserted that the definition of literacy should include the ability to use tools such as web browsers, word processing programs, and text messages. This expanded skill of education has been termed computer literacy. Apart from this, some other types of skills are also advocated while defining literacy in question. These are health literacy, cultural literacy, financial literacy, numeracy, racial literacy, scientific literacy, statistical literacy, visual literacy etc.

However, while dealing with the problem of illiteracy in the Indian context the study will confine to the definition as given by the Indian Census, the ultimate source of literacy data in the country. The Indian Census defines a person as literate if he or she has the ability to read and write with understanding in any language, such as Hindi, Kannada, Marathi, Tamil or any other official Indian languages. But it is not always very easy to measure one's ability of reading with understanding. The census enumerator normally asks the respondent of a household (normally the male head of household) whether each household member is literate. A positive response is given if the individual has had any schooling at all. In India, the census enumerator is instructed, when there is doubt, to see whether a person listed as literate can read any part of the enumerator's manual and write a simple sentence. The census also requires to record details about employment, family size, housing, and many other topics. While conducting the survey, this definition was kept in view and care was taken to find out the literacy skill of the family members as accurately as was possible. With the information gathered from the survey, Table 6.6 has been constructed to reflect the literacy scenario of the study villages.

A simple interpretation of Table-6.6 reveals some important features. Firstly, there has been a remarkable progress in the overall literacy rate of the villages since 2001. A highest of 35 per cent literacy jump has been observed in Kantigach and the lowest is in Uttar Bhagalpur (9.5%). The same is 16% for Juropani and around 13% in Dangipara. It may be noted here that the survey was carried out during the year 2006-07. The literacy jump thus seems to be quite interesting. This simply shows that the growth in literacy rate is more sensitive in the low literate villages.

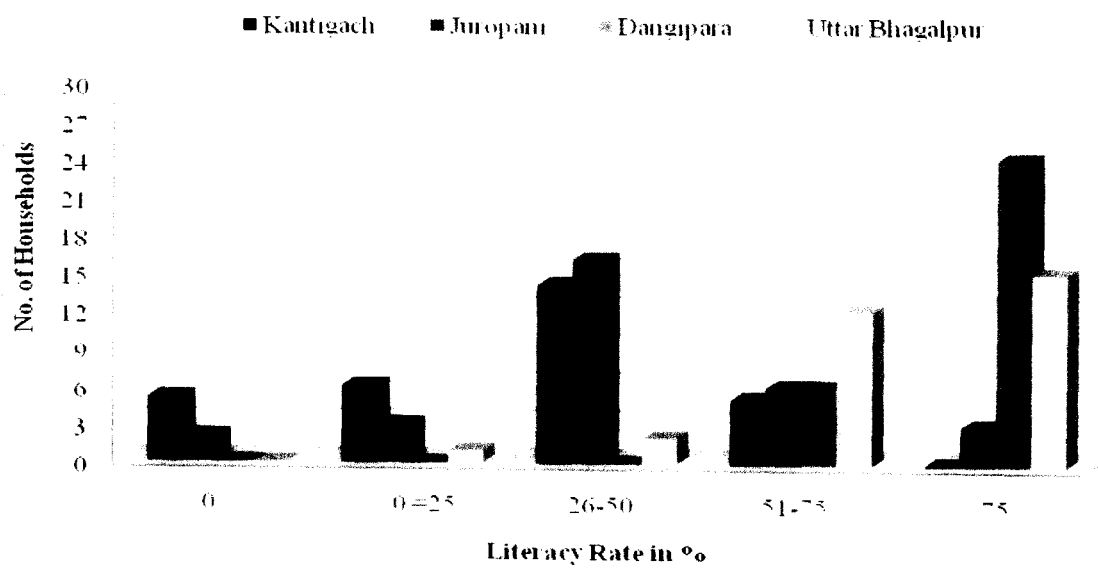
Secondly, in all the villages (except Kantigach where both male and female literacy rates were almost negligible in 2001 Census) the literacy increase has been higher for the females than for the males thereby showing a trend of narrowing the gender gap in literacy rate in the near future. Thirdly, the gender gap in literacy rate is found to be comparatively higher in the low literate villages. This suggests that the higher the overall literacy rate of any area, the lower will be the gender gap. Finally, the overall literacy rate of the sample villages in total is 62.7%, which is comparable to the state average rural literacy rate (63.42 as per Census 2001).

Table-6.6: Literacy Profile of the Villages

Name of the Village	Population		6^+ age Population		6^+ age Literate		6^+ age Literacy Rate		
	Male	Female	Male	Female	Male	Female	Male	Female	Total
Kantigach	102	79	71	56	31	17	43.7 (3.5)	30.4 (0.98)	37.8 (02.2)
Juropani	94	92	75	73	40	27	53.3 (41.3)	37.0 (17.8)	45.0 (29.2)
Dangipara	106	93	92	77	84	60	91.3 (83.9)	77.9 (58.8)	85.2 (72.5)
Uttar Bhagalpur	87	75	73	69	58	51	79.5 (78.6)	73.9 (54.2)	76.8 (67.3)
Total of 4 Villages	389	339	311	275	213	155	68.5	56.4	62.7

Source: - Field Survey (figure in the parenthesis represents the literacy rate 2001)

Fig-6.1: Distribution of Households as per the Literacy Rate



Source: Field Survey

Apart from village level variation, as shown in Table-6.6, it is also noticed that there is a considerable variation in literacy achievement among the households too. A brief sketch of household level literacy character is now being presented in Figure-6.1.

The households have been categorized as per the literacy rate and five categories have emerged from the survey data that reflect upon the educational level of the households. So far the literacy attainment of the households in the sample villages is concerned, it is the Kantigach village where every household has been found with a literacy rate below 75%. Again, 40% of the total number of households (11 out of 30) has hardly acquired a minimum of 25% literacy rate. Half of the family members are found to be illiterate in case of more than 80% households (25 out of 30) in this particular village. On the other hand, the households of village Dangipara from the same block

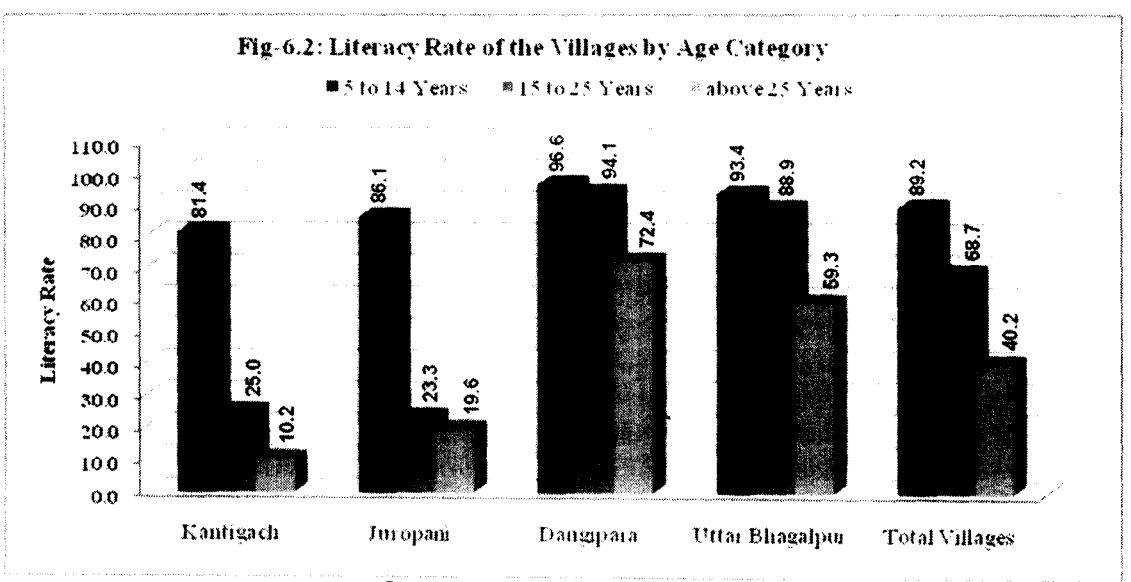
(Goalpokhar-I) show far more developed in this connection. All the households in this village at least maintain a literacy rate of more than 50%. Again, 80% of the total sample households show a literacy rate more than 75% in this village. Juropani, a tribal village, is also a comparative low literate village where 75% of the households are lying within a literacy range of below 50% level.

Beyond this simple interpretation of literacy rate, age-wise literacy character is also constructed for disaggregated analysis. Out of total literate members in all villages (407), 54.8% (223) belongs to the category of 5 to 14 years age group, 19.4% (69) in 15-25 years age group and the remaining 25.8% (115) in 25 years and above. The literacy rate among the family members aged 25 years and above (expected parental group) is found to be the lowest in Kantigach (10.2%) followed by Juropani (19.6%), Uttar Bhagalpur (59.3%) and it is the highest in Dangipara (72.4%). The above provides the overall educational development pattern of the villages too. Further analysis of literacy variation has been undertaken in a later chapter.

6.5 Relevant Data & Variables of the Econometric Application

6.5.1 Household Income/Expenditures

Household income, as it has been pointed out, is an important economic variable that bears strong statistical relationship with educational attainment. But at household level it is rather difficult to assess the family income in the absence of accurate information. It has been observed during the survey that there is a huge discrepancy between income earned (Y), consumption (C) and savings (S) of a particular household. Three types of information were asked - total earnings, total expenditure and total savings. An interesting observation is that total expenditure and total savings altogether was much higher than the total earnings for most of the households. Even for some households, expenditure is much higher than the level of income although the families have not taken any loans. This indicates that that there is either a tendency for the respondent to provide inaccurate information or that they are unable to calculate their earnings on a monthly basis. Thus an alternative process was adopted to capture the issue. For this only the expenditure level of the households was retained for the sake of analysis. Though it is not possible to include all the parts of expenses made by a household, five major expenditures were built-in that were commonly incurred by all the families. They are - expenses (annual) on food, clothing, housing, education and other consumer goods. Owing to the irregular nature authenticity response on health expenditure appears to be doubtful. As such, health expenditure, although an important part, has not been included. On the basis of information on the annual expenditure, monthly per capita expenditure is calculated. The summary of the data is shown in Figure- 6.2 below.



Source: - calculated from Field Survey

Taking expenditure of the households as a surrogate of income, it is seen that in all the villages, expenditure on food is a major share of total expenditure. In the low literate villages (Kantigach

and Juropani) more than 75% of total expenditure is accounted for by food expenditure. In Dangipara, it is the lowest followed by Uttar Bhagalpur and Juropani. This high proportion of food expenditure in total expenditure has an important implication. It is generally observed that lower the level of income, the higher will be the proportion of expenditure on food. As such it seems that, the village Kantigach suffers from acute poverty (monthly per capita expenditure is found to be the lowest at Rs.333, Table-6.7) although the other villages also have the problem of persistent poverty too. Except Dangipara, in all the villages, it is the expenditure on clothing that occupies the second largest share in total expenditure. In Dangipara, educational expenditure gets more prominence, standing next to food expenditure. The above information regarding expenditure of households is a pointer to the fact that the households in most villages have to earn for their basic needs of food and clothing, thus making education of a child an option and not a necessity.

For a detailed understanding of expenditure of the households, the present exercise has calculated an induced expenditure variable apart from traditional monthly per capita expenditure (MPCE). This is educational expenditure as proportion of total expenditure and per capita educational expenditure which is shown in Table-6.7. It is seen from the table that all the induced expenditure variables bear strong relationships with the overall literacy rate at village level. However, all expenditure variables are strongly correlated with each other. Accordingly, introduction of all these independent variables may precipitate the problem of multicollinearity in the regression model. For this, the conventional method of checking the multicolinearity problem will be applied.

Table-6.7: Village Level Expenditure

Name of the Villages	Total	Total Population	Total Literacy Rate	Monthly Per capita Expenditure	Education as % of Total Expenditure	Per capita Educational Expenditure
Kantigach	722440	181	37.8	333	4.72	775
Juropani	865775	186	45.0	388	6.73	987
Dangipara	1027695	199	85.2	430	8.08	1483
Uttar Bhagalpur	767210	162	76.8	395	7.36	1065

Source: - calculated from Field Survey

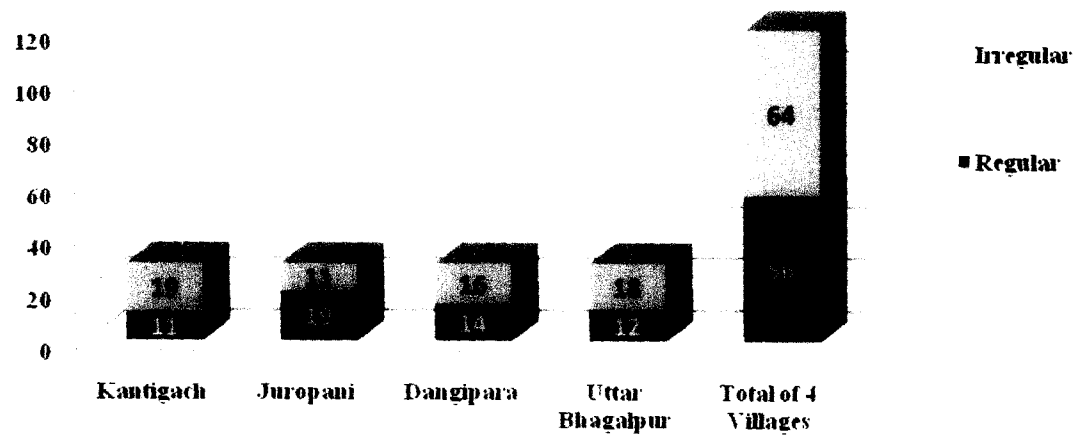
6.5.2 Occupational Pattern and Literacy Attainment

A common classification of occupational pattern is to distinguish between earners as agriculture labourer, cultivator, household industry worker and other worker (Census classificaion). This can be broadly divided into agricultural households and non-agricultural households. In the present study differential occupational pattern is also captured with classification of the households into two new categories - those who have regular/permanent income and the other with irregular income, irrespective of the total income earned by the households.

In order to differentiate between regular and irregular nature and source of income, assessment of households into two broad categories depending upon their nature of work has been undertaken. Firstly, the self employed workers, salaried earners, working in unorganized sectors or in any other types of work that generate a steady flow of income and, secondly, the workers having uncertainty so far as flow of income is concerned, i.e. those who have to rely on saonal income flow. In the latter category, one usually finds the small cultivators, agricultural labourers, semi-skilled labourers. Thus, occupational pattern as reflected by patterns of flow of income help in distinguishing between households which have the security of a steady income from those who do not. The logic behind such exercise is that a household earning income on a regular basis will enjoy economic security along with a guarantee of a future income, which in turn makes decision making and planning for the future less complicated from those who do not have a regular source of income. Education of the children is thus included in the decision making and planning and it has an important bearing on the decision to send the wards to school.

Briefly, out of 120 households in four villages, 56 have been found to earn a regular income and the remaining 64 households are subject to fluctuations or irregularities in income earning. In Juropani, in spite of having lower MPCE, more than 60 percent of the households (19 out of 30) earn income on a regular basis. It has earlier been noted that most of the earners of this village are engaged as tea garden workers getting their wages either monthly or on weekly basis with also having access to the Provident Fund, rations, medical facilities, etc. On the other hand, in Kantigach, the workers frequently migrated to other districts and states for earning and thus being subjected to uncertainty in income earning and often having to face deception from employers. Certainty in getting employed in a distant region is also dependent on several factors like networking, information base, security, connectivity, etc., thus making the process very time dependent. In the other two villages of Dangipara and Uttar Bhagalpur, there is a blend of regular and irregular income holders with the latter being higher than the former in both the villages. The findings are provided in Figure 6.3.

Fig-6.3: Occupational Pattern of the Households



Source: Field Survey

6.5.3 Opportunity Cost of Sending the Children to School

In studying the schooling characteristics of the children, it appears that where income poverty is more acute and where there is opportunity for a child to be engaged in paid work, the parents were under pressure to choose one option from a set of two - either to send the child to school or to send the to work to supplement family income. Considering the income poverty of the families, the propensity to exercise the second option is expected to be stronger and more realistic for the household to alleviate income poverty. This appears to generate a two-way relation in the following manner-

→

INCOME POVERTY EDUCATIONAL POVERTY

←

A strong debate exists in arriving at such a conclusion, i.e. economic security/income will always have positive educational or schooling outcomes. It is common in the literature on education where one can find higher literacy rate among relatively lower income groups. As such, in the study, to overcome such issues, opportunity cost of sending a child (below18 years age) to school has been introduced in the model. This has been done as follows-

The children of the family with irregular income nature usually need to assist their parents for earning. In the process they either assist their parents or they are directly engaged in paid work. With this assumption the collected data from each household include the work status of each family member. The information on work status of the children aged up to 18 years provides information on whether there is scope for children to get access to paid work in the labour market

or not. If it is in the affirmative, it is assumed that there is some opportunity cost of sending children to the school. The implicit opportunity cost through work status has been represented in Table-6.8.

Table-6.8: Work Status of Children (Up to the age 18 years)

Name of the Village	No. of households		
	Total	Children in Paid Work	Children not in Paid Work
Kantigach	30	13	17
Juropani	30	16	14
Dangipara	30	8	22
Uttar Bhagalpur	30	10	20
Total of 4 Villages	120	47	73

Source: Field Survey

It is observed that out of total 120 households, children of 47 households were found to be engaged in paid work. The extent of children in paid work is highest in Juropani (tea garden dominated tribal area) followed by Kantigach and Uttar Bhagalpur. The opportunity cost of sending the children to school may additionally capture the character of income poverty with the prevalence of child labour with a priori assumption that where there is work opportunity for children, lower will be the school enrolment rate of the children. The households are assigned the value "1" if at least 1 child is found in the labour market and "0" other wise, thus assigning it as a qualitative dummy variable with binary values.

6.5.4 Dependency Ratio

Among the family members, children belonging to the age group of 0-5 years and the elderly population of 60 years and above are economically dependent in the sense that they are non-earning members in the household. The presence of this group of family members creates two types of dependencies. As they are non-earners, they are financially dependent and their financial burden is usually borne by the adult earners. This may be termed as economic dependency of a household that may have some effect on child schooling and literacy rate of the family. Again, physically, they are also dependent (except some elderly members) on other able-bodied members of the household since the latter is expected to take care of the sick and elderly in the family. This physical dependency may again be termed as household dependency which is a non-financial burden. This non-financial burden is usually shouldered by other non-earners, primarily by the school going age children or by the female members of the family unit. It may generally be assumed that larger the dependency (both economic and household dependency) of a family, lower will be the chance of a child to be schooled. Considering this assumption, the extent of both types of dependency ratio has been incorporated as an additional explanatory variable in the regression exercise. A brief picture of the dependency character as obtained from the sample households is depicted in Table-6.9 below.

Table-6.9: Village Level Dependency Ratio

Name of the Villages	Schooling age 5-14 Children	Dependent member	Total Earners	Household dependency Ratio	Economic Dependency Ratio
Kantigach	59	48	42	0.81	1.14
Juropani	72	47	72	0.65	0.65
Dangipara	58	28	55	0.48	0.51
Uttar Bhagalpur	61	35	52	0.57	0.67
Total of 4 Villages	250	158	221	0.63	0.71

Source: - calculated from Field Survey

It may be noted here that household dependency ratio is calculated as ratio of dependent members to schooling age children following the above mentioned argument. The economic dependency has been estimated by the ratio between the dependent members to earners in the family. It is seen that both household and economic dependency ratio are the highest in the least literate village of Kantigach followed by the higher literate village. The ratios have an explanatory capacity in capturing the variations in literacy character of the households, thus amply justifying the inclusion of this variable in the regression analysis.

6.5.5 Role of Female Members in the Household

Mothers' role is undeniably significant in determining the educational status in a family. In order to elucidate on this crucial role, female Work Participation Rate (FWPR) has been considered in the model as an explanatory variable acting on the premise that it will negatively impact upon the educational outcomes of the children in the household. Taking FWPR for all the villages under survey, around 19% of the female members were found to be working. A very low rate of FWP was found in the village of Kantigach where the literacy level is also very low and the same is reflected in the other villages too. This however is not in line with economic arguments present in literature where low literacy of women is associated with high WPR especially in informal sectors. However, the figure is for all the villages which may or may not be supported at household level. The regression coefficients will provide more robust results. In the regression model, FWPR has been incorporated as a qualitative dummy variable assigning the value '1' if any female member is working and '0' otherwise.

Table-6.10: Work Participation

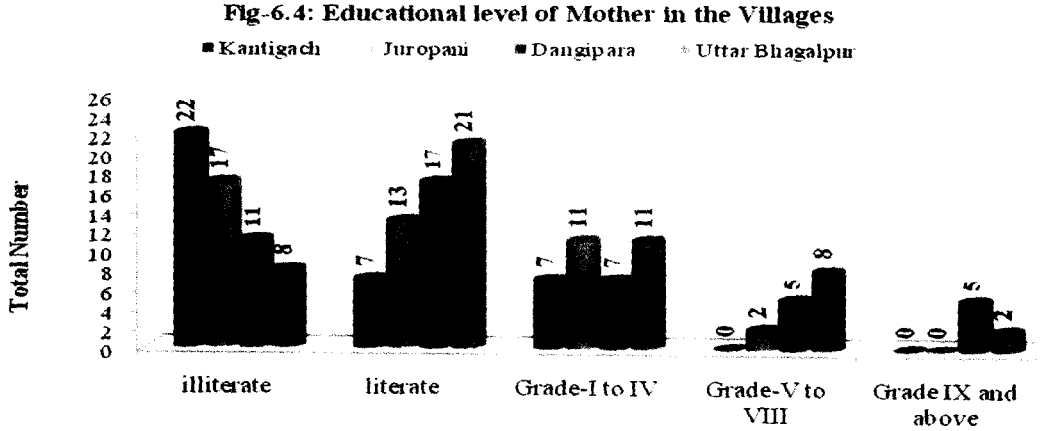
Name of the Villages	No of Earning members			WPRF
	M	F	T	
Kantigach	40	2	42	2.53
Juropani	35	37	72	40.22
Dangipara	47	8	55	8.60
Uttar Bhagalpur	35	17	52	22.67
Total of 4 Villages	157	64	221	18.88

Source: - calculated from Field Survey

6.5.6 Parental Educational Level and Schooling of Children

Studies on educational status of the parents show that mothers' educational level has a positive relationship with the households' educational attainment and accordingly this has also been introduced in the present schema of analysis. It has been found from the sample data that out of total 120 households there were 116 mothers whose level of education had been recorded (Fig-6.4).

Fig-6.4: Educational level of Mother in the Villages



Source: - calculated from Field Survey

Out of this, 58 were found to be literate of which 36 were literate with below primary level of education, 15 had received education till upper primary and the remaining 7 mothers were found to have moved beyond upper primary level of education. Educational level of mothers in Kantigach village is most appalling with only 7 mothers (out of 29) who are literate with below primary level of education.

Over the past few decades the empowerment of women has been an important issue that has been adopted not only in academic research but also in policy matters. Among all the factors that can empower women, education is understood to be the all important factor that can enhance the capability of the person thereby empowering the person. To quantify the level of empowerment, an Index of Parental Empowerment as suggested by Chakrabarti and Sharma Biswas, (2008) has been adopted to serve the present purpose. The index takes into account whether decisionwithin the household is taken by the respondent (mother) herself or jointly with husband or other family members (Yes = 1; No = 0). This index has been constructed for two dimensions - index for household matters (e.g.cooking, health, allowed to have money) and index for freedom to go outside the house (e.g. freedom to go to the market, visiting relative/friend's house, right to spend money, right to purchase jewelry, decision to stay at parental house). Following this, this index has been introduced as a variable in a modified form. For this, a simple query covering 4 aspects of decision making had been included under the principle query. The four specific issues with three pre-defined possible responses were – who decide/s in the event of enrolment of children, continuation of child education, economic matters and matrimonial and other socio cultural affairs. Although there are other related issues on which decisions usually taken, but for capturing the educational purpose, only these four specific issues have been covered. The possible three responses were - in some issues father alone, in some issues mother alone and some decisions may be taken jointly by both parents. Each parent is allotted 01 mark for taking part in decision making sphere. As such, a parent (father/mother) may be assigned a maximum of 4 marks if he/she takes part in all the familial issues and in each village 120 is the maximum level of score (Full Score; if all the 30 parents are alive) that the fathers or the mothers of that particular village can obtain. Finally, on the basis of the score/marks obtained by each parent in a household, an index assigned to each parent of the households has been calculated as follows –

$$\textbf{Parental Empowerment Index (PEI)} = \frac{\textbf{(Marks or score obtained by the parent)}}{\textbf{Total Score}} \times 100$$

This index has been termed as Parental Empowerment Index (PEI). Accordingly, if a parent takes part in all the four specified decision-making spheres, then his/her empowerment index will be 100%. Finally, after calculating the value of the index for both of the parents in a particular household, it is again observed whether the value of the index of mother is greater or equal to that of father. In that case, it is obvious that the mothers are enjoying equal or even more decisive power in the family and accordingly it may be said that the mothers are empowered in that very particular family. In technical terms, if mothers are empowered then they are assigned with yes or '1' value and '0' if not.

The value of the index for the four villages is shown in Table-6.11. Overall, the decisive power of the fathers is 86.3%, i.e., at the event of taking familial decision, in 86.3% issues, the fathers are taking active part while the same is 76.5% for the mothers. It is observed that the PEI varies across the villages. Interestingly it is found that the PEI associated with the mothers has a unique positive correlation with the overall literacy rate of the villages. In Kantigach, the mothers are less empowered than all the villages where the literacy rate has also been found to be lower than the other three villages under study. The only exception found here is that in Juropani while mothers were highly empowered, the literacy rate does not reflect the level of empowerment of the women in the family. In Juropani, the mothers play an important role in familial decision-making which may be because of the culture of a tribal dominated village.

Noticeably it is seen that in economic matter, the fathers' role is more dominating than the other issues. Considering such an interesting association of this PEI with literacy attainment, it has been decided to incorporate the index as an additional explanatory variable to see the role of women empowerment in educational attainment.

Table-6.11: Construction of Parental Empowerment Index

Area of Empowerment	Kantigach		Juropani		Dangipara		Uttar Bhagalpur		All Villages	
	Father	Mother	Father	Mother	Father	Mother	Father	Mother	Father	Mother
Enrolment of children	23	14	23	27	25	28	27	21	98	90
Continuation of education	22	14	25	25	23	28	27	20	97	87
Economic matter	28	6	25	27	27	27	30	21	110	81
Matrimonial/other socio-cultural purpose	27	15	25	29	26	29	28	24	106	97
Total Score of the village	100	49	98	108	101	112	112	86	411	355
Full Score	120	116	116	120	120	112	120	116	476	464
Index of Empowerment	83.33	42.24	84.48	90.00	84.17	100.00	93.33	74.14	86.34	76.51

Source: - calculated from Field Survey

6.6 Factors Explaining Literacy Rate

In order to identify the factors influencing the literacy levels a set of explanatory variables (demand side) are incorporated in the multiple regression analysis where households are taken as unit of analysis.

A regression analysis using PLR, MLR and FLR as the outcome variables representing the literacy rates of person, male and female has been performed. The right hand side variables in the equations are predictors, the definitions of which along with the other details have already been discussed in the earlier section of this chapter. Theoretically we expect that higher literacy attainment would be associated with higher proportion of educational expenditures (EDNTOTEX), better educated parents (FTHEDN & MTHEDN), households with having more regular income (INREGIRREG) and the families where the mothers enjoy more or at least equal level of empowerment as compare to the fathers (EMPMTH) . On contrary, a negative association is expected with the households where economic dependency of the earner (ECONDEP) and household dependency of the children found to be higher and where the mothers ought to work in the labour market (FWPR). Finally, a priori assumption is also made that the literacy attainment of the households will be lower if there is a scope of the children to be engaged in labour market (OPPTNTCOST). It is also assumed that literacy rate is a linear function of all the predictors and considering the above theoretical background, three separate regression equations have been constructed for analysis in the following way

Regression Equation

$$PLR_i = \beta_0 + \beta_1 EDNTOTEX_i + \beta_2 FWPR_i + \beta_3 ECONDEP_i + \beta_4 HHDEP_i + \beta_5 INREGIRREG_i + \beta_6 OPTNTCOST_i + \beta_7 FTHEDN_i + \beta_8 MTHEDN_i + \beta_9 EMPMTT_i + \epsilon_i - -(1)$$

$$MLR_i = \beta_0 + \beta_1 EDNTOTEX_i + \beta_2 FWPR_i + \beta_3 ECONDEP_i + \beta_4 HHDEP_i + \beta_5 INREGIRREG_i + \beta_6 OPTNTCOST_i + \beta_7 FTHEDN_i + \beta_8 MTHEDN_i + \beta_9 EMPMTT_i + \epsilon_i - -(2)$$

$$FLR_i = \beta_0 + \beta_1 EDNTOTEX_i + \beta_2 FWPR_i + \beta_3 ECONDEP_i + \beta_4 HHDEP_i + \beta_5 INREGIRREG_i + \beta_6 OPTNTCOST_i + \beta_7 FTHEDN_i + \beta_8 MTHEDN_i + \beta_9 EMPMTT_i + \epsilon_i - -(3)$$

Here, i=1, 2, 3.....120 and ε_1 is the error term which accommodate all the other influences on the outcome variable not accounted for in the assumed model. This error term is assumed to be a random variable εwith a mean of zero, conditional on the explanatory variables with the property that the variance of the error is the same at any level of explanatory variables, i.e., E (u 1i) ~ N (0, σ2). $\beta_0, \beta_1, - - - - - - \beta_9$ are unknown parameters which are estimated by "fitting" the equation to the data using least-squares. The variables used for regression analysis are listed below (Table-6.12):

Table-6.12: Description of Variables

Variables	Notation	Description
Dependent variables		
Literacy Rate (Person)	PLR	Percentage of total population literate in the age group 7 years and above
Literacy Rate (Male)	MLR	Percentage of male population literate in the age group 7 years and above
Literacy Rate (Female)	FLR	Percentage of female population literate in the age group 7 years and above
Independent covariates		
Economic dependency ratio	ECONDEP	Ratio of non-earners (old-ages and siblings) to earners
Household dependency ratio	HHDEP	Ratio of non-earners (old-ages and siblings) to schooling age children
Proportion of educational expenditure	EDNTOTEX	Educational expenditure as a % of total expenditure
Mothers' empowerment	EMPMTH	1= if mothers' empowerment index is greater than or equal to that of father's, 0 = less than that of father's
Fathers' education level	FTHEDN	education in completed number of years
Female work participation	FWPR	1= if female members work, 0= no female works
Pattern of income	INREGIRREG	1= if the main income of a household is regular, 0= if the main income is not regular
Mother's education level	MTHEDN	education in completed number of years
Opportunity cost of schooling	OPTNTCOST	1= if there is any children in the age group below 18 years and working, 0= if there is no children in the age group below 18 years and working

It is expected that as proportion of educational expenditure to total expenditure (EDNTOTEX), monthly per capita expenditure of the household (MPCE), educational level of father and mother (FTHEDN, MTHEDN) increase, the literacy rate will also have an increasing trend and households with regular income (INREGIRREG) and with mothers' empowerment will also show the increase in literacy rate. However it is assumed that increasing economic dependency representing the ratio of non-earners to earners (ECONDEP) and household dependency measured as the ratio of old age and sibling members to total number of schooling age children (HHDEP) will lower the literacy rate. it is also assumed that the literacy rate will be lower for the households where the female members work (FWP) in the paid market and where there is an opportunity for the children to be engaged as child labour (OPTNTCOST).

6.6.1 Regression Analysis

It has earlier been stated that multiple regression analysis as a method of estimation is being applied in the present study. However, there are different forms of regression methods such as Standard multiple regression, Hierarchical multiple regression and Stepwise multiple regression. Standard multiple regression is used to evaluate the relationships between a set of independent variables (metric or dichotomous) and a dependent variable (metric). Hierarchical, or sequential, regression is used to examine the relationships between a set of independent variables and a

dependent variable, after controlling for the effects of some other independent variables on the dependent variable. Stepwise regression is used to identify the subset of independent variables that has the strongest relationship to a dependent variable. So far the present analysis is concerned, standard multiple regression is resorted to where all the independent variables are entered into the regression equation at a time. Here, beyond normal goodness of fit test (R^2 Adjusted) and statistical significance test ('t' test), an 'F' test is used to determine if the relationship can be generalized to the population represented by the sample (as because sampling method has been applied in the process of data collection).

6.6.2 Initial Observations on the Multiple Regression Result

In order to test the viability of sample size, the minimum ratio of valid observation to independent variables for multiple regression is normally suggested 5 to 1. In our model with 120 valid cases and 09 independent variables, the ratio for this analysis has become 13 to 1, which is more than the minimum requirement. Another rule of thumb has been suggested in the manner that when there is a number of cases (N) relative to the number of predictor variables (k), N should be > 50 + 8k when testing R2 and N > 104 + k when testing individual Bj values (Green, 1991).

Thus it appears that the selection of sample size and independent variables holds satisfied so far the statistical requirement is concerned.

Table-6.13: Descriptive Statistics

Variables	Mean	Std. Deviation	N
PLR	62.61	30.25	120
MLR	69.05	36.69	120
FLR	56.17	40.95	120
EDNTOTEX	6.82	4.69	120
MPCE	392.80	124.49	120
FWPR	0.47	0.50	120
ECONDEP	0.89	0.98	120
HHDEP	0.70	0.69	120
INREGIRREG	0.47	0.50	120
OPTNTCOST	0.35	0.48	120
FTHEDN	3.47	4.01	120
MTHEDN	2.08	2.86	120

Multiple correlation coefficient- Goodness of fit

Before reporting the regression results which usually contain the estimates of the parameters and their standard errors or t-statistics, it has been tried to gather some information that tells how closely the regression line fits the survey data. Putting it in other words, how much variation in the response variable is being explained by the independent variables that are included in the analysis. It is earlier stated that R Square or Adjusted R Square is a statistic that measures the above proportion in any multiple regression analysis.

Three literacy variables have been regressed upon in the model and as such multiple correlation coefficient for each of the three equations have been found and shown in detail in Table 6.14. Adjusted R Square is about 0.553 for overall literacy rate (PLR) which tells us that 55 % variation in literacy rate (41% and 38% in female and male literacy rate) is accounted for by the predictors included in our model. This apparently indicates that variation in literacy rate is not much explained by the household level socio economic factors (predictors in the model). But it should be noted here that cross sectional data has been used in the analysis. In such a case individual differences are caused by many factors that often cannot be measured. Moreover, so far the number of observations of the analysis (120 observations in total) and limitation of not including the institutional factors in the model are concerned, it appears that the exercise proved to be a good model for explaining the literacy rate. However a better explanation in this respect may be found

from the analysis of variance given by the F- ratio which incorporates sample size and number of predictors in assessing the significance of the relationship. It actually measures whether all the independent variables, taken as a group, (linearly) influence the dependent variable or not. That is, it tests the null hypothes ($H_0=\beta_1 = \beta_2 = \beta_3 = - - - \beta_k = 0$) against the alternative hypothesis, ($\beta_1 = \beta_2 = \beta_3 = - - - \beta_k \neq 0$) that at least one of the coefficients is not zero. If the H_0 is true, clearly the model does not explain the behavior of dependent variable and the regression model then reduces to-

$$\text{Literacy Rate} = \beta_0 + \epsilon_i$$

Table-6.14: Multiple correlation coefficient/ Goodness of fit

R	R Square	Adjusted R Square	Std. Error of the Estimate	Dependent Variable
0.766	0.586	0.553	20.238	PLR
0.656	0.431	0.384	28.792	MLR
0.674	0.454	0.409	31.481	FLR

The result of ANOVA as found for the three regression equation is briefly summarized in table-6.15. From the analysis of variance (ANOVA) it is found that the probability of the F statistic in all the three equations estimated is less than 0.001 which is again less than or equal to the level of significance of 0.05. Thus it appears that a statistically significant F- ratio as found for the present analysis indicates that the model is quite good. Thus, the results support the research hypothesis that there is a statistically significant relationship between the set of independent variables included in the model with that of the dependent variables, thereby rejecting the null hypothesis.

Finally, before interpreting the regression coefficients, check on the presence of collinearity among the regressors has been done. In order to find out the actual effect of a particular independent variable, multiple regression analysis allows for control over the effects of other independent variables on the dependent variable. Thus it is essential in multiple regression analysis that the key explanatory variable not be correlated perfectly with one or more of the other explanatory variables. If there is to be perfect correlation between two independent variables, it would not be possible to separate out the effect of the key variable on the dependent variable from the effect of the other independent variable/s.

Table-6.15: ANOVA

	Sum of Squares	df	Mean Square	F	Sig.	Dependent Variable
Regression	63872.367	9	7096.930	17.328	0.000(a)	PLR
Residual	45051.867	110	409.562			
Regression	69007.047	9	7667.450	9.249	0.000(a)	MLR
Residual	91187.302	110	828.975			
Regression	90552.604	9	10061.400	10.153	0.000(a)	FLR
Residual	109012.705	110	991.025			

a Predictors: (Constant), MTHEDN, HHDEP, FWPR, OPTNTCOST, INREGIRREG, MPCE, EDNTOTEX, FTHEDN, ECONDEP

It is equally important that suppose the explanatory variables (e.g. in the model EDNTOTEX and other independent variables) are not correlated, then including one particular explanatory variable (e.g. EDNTOTEX) in the model would not have an effect on estimated coefficients of the other

independent variables (e.g. EMPMTH, MTHEDN, HHDEP, INREGIRREG, FWPR etc.) in explaining the dependent variable (e.g. literacy rate in our model). In other words, if the independent variables are not correlated, then there is no need to perform multiple regression. Bivariate analysis may capture the true effect of a particular variable. In practice, economic variables will almost always have some correlation with one another. When two or more variables are highly, but not perfectly, correlated - that is, when there is multicollinearity - the regression can be estimated, but some concerns remain. The basic task is to see whether the correlation among the predictors is tolerable. For this one may compute correlation coefficients of independent variables. But high correlation coefficients do not always necessarily imply multicollinearity. We can make a judgment by checking related statistics, such as tolerance value or variance inflation factor (VIF), Eigen value, and condition number.

Table 6.16 and Table 6.17 have been constructed to make a judgment on the presence of the magnitude of multicollinearity in the preferred model. From the tolerance values it is seen that all are ranging from a low of 0.4 to as high as 0.9 thereby generalizing that the tolerance values of all independent variables is greater than 0.1 which again limits the values of VIF sufficiently less than 10 (Kleinbaum et al 1988). Primarily, it may therefore be expected that the problem of multicollinearity is not severe in the model. It may be worthy to note here that VIF shows how multicollinearity has increased the instability of the coefficient estimates (Freund and Littell 2000: 98). Putting it differently, Variance Inflation Factor (VIF) provides a measure of how much the variance of the estimated regression coefficient is inflated as compared to when the variables are not linearly related.

There is another use of VIF that relates to Multicollinearity of the model. If the mean value of VIFs related with the regressors is much greater than 1, serious problems are indicated (http://www.stat.lsu.edu /faculty/geaghan/EXST7034/Fall2005/E34MaterialsFall2005.html). In the present model the value is found be as low as 1.6 which seems to be not much higher than 1. This suggests that no serious concern remains so far as multicollinearity is concerned.

Table-6.16: Tolerance and Multicollinearity

Variables	Tolerance	VIF
EDNTOTEX	0.762	1.312
MPCE	0.691	1.446
FWPR	0.850	1.177
ECONDEP	0.445	2.249
HHDEP	0.456	2.194
INREGIRREG	0.806	1.240
OPTNTCOST	0.867	1.153
FTHEDN	0.504	1.982
MTHEDN	0.596	1.677
Mean of VIFs		1.604

In order to have an alternative test of the Multicollinearity problem, Collinearity Diagnostics has been analysed in our model. Statistically, a set of eigenvalues of relatively equal magnitudes indicates that there is little multicollinearity (Freund and Littell 2000: 99). Condition numbers or condition indices are square roots of the ratios of the largest eigen values to individual ith eigen values. Conventionally, an eigen value close to zero (say less than .01) or condition number greater than 50 (30 for conservative persons) indicates significant multicollinearity. Belsley, Kuh, and Welsch (1980) insist 10 to 100 as a beginning and serious points when collinearity affects estimates. In both the cases, the Collinearity Diagnostics table (Table-6.17) as calculated for the survey data does not seem to be indicative that there is a severe multicolloinearity that needs to be addressed in the model.

Table-6.17: Collinearity Diagnostics

Dimension	Eigen value	Condition Index
1	6.123	1.000
2	1.216	2.244
3	0.760	2.838
4	0.599	3.198
5	0.469	3.613
6	0.299	4.526
7	0.210	5.394
8	0.168	6.039
9	0.125	6.992
10	0.031	14.042

6.6.3 Interpretation of Regression Coefficients

The remaining task is now to determine which variables statistically and significantly contribute to the model. The significance tests for individual regression coefficients assess the significance of each predictor variable assuming that all other predictors are included in the regression equation. The level of statistical significance required to reject the null hypothesis (i.e., to obtain a statistically significant result) is set conventionally at .05, or 5%. While the 5% criterion is typical, reporting of more stringent 1% significance tests or less stringent 10% tests can also provide useful information. In the present analysis the level of statistical significance has been assumed at the level of less stringent 10%. It is also to be noted here that if the sample is random and large enough, one can assume the distribution of the errors of estimation would be normally distributed. As the sample becomes smaller, the same will follow the't' distribution. So far the sample size of the present study is concerned (120 out of 477 households, i.e. only 25%), it is reasonable to use the t- statistics for obtaining the statistical significance of the model.

Out of nine explanatory variables that are incorporated in the model, seven variables namely, EDNTOTEX, MPCE, ECONDEP, INREGIRREG, OPTNTCOST, FTHEDN and MTHEDN, show statistically significant relationship with the literacy characters (PLR, MLR, FLR) of the households. However, all the explanatory variables are not found to be equally significant and enabling in determining each of the dependent variables for which the explanatory variables were used. The regression coefficients (un-standardized) along with their standard errors, t-values and the level of significance (p-values) have been shown in Table 6.18.

However, by applying the stepwise regression technique (backward or forward) one could drop the so-called insignificant variables. But one should also be aware about the potential costs that may arise if the so-called insignificant variables were excluded from the equation. In reality, insignificance arises if the p-value is greater than some "critical" level (e.g. 0.10 as assumed in the model). Under such an assumption, in a standard t-test, one can generally conclude that the variable has no explanatory power in the model if the p-value is 0.10001. Such a conclusion is probably not always justified, because, the so-called insignificance is a statement about precision in estimation and not about causality. Actually, the larger the p-value, the more likely the null hypothesis is true. In a justified manner, one can conclude that the probability of making a Type I error (i.e. rejecting the null hypothesis) is slightly higher than what is otherwise likely to be assumed. But in no way does the decision to reject or not, mean that the variable is uninformative. Moreover, the dropped variables are now part of the error term and to the extent these dropped variables are correlated with any included regressors will be critical since the coefficients on those regressors, under such circumstances, will be estimated with bias. This is a much more serious problem than including regressors with imprecisely estimated coefficients. In conclusion, the decision to mark a variable as insignificant is a subjective choice and one is liable to falsely think that p-values are more objectively applied. For this technical point, the present regression exercise

has included all the nine (09) explanatory variables without dropping any. However, attempt will be made to assign economic significance to not-so-significant variables.

Table-6.18: Regression Result

Variables	Model-1 PLR			
	Un-standardized co-efficient	Std. Error	't' value	Sig
(CONSTANT)	26.367	8.125	3.245	0.002
EDNTOTEX	0.909**	0.453	2.009	0.047
MPCE	0.053***	0.018	2.932	0.004
FWPR	5.571	4.017	1.387	0.168
ECONDEP	-4.780*	2.834	-1.686	0.095
HHDEP	1.296	4.002	0.324	0.747
INREGIRREG	9.537**	4.124	2.312	0.023
OPTNTCOST	-11.623***	4.160	-2.794	0.006
FTHEDN	0.770	0.652	1.181	0.240
MTHEDN	3.408***	0.839	4.060	0.000

R2 = 0.586; Adj. R2 = 0.553; N = 120

Variables	Model-2 MLR			
	Un-standardized co-efficient	Std. Error	't' value	Sig
(CONSTANT)	22.893	11.560	1.980	0.050
EDNTOTEX	0.379	0.644	0.589	0.557
MPCE	0.085***	0.025	3.330	0.001
FWPR	-1.007	5.715	-0.176	0.860
ECONDEP	-4.827	4.032	-1.197	0.234
HHDEP	3.472	5.694	0.610	0.543
INREGIRREG	14.563**	5.867	2.482	0.015
OPTNTCOST	-9.324	5.918	-1.575	0.118
FTHEDN	2.201**	0.928	2.372	0.019
MTHEDN	0.661	1.194	0.553	0.581

R2 = 0.431; Adj. R2 = 0.384; N = 120

Variables	Model-3 FLR			
	Un-standardized co-efficient	Std. Error	't' value	Sig
(CONSTANT)	25.348	12.639	2.006	0.047
EDNTOTEX	2.141***	0.704	3.039	0.003
MPCE	0.028	0.028	1.021	0.309
FWPR	11.587*	6.249	1.854	0.066
ECONDEP	-2.950	4.409	-0.669	0.505
HHDEP	-4.580	6.225	-0.736	0.463
INREGIRREG	-4.600	6.415	-0.717	0.475
OPTNTCOST	-12.428*	6.471	-1.921	0.057
FTHEDN	-0.756	1.014	-0.745	0.458
MTHEDN	7.015***	1.306	5.372	0.000

R^2 = 0.454; Adj. R^2 = 0.409; N = 120

Note: *** significant at 1% level, ** at 5% and * at 10%

Another point related with the model estimation is that the values of the least-squares estimators vary from sample to sample. If continued collection of more and more samples are generated as additional estimates, as might happen when new data become available over time, the estimates of each parameter would follow a probability distribution. This probability distribution can be summarized by a mean and a measure of dispersion around the mean, a standard deviation that usually is referred to as the standard error of the coefficient, or the standard error. Small standard errors imply results that are likely to be similar from sample to sample, while results with large standard errors show more variability. In the present model, the variables that are statistically

significant show small standard error of estimates thereby reflecting small variability in connection with the values of the estimators (quite consistent estimators) associated with each significant explanatory variable.

6.6.4 Discussion on Regression Result

Interpretation of the regression results in statistical terms is done through the use of significance tests and other specifications. However, the result can be interpreted in a more practical, non-statistical manner, i.e., how far the statistically significant results sustain with the practical significance in question. This is much more important than a mere interpretation of the statistical model. In analyzing the present result, we shall compare it with the results that have been found in our earlier chapters (chapter-IV & chapter-V). In multiple regression, the regression coefficients describe the effect of an explanatory variable while controlling effects of the other explanatory variables in the model. Based on this statistical concept, the result of the regression has been analysed in a more detailed manner by capturing the practical significance as observed during the data collection.

In explaining the literacy rate at district and block level in chapter-IV, asset holding has been treated as a surrogate for income. There it was observed that holding of no assets by the household bears a significant negative impact on both male and female literacy rate at block level in the state. However, it does not bear any significant impact on female literacy rate at district level. In the present case, two expenditure related variables has been tested to observe the impact on literacy achievement namely Monthly per capita expenditure (MPCE) and expenditure on education as proportion of total expenditure. MPCE is commonly used as a proxy of income and remains as a significant variable in determining the overall literacy rate (PLR) and male literacy rate. On the other hand, expenditure on education as proportion of total expenditure (EDNTOTEX) has been found to significantly explain the overall literacy rate and female literacy rate. The basic objective of incorporating expenditure variables in two different dimensions is to observe the effect of marking a larger share of expenditure for literacy achievement. Some interesting results come out from the regression result. Both the expenditure variables have been found to be significant in explaining the overall literacy rate. But the explanatory power of EDNTOTEX (0.909) is much higher than the conventional MPCE (0.053). This is probably because the share of expenditure of a household on education is a focused variable than the actual MPCE which includes expenditure on basic needs with food expenditure taking a prime share. Again, EDNTOTEX is also significant in analyzing the female literacy rate while it is insignificant in explaining the male literacy rate. This gender difference explains to some extent that an intention of spending more income on education may enhance the female literacy rate.

In Chapter-IV, while analyzing the district level and block level regression equation for West Bengal, the variable Female Work Participation Rate (FWPR) was found to be significant in exerting its negative impact on literacy rate. At mouza level (in Chapter-V) it was found that the variable is significant for Goalpokhar-I block but insignificant for Chopra block. Several socio-economic factors are responsible for female work participation. It may be due to the fact that the female members are well equipped with skill and knowledge for participation in the labour market. On the other hand, it may well be that the earning of the male members of the household is not sufficient to maintain the family. From the cross tabulation of data on mothers' education level it is observed that out of 116 mothers, 58 mothers are completely illiterate and 36 out of remaining 58 literate mother hardly have the primary level of education. As such it may be assumed that the latter reason seems to be more acceptable for work participation of the female members in the study area. With a poor educational level, it would be very optimistic to assume that the mothers are well aware of the economic and social benefits of education. Consequently, in the present primary data analysis, FWPR does not impact literacy rate in a significant manner.

As mentioned earlier, economic dependency ratio is calculated as the ratio of non-earners to earners. The higher value of the ratio represents the magnitude of economic burden on a particular household. It explains that households with larger number of non-workers imply increased economic pressure on the earner/s of the family. Economic dependency ratio (ECONDEP) has

been tested at mouza level where it was found that the ratio is significant in those blocks where the value of the ratio is comparatively low (Chapter-V). In the present analysis the overall value of the ratio is calculated to be 1.87 (Table 6.10) which is considerably higher than the overall district average (1.67) and accordingly it may be expected that the said variable could barely show its effect on literacy variation at household level analysis. Although ECONDEP is found to be significant in explaining the overall literacy rate, in isolation, it does not bear any statistically significant effect on literacy rate. It may however be noted here that economic dependency is directly related to the level of poverty, both moving in the same direction. To sustain with such extreme poverty is the immediate pressure on the household even at the cost of engaging the schooling age children in the labour market. The household dependency which calculates the non-economic burden of a family as discussed in the earlier section of this chapter may have some effects on literacy attainment of a household. Thus in addition to economic dependency, household dependency ratio has been incorporated as an additional explanatory variable in the regression equation. However, the variable does not have any statistically significant effect on the literacy rate. This may be due to the fact that the non-economic burden of taking care of the siblings and old-age members does not stand as an additional barrier for educational development in the study area.

Among the parental characteristics, the educational level of both the parents has been incorporated as variables affecting literacy rate. In explaining the overall and female literacy rate, educational level of mothers is found to have statistically significant impact on the literacy variable. Mothers' educational level however do not affect the male literacy rate in a significant way while it has strong positive effect on female literacy rate. Similarly, father's education remains insignificant in determining the female literacy rate, but it exerts a positive role on male literacy rate. It explains that in order to enhance female literacy levels in an educationally deprived area, it is necessary for the present generation girl children to be schooled so that these educated future mothers in the society will further the education of next generation learners.

Occupational diversification has been captured by including a dummy variable with 0, 1 options. Households having their primary income in a regular pattern have been coded numerically as 1, while it is 0 otherwise. This variable shows its positive association with male literacy rate. This means that a higher male literacy is found in the households having regular pattern of income. The households having a certainty in income pattern enjoys economic security. For example, in Kantigach village, there are lot of earners who work as tea garden workers basically engaged either on permanent basis or as casual workers. The permanent workers are provided with ration, medical and provident fund facilities. Casual workers are economically insecure and this insecurity compels the other family members to participate in the labour market for maintaining their livelihood. In case of marginal cultivators and agricultural workers (observed mostly in Kantigach village) an irregular pattern in earning is also very common. In such a situation the earners, mostly the male members of a household migrate to other states like Haryana, Delhi, Sikkim etc., in search of work as skilled or unskilled labourers. This economic insecurity compels the household to engage more number of male members in the labour market for enhancing the income. This may be responsible for poor male literacy rate in the village. However, the income pattern does not show any significant relation with the female literacy rate whatsoever.

Opportunity cost of schooling is included in the model to observe its effect on literacy rate. In order to capture whether a household considers the opportunity cost of sending a child to school or not, was achieved by collecting data on children in the age group up to 18 years participating in the labour market from a particular household. After assigning a numeric code 1 to those households where at least one child has been working in labour market and 0 other wise, it was found from the regression analysis that this dummy variable exerts a strong negative effect in explaining the female literacy rate. The same is not found to be significant in determining the male literacy rate. This represents that if there is enough scope of girl children to be engaged in labour market, this will adversely affect the female literacy rate. The magnitude of child labour is found to be the highest in Juropani village where it is observed that the girls in this age group usually work in the tea garden as casual workers, especially during the plucking period (generally from the April to

September in a calendar year) of a year. The female literacy rate of this village is also considerably low (37.0%) as compared to the other villages (Table-6.9). There is thus a strong indication that there exists a negative association between child labour and female literacy, frequent evidence of which is found in the literature on education.

6.7 Comparison of the Relative Magnitude of Regression Coefficients

So far discussion has been limited to the overall relationship between literacy rate and the predictor variables that are included in the model and the magnitude of the variance in literacy rate that is accounted for by the predictor variables (ANOVA and Multiple Correlation R2). Interpretation of the direction and magnitude of the effect of each predictor variable on literacy rate (explaining the regression coefficients) is also undertaken. But frequently, the predictor variables in a regression equation (Xk) are in different scales of measurement. For example, INREGIRREG and OPTNTCOST are categorical assuming only two values (0 or 1), MPCE is expressed in money terms, parental educational level is expressed by the completed years of education, EDNTOTEX in percentage and so on. Now by comparing the regression coefficients of each predicator, it is found that the regression coefficient for INREGIRREG is the largest (Table-6.13). From this, it can't be concluded that INREGIRREG has the greatest impact on literacy rate. For the task of comparing the various regression coefficients in a regression equation that contains variables in different scales of measurement, the coefficients must first be standardized and the standardized regression weights will help in comparing the Relative Magnitude of Regression Coefficients. But there are some difficulties in using the standardized regression weights. Beta weights are interpretable and useful when the independent variables are not collinear. The magnitude of the beta weights can only be compared with the variables in the equation. If additional variables are added to the equation, the beta weights of the previous variables are likely to change, particularly if they are collinear with the newly added variables. Similarly, the beta weights in an equation cannot be compared with the beta weights found in similar studies, since their value is sample specific. Moreover, beta weights are sensitive to the ranges of the independent variables. Any change in ranges will change the estimated beta weights as well. Owing to the caveats in interpreting standardized regression coefficients, one can alternatively suggest (Bryman and Cramer, 2005) to calculate the Partial and Semi-Partial correlation coefficients between the dependent and explanatory variables separately for each predictor. An attempt will now be made to describe and give a brief interpretation of partial and semi-partial correlation coefficients in respect of the present model. Using SPSS, partial correlations are labeled as "Partial", and semipartial correlations are labeled "Part."

Partial correlation represents the correlation between the dependent variable and a predictor after common variance with other predictors has been removed from both the dependent variable and the predictors of interest. That is, after removing variance that the dependent variable and the predictor have in common with other predictors, the partial expresses the correlation between the residual predictor and the residual criterion.

A semipartial correlation coefficient represents the correlation between the dependent variable and a predictor that has been residualized with respect to all other predictors in the equation. It may be noted here that the dependent variable remains unaltered in the semipartial. Only the predictor is residualized. An important advantage of the semipartial is that the denominator of the coefficient (the total variance of the dependent variable) remains the same no matter which predictor is being examined. This makes the semipartial very interpretable (Neil, 2007).

The partial and semipartial correlation coefficient of literacy rate (PLR, MLR and FLR) with the predictors is given in Table-6.9 above. A careful observation of the Table reflects that the educational level of mothers is dominant in explaining the overall (PLR) and female literacy rate (FLR). Monthly per capita expenditure used as a surrogate for income is the second relatively important factor in explaining the overall literacy rate. However, this income factor is more important in explaining the male literacy. The opportunity cost of sending a child to school occupies the third position among all the predictors in explaining the PLR. The pattern of income (INREGIRREG) which has the dominant regression coefficient is not significant in part and partial

correlation analysis, although it was found to have considerable importance in explain the PLR and MLR from the previous regression analysis.

Table-6.19: Partial and Semi-partial Correlation Coefficients for the Literacy Rate

Predictors	Model-1 PLR			Model-2 MLR			Model-3 FLR		
	Zero-order	Partial	Part	Zero-order	Partial	Part	Zero-order	Partial	Part
EDNTOTEX	0.384	0.188	0.123	0.268	0.056	0.042	0.390	0.278	0.214
MPCE	0.520	0.269	0.180	0.517	0.303	0.240	0.334	0.097	0.072
FWPR	0.242	0.131	0.085	0.117	-0.017	-0.013	0.237	0.174	0.131
ECONDEP	-0.301	-0.159	-0.103	-0.209	-0.113	-0.086	-0.264	-0.064	-0.047
HHDEP	-0.196	0.031	0.020	-0.113	0.058	0.044	-0.229	-0.070	-0.052
INREGIRREG	0.403	0.215	0.142	0.390	0.230	0.179	0.172	-0.068	-0.051
OPTNTCOST	-0.348	-0.257	-0.171	-0.304	-0.149	-0.113	-0.222	-0.180	-0.135
FTHEDN	0.537	0.112	0.072	0.491	0.221	0.171	0.370	-0.071	-0.052
MTHEDN	0.540	0.361	0.249	0.352	0.053	0.040	0.539	0.456	0.379

In explaining the MLR, MPCE is found to be the most important variable along with income pattern (INREGIRREG) and educational level of father. In explaining the effects on FLR, mother's educational achievement followed by proportion of educational expenditure plays a vital role than the other predictors included in the model. It is interesting to note that female work participation affects the FLR adversely and it has a large magnitude too. This variable however has negligible effect on other literacy characters.

6.8 Conclusion

Summarising, it may be said that the demand side socio-economic factors have considerable impact on determining the literacy rate in an educationally deprived area. Parental education, especially mothers' education, is found to be important in determining the literacy character in a deprived area. Apart from the supply related educational amenities, it is the educational level of mothers that can alone enhance the female literacy rate to a considerable extent than other household related factors. So girls' education needs to be given more priority in order to have a balanced literacy development in a near future in such educationally backward areas. Monthly per-capita expenditure (proxy of income) remains important but what is more important is that how much income a household is willing to spend on education. This proportion is more important in enhancing the female literacy rate and as means of removing gender differential in literacy rate. Availability of opportunity to work in the paid labour market impedes literacy development in backward areas. Economic crisis of a family compels the parent to engage their children in such paid activities to ease economic pressure on the family. Thus, policy related to abolition of child labour is a way to ensure schooling of children and also a measure to prohibit employers from engaging a child worker.

References

Baba, Kunihiro; Ritei Shibata & Masaaki Sibuya 2004. Partial correlation and conditional correlation as measures of conditional independence, *Australian and New Zealand Journal of Statistics* 46 (4): 657–664. doi:10.1111/j.1467-842X.2004.00360.x published Online: 13 Dec 2004

Belsley, David. A., Edwin. Kuh, and Roy. E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, New York: John Wiley and Sons.

Bryman Alan and Duncan Cramer 2005. *Quantitative data analysis with SPSS 12 and 13: A guide for social scientists,* Routledge, 27 Church Road, Have, East Sussex, BN3 2FA

Charles M. Friel. *Multiple Linear Regression*, Criminal Justice Center, Sam Houston State University

Chakrabarti Singdha and Chaiti Sharma Biswas 2008. *Women Empowerment, Household Condition and Personal Characteristics: Their Interdependencies in Developing Countries,* Discussion Paper, Economic Research Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata 700108, India

Debi, Sailabala 2001. Inequality of Access to Elementary Education in Orissa: An Inter- and Intra-spatial Analysis, in Vaidyanathan A. and P R Gopinathan Nair (eds) '*Elementary Education in Rural India: A Grassroots View*', Sage Publication, New Delhi.

Dreze, Jean and Amartya, Sen 2002. *India: Development and Participation*, Oxford University Press, New Delhi

Duraisamy, Malathy 2001. Demand for and Access to Schooling in Tamil Nadu, in Vaidyanathan A. and P R Gopinathan Nair (eds) *Elementary Education in Rural India: A Grassroots View,* Sage Publication, New Delhi.

Ethington, C. A., Thomas, S. L., & Pike, G. R. 2002. Back to the basics: Regression as it should be" in J. C. Smart (Ed.), *Higher education: Handbook of theory and research*, Vol. 17. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Federal Bureau of Statistics, Government of Pakistan 1998. *Education sector performance in the 1990s: Analysis from the PIHS*, Islamabad, Government of Pakistan

Freund, Rudolf J. and Ramon C. Littell. 2002. *SAS for Linear Models, Fourth Edition* (Fourth edition). SAS Publishing.

Galton, F. 1885. Regression toward mediocrity in heredity stature, Journal *of the Anthropological Institute, 15*, 246-263

Green, S. B. 1991. How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research, 26*, 499-510.

Hazarika, G. 2001. The sensitivity of primary school enrollment to the costs of post-primary schooling in rural Pakistan: A gender perspective, *Education Economics,* 9(3), 237–244.

Jabbi, M K. and C. Rajyalakshmi 2001. Education of Marginalised Social Group s of Bihar, in Vaidyanathan A. and P R Gopinathan Nair (eds) '*Elementary Education in Rural India: A Grassroots View*', Sage Publication, New Delhi.

Jejeebhoy, Shireen J 1993. Family Size, Outcomes for Children and Gender Disparities: Case for Rural Maharastra, *Economic and Political Weekly*, August 28: 1811-21.

Knodel, John and Malinee, Wongsith 1990. Family Size and Children's Education in Thailand: Evidence from a National Sample, *Demography*, 28(1): 119-31.

Krishnaji, N. 2001. Poverty, Gender and Schooling in Andhra Pradesh, in Vaidyanathan A. and P R Gopinathan Nair (eds) *Elementary Education in Rural India: A Grassroots View*, Sage Publication, New Delhi.

Levine, R. and Renelt, D.R. 1992. A sensitivity analysis of cross country regression, *American Economic Review* 82, 942-963

Lincove, J. Arnold 2009. Determinants of schooling for boys and girls in Nigeria under a policy of free primary education, *Economics of Education Review*, Vol. 28, Issue 4, (August 2009) Pages 474-484 www.elsevier.com/locate/econedurev

McMahon, Walter W. 1998. Education and Growth in east Asia, *Economics of Education Review*, Vol. 17, No. 2, pp 159-172, www.elsevier.com/locate/econedurev

Mukherjee, Dipa 2005. Educational Attainment in India' *Journal of Educational Planning and Administration*, Vol. XIX, No.4: 523-541; National Institute of Educational Planning and Administration, New Delhi.

Mukhopadhyay, Asim 1994. Girl Child in Three Indian States, *Economic and Political Weekly*, June 4, 1379-82

Neil Salkind (Ed.) 2007. *Encyclopedia of Measurement and Statistics*, Thousand Oaks (CA): Sage.

Pandey, G D. 1990. Some Correlates of Literacy and Educational Attainment among Tribal Children in Rural Madhya Pradesh', Demography India, 19(2): 289-300.

Myers, R.H. 1990. *Classical and modern regression with applications*, PWS-KENT Publishing Co. Boston

Pande, Anuradha 2000. Education of Rural Children in UP Himalayas' in Rekha Wazir (ed) '*The Gender Gap in Basic Education*', Sage Publication, New Delhi.

Petrakis P.E. and Stamatakis D. 2002. Growth and educational levels: a comparative analysis' *Economics of Education Review* 21 (2002) 513–521, www.elsevier.com/locate/econedurev

Public Report on Basic Education in India (1999); Oxford University Press, New Delhi.

Psacharopouslos, George and Ana Maria Arrigada 1989. The Determinants of Early Age Human Capital Formation: Evidence from Brazil, *Economic Development and Cultural Change*, 37(4): 683-708.

Reddy, V R and R Nageswar Rao 2003. Primary Education: Progress and Constraints, *Economic and Political Weekly*, Vol. XXVIII No.12&13.

Sathar, Z. A., & Lloyd, C. B. 1994. Who gets primary schooling in Pakistan: Inequalities among and within families, *Pakistan Development Review*, 33(2), 103–34

Sengupta, Piyali and Jaba Guha 2002. Enrolment, Dropout and Grade Completion of Girl Child in West Bengal, *Economic and Political Weekly*, Vol. XXVII No. 17, 1621-37.

Stromquist, N. 2005. The political benefits of adult literacy, Background Paper for *EFA Global Monitoring Report* 2006; p. 12

Tansel, Aysit 1997. Schooling Attainment, Parental Education and Gender in Cote d' Ivoire and Ghana, *Economic Development and Cultural Change*, 45(4): 825-56

Thind, Satinderjit and Sushma Jaiswal: 'Gender Bias in Education Among the Rural Sikhs of Punjab', Journal of Educational Planning and Administration, Vol. XVIII, No.1, National Institute of Educational Planning and Administration, New Delhi.

Thomas Joseph A. 2000. Dynamics of Educational Development: A case Study of Selected Backward Villages in Kerala, in Rekha Wazir,(eds) '*The Gender Gap in Basic Education*', Sage Publication, New Delhi.

UNESCO 2005. *EFA Global Monitoring Report 2006*, United Nations Educational, Scientific and Cultural Organization, 7, Place de Fontenoy, 75352 Paris 07 SP, France

Unni, Jeemol 1998. Gender Differential in Schooling, in M. Krishnaraj, Ratna M. Sudarshan and Abusaleh Sharrif (eds), *Gender, Population and Development*, Oxford University Press, New Delhi.

Venkatanarayana M. 2004. *Educational Deprivation Of Children In Andhra Pradesh- Levels and Trends, Disparities and Associative Factors*, Working Paper 362 Centre for Development Studies, Prasanth Nagar,Ulloor, , Trivandrum 695 011, Kerala, India.

Weiner, Myron 1996. Child Labour in India: Putting Compulsory Primary Education on the Political Agenda, *Economic and Political Weekly*, Vol. 9: 3007-16.

Wazir, Rekha 2000. Profiling the Problem, in Rekha Wazir,(eds) '*The Gender Gap in Basic Education*', Sage Publication, New Delhi.

World Bank 2002. *Pakistan poverty assessment; poverty in Pakistan: Vulnerabilities, social gaps and rural dynamics*. Report No. 24296–PAK. Washington DC.

APPENDIX-6-I

The Determinants of Literacy and Educational Attainment: An Investigation of Regional Patterns of Elementary Education in West Bengal :

Household Census, 2005-06

District:- Uttar Dinajpur, CD Block:- _____ village:- Kantigach, Date of Survey:- _____

Sl No	HH No	Name of the Head	Age	Sex	No. of 5-14 age group children			Caste	Religion
					M	F	T		

The Determinants of Literacy and Educational Attainment: An Investigation of Regional Patterns of Elementary Education in West Bengal: Household Census, 2005-2006

APPENDIX-6.I (cont)

The Determinants of Literacy and Educational Attainment: An Investigation of Regional Patterns of Elementary Education in West Bengal: Household Census, 2005 - 2006

Parental Information

Father			Mother	
Name			Name	
Age			Age	
Caste			Caste	
Religion			Religion	
Sub-community			Sub-community	
Mother Tongue			Mother Tongue	
Educational Level			Educational Level	
Occupation			Occupation	
a. Primary			a. Primary	
b. Secondary			b. Secondary	
Income (Yearly)	Rs.		Income (Yearly)	Rs.
Aspiration in educating the children	Upto Pri / U Pri / Secondary / Higher Secnodary / Graduate & above		Aspiration in educating the children	Upto Pri / U Pri / Secondary / Higher Secnodary / Graduate & above
Level of awareness regarding direct / indirect benefit of education	**DIRECT:** increases employment prospects, enhances earnings, raises social status, nil, others		Level of awareness regarding direct / indirect benefit of education	**DIRECT:** increases employment prospects, enhances earnings, raises social status, nil, others
	INDIRECT: incentives generates indirect monetary benefits, improves marriage prospects, develops good manners and habits, future resort of the parents, ensures family planning, nil, others			**INDIRECT:** incentives generates indirect monetary benefits, improves marriage prospects, develops good manners and habits, future resort of the parents, ensures family planning, nil, others
Reasons for non-realisation of own aspiration	I) II) III)		Reasons for non-realisation of own aspiration	I) II) III)
Opinion about measures to be taken for realising the aspiration	I) II) III)		Opinion about measures to be taken for realising the aspiration	I) II) III)
Who decides?	Enrolment of Children	continuation of schooling	In Economic matter	In matrimonial purpose
	Father / Mother	Father / Mother	Father / Mother	Father / Mother

APPENDIX-6.I (cont)

The Determinants of Literacy and Educational Attainment: An Investigation of Regional Patterns of Elementary Education in West Bengal: Household Census, 2005-2006

Basic Informations												
1.Village Informations												
Name of the State	West Bengal											
Name of the District	Uttar Dinajpur											
Name of the CD Block												
Name of the Gram Panchyat												
Name of the Census Village												
Gram Sansad No												
Total No of Households												
2. Basic Amenities												
Medical Amenities	Distance Range of nearest Primary Health Centre (in Kms)											
Communication Amenities	Distance Range of nearest Bus Connectivity (in Kms)											
Banking	Distance Range of nearest Bank / Credit Society / others (in Kms)											
Presence of Educational Institution within the Mouza/village	Pre-primary School				Primary School				Upper-primary / Secondary School			
	Nil	Govt	Private	Others	Nil	Govt	Private	Others	Nil	Govt	Private	Others
3. School Informations												
Educational Institution	Distance Range of nearest Pre-Primary Shool											
	Distance Range of nearest Primary Shool											
	Distance Range of nearest Upper Primary Shool											

	No of students		No of Units	No of Class Rooms		No of Teachers		Pupil-Teacher Ratio	% of Female Teachers
	M	F				M	F		
Primary School									
U Primary School									

School Infrastructure at village level	Sanitation Facility		Playground		DW Facility		Blackboard in every classroom		Library		Cooked Meal	
Type of School	Pri	U. Pri	Pri	U. Pri	Pri	U. Pri	Pri	U Pri	Pri	U. Pri	Pri	U. Pri
Available & Functional												
Available & Non-Functional												
Not available												

The Determinants of Literacy and Educational Attainment: An Investigation of Regional Patterns of Elementary Education in West Bengal: Household Census, 2005 – 2006.

Household Schedule

Household No.-

Name	Sex	Age (in complete year)	Relation with the Head	Educational Level	Present Educational Status	Work Status	Place of work	Reason for working outside the village	Reason for dropout / never enrolment	Reason for enrolment	Marital Status	Mean age at marriage (Expected if not married)	Amount of Dowry paid / received (expected if not married), in Rs.	Earner/ Dependant	Cause of Dependency	Pre-primary Schooling status	Type of Pre-primary School
5-14 Years	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1																	
2																	
3																	
4																	
5																	
Others																	
1																	
2																	
3																	
4																	
5																	

APPENDIX-6.I (cont)

The Determinants of Literacy and Educational Attainment: An Investigation of Regional Patterns of Elementary Education in West Bengal: Household Survey

Technical Note

Sex	1 Female 2 Male 3 other Brother
Relation with the Head	0 Head of family 1 Father 2 Mother 3 Brother 4 Sister 5 Son 6 Daughter 7 Husband 8 Wife 9 Other
Educational Level	0 Illiterate, 1 up to class-I, 2 up to class-II, 3 up to class-III and so on
Present Educational Status	1 currently enrolled, 2 out of school & never enrolled, 3 out of school after completing at least 8 years schooling, 4 dropout at elementary level
Work Status	1 sibling care 2 house- hold work 3 working for enhancing family income 4 studying in school 5 not in school nor in work
Place of work	1 within the household 2 within the Village 3 outside the village
Reason for working outside the village	1 Job opportunity is not available within the village 2 for higher wages 3 other (to be specified)
Reason for Non-enrolment/ dropout	1. Child is needed for other activities 2. School is too expensive 3. School did not seem to be interested 4. Child could not qualify 5. Child not interested in study 6 Poor family conditions 7. School is too far way 8. Child is disabled 9 Parental reluctance 10 others (to be specified)
Reason for enrolment	ECONOMIC: E1. Increases employment prospects E2.school incentives generates indirect monetary benefits E3.enhances earnings E4. Raises productivity E5. Others E6. No reason given
	NON-ECONOMIC: S1. Raises social status S2. Improves marriage prospects S3. Develops good manners and habits S4. Ensures family planning S5. Enhances the feeling of security S6. Education helps to communicate with the modern world S7. Others S8. No reason given
Marital Status	1 never married 2 Currently married 3 Widowed 4 Separated/divorced 5 others
Earner/ Dependant	1 for Earner 2 for Dependant
Cause of Dependency	1 Minor 2 Old age member 3 Physically incapable 4 involved in un-paid family work 5 Student 6 Unemployed
Pre-primary Schooling status	1 for yes 2 for no
Type of Pre-primary school	1 for Government 2 for Private