

CHAPTER- 6

Factors determining the learning achievement score of the students

In this chapter we are making an attempt to identify the factors that can influence the students' learning achievement scores. These variables may be either the ratio scale variables, like income of the households to which the students belong, or the qualitative variables or the nominal scale variables such as educational qualification of the parents or the Head of the family, sex, caste, school management, location of the school, class attendance, school infrastructure and the like. In our problem, we have considered both the ratio scale or quantitative and nominal scale or the qualitative variables as the most likely factors that can influence the dependent variable, or the regress and which is the learning achievement score of the students of the primary schools located in rural as well as urban areas in Midnapore town, and in its surrounding rural areas.

6.1: The total number of schools (primary) considered in our study is sixteen (16), out of which eight (8) are private schools and eight (8) are public schools (Government - aided). Again, of these 16 schools, 8 are in the rural area and 8 in the urban area. The total number of students constituting our sample is 154. This means that the size of our sample is 154. Data on various quantitative and qualitative variables have been collected for these 154 students. The qualitative variables considered in our study are, namely, academic qualifications of the parents; sources of income of the parents or the Head of the family (such as service, business, farming and others); school infrastructure levels such as excellent, very good, good and not good; private coaching; organization or management of school (private and public); area or location such as rural or urban; gender such as male or female; caste such as general and others including SC, ST and OBCs; class attendance of the students. Annual income of the parents is only the quantitative variable.

All the variables except the last one, i.e., annual income of the parents, are the qualitative variables. The dummy variables are represented by $D_2, D_3, D_4, \dots, D_{16}$, and the income variable by X_2 .

The model consisting of the dependent variables (the learning achievement score of all the students under consideration) and the independent variables, is presented below (The descriptions of the variables are given in the chapter 3).

$$y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 D_{4i} + \beta_5 D_{5i} + \beta_6 D_{6i} + \beta_7 D_{7i} + \beta_8 D_{8i} + \beta_9 D_{9i} + \beta_{10} D_{10i} + \beta_{11} D_{11i} + \beta_{12} D_{12i} + \beta_{13} D_{13i} + \beta_{14} D_{14i} + \beta_{15} D_{15i} + \beta_{16} D_{16i} + \gamma X_{2i} + u_i \dots\dots\dots(6.1)$$

Description of the variables:

Y_i = Learning achievement scores of the students who constitute a cross-section of our sample, drawn from the population for our empirical study. The male students constitute another cross-section of our sample.

$D_{2i} = 1$ if the academic qualification of the parents of the student i is graduate or above
 $= 0$ otherwise

$D_{3i} = 1$ if the academic qualification of the parents of the student i is Madhyamik or Higher Secondary
 $= 0$ otherwise

$D_{4i} = 1$ if the source of income of the parents of the student i is service (government or private)
 $= 0$ otherwise

$D_{5i} = 1$ if the source of income of the parents of the student i is business
 $= 0$ otherwise

$D_{6i} = 1$ if the source of income of the parents of the students i is farming

= 0 otherwise

$D_{7i} = 1$ if the infrastructure of the school to which the student i belongs is excellent (0.75 to 1)

= 0 otherwise

$D_{8i} = 1$ if the infrastructure of the school to which the student i belongs is very good (0.6 to 0.75)

= 0 otherwise

$D_{9i} = 1$ if the infrastructure of the school to which the student i belongs is good (0.4 to 0.6)

=0 otherwise

$D_{10i} = 1$ if the student i undertakes private coaching

= 0 otherwise

$D_{11i} = 1$ if the school of the i student is a private school (it indicates school management)

= 0 otherwise

$D_{12i} = 1$ if the school of the student i is situated in urban area

=0 otherwise

$D_{13i} = 1$ if the student i is a girl

=0 otherwise

$D_{14i} = 1$ if the student i belongs to general caste

=0 otherwise

$D_{15i} = 1$ if class attendance of the student i is very good (When attendance 75 % and above)

= 0 otherwise

$D_{16i} = 1$ if class attendance of the student i is good (When attendance between 60 % and 75 %)
 $= 0$ otherwise

X_{2i} = income of the parents of the student i

All the variables are defined in chapter 3: Database and Research Methodology.

U_i = disturbance term which satisfies all the assumptions of CLRM such as $E(u_i)=0$, $E(u_i, u_j)=0, i \neq j$, $E(u_i, u_j)=\sigma^2(\text{constant}) (i=j)$ and X is non-stochastic and has finite variance.

6.2. Interpretation of the coefficients of the model

The influence of the different characteristics of the dummy variable-academic qualification of the parent of the student i represented by D_{2i} and D_{3i} which represent the graduate and above and Higher Secondary (H.S) & Madhyamik Pariksha (M.P.) passed respectively .

Taking expectation of the equation (6.1) we obtain $E(Y_i/D_{2i}=0, D_{3i}=0) = \beta_1$ (holding all other variables in the model constant).

$$\left. \begin{aligned} \text{Likewise } E(Y_i/D_{2i}=1, D_{3i}=0) &= \beta_1 + \beta_2 \\ \text{And } E(Y_i/D_{3i}=1, D_{2i}=0) &= \beta_1 + \beta_3 \end{aligned} \right\} \text{ (A)}$$

Therefore, the mean value of Y_i , i.e., the mean learning achievement score of the students whose parents are graduates and above is given by $\beta_1 + \beta_2$ and that of the students whose parents are H.S or M.P passed is given by $\beta_1 + \beta_3$.

Here for the variable – academic qualification of the parents of the students, the category for which no dummy variable is assigned is ‘below H.S/M.P’. This is known as the base or bench - mark or reference category.

The coefficients of dummy variables in (6.1) are known as the differential intercept coefficients (differential because we know from it by how much the value of the intercept coefficient that receives the value of 1 (one) differs from the intercept coefficient of the bench - mark category). For example, here β_2 , the coefficient of D_{2i} (when $D_{2i}=1$) measures

by how much the mean learning achievement score of the students of the category of our interest differs from the intercept coefficient of the bench mark category, which is β_1 , which stands for the mean learning achievement score of the students whose parents are comparatively academically less qualified.

Note that $E(Y_i / D_{2i} = 0, D_{3i} = 0) = \beta_1$

We have explained in (A) what the expected or mean learning achievement score of the students will be when their parents' academic qualification is at least graduates or higher.

While we want to know this, we keep all other variables, including all the characteristics, constant. Thus,

$E(Y_i / D_{2i} = 1, D_{3i} = 0, D_{4i} = 0, \dots) = \beta_1 + \beta_2 =$ {the mean learning achievement score of the students whose parents are graduates(atleast)}

$E(Y_i / D_{2i} = 0, D_{3i} = 1, D_{4i} = 0, \dots) = \beta_1 + \beta_3 =$ {the mean learning achievement score of the students whose parents are H.S/M.P passed.}

Note that, $\beta_1 =$ the mean learning achievement score of the students whose parents are below M.P standard.

We can, therefore write $\beta_2 = (\beta_1 + \beta_2) - \beta_1$. This means that β_2 measures the difference between the mean score of the students whose parents are at least graduates and the mean score of the students whose parents are having minimum education or no education at all i.e. below secondary level. In other words, β_2 measures how much better is the learning achievement score of the students whose parents are highly qualified (i.e. at least graduates) than the learning achievement score of the students of whose parents' education level is below secondary level (M P).

The question 'how much better' can be easily answered by testing the null hypothesis

$H_0: \beta_2 = 0$ against

$H_1: \beta_2 \neq 0$

If H_0 is accepted, then it is statistically established that parents' education levels do not influence the performance of the students. Similarly, by testing the $H_0: \beta_2=0$, against $H_1: \beta_2 \neq 0$, we can know whether the learning achievement score of the student whose parents are H.S or M.P passed is better than that of the students whose parents are below M.P. standard. In this way we can statistically test whether the different characteristics of the other dummy variables have any influence on the learning achievements of the students.

Following these rules we describe below which characteristics of different variables have statistically significant effect upon the students' learning achievement score. For this purpose we refer to the p values corresponding to each coefficient of the characteristics or variables concerned. In testing the hypothesis we set the significance level at 10 percent. Though 5% is considered as standard level, we consider probability of 0.10 as the chosen level of significance in the statistical sense. We do it without much risk, because as much as 90% confidence in making the decisions correct is a safe level of confidence so far as the social variables or institutional variables considered in our study are concerned.

6.3 In this section we consider the whole set of sample data (N=154), that include **both boys and girls**. These data pertain to all the students; each of them is a unit of the sample. The model purports to identify the factors that have statistically significant effects on the dependent variable, Y_i , which represents the learning achievement score of all 154 students.

We have total 16 explanatory variables and one of them is quantitative, which is the total annual income of the parents or the head of the family of the students, as the case may be. The other 15 variables are qualitative or dummy variables. The dummy variables are binary in nature assuming only two values – 0 and 1.

The estimated regression equation is given by

$$\hat{Y}_i = 26.240 + 13.169D_2 (\text{GRADUATE}) + 6.129D_3 (\text{H.S/M.P}) - 40814D_4(\text{SERVICE})$$

t= 3.756	2.167	1.574	-0.708
----------	-------	-------	--------

p= .000	.032	.118	.480
---------	------	------	------

$$+ 0.928D_5 (\text{BUSINESS}) + 24.906D_6 (\text{FARMING}) + 7.788D_7 (\text{INFRA EXCEL}) + 13.399D_8 (\text{INFRA V$$

GOOD)

t= 0.221	5.548	1.216	2.380
----------	-------	-------	-------

p= .825	.000	.226	.019
---------	------	------	------

$$+ 9.326D_9 (\text{INFRA GOOD}) - 0.169D_{10} (\text{PRIVATE TUTION}) + 5.700D_{11}(\text{PRIVATE SCHOOL})$$

t=1.885	-0.056	1.279	
---------	--------	-------	--

p=.062	.955	.203	
--------	------	------	--

$$+ 18.468D_{12} (\text{URBAN SCHOOL}) + 0.831D_{13} (\text{GENDER}) - 1.153D_{14} (\text{CASTE}) +$$

t= 4.637	0.279	-0.374	
----------	-------	--------	--

p=.000	.781	.709	
--------	------	------	--

$$13.179D_{15}(\text{ATTENDANCE VGOOD}) + 5.609D_{16} (\text{ATTENDANCE GOOD}) + 1.277 E-5X_2(\text{Household$$

income)

t= 2.166	1.007	0.827	
----------	-------	-------	--

p= .032	.316	.410	
---------	------	------	--

$$R^2 = 0.433, R^{-2} = 0.367, F = 6.543, \text{Sig} = .000, \text{df} = (N-K) = (154-17) = 137, \text{DW} = 1.468$$

..... (6.2)

The total number of explanatory variables considered in the above model (6.2) is 16. Of these 16 variables, as many as 15 variables are qualitative and only one is the quantitative variable that represents the total income of the family of the students.

Of these 15 qualitative variables, which have been defined in Section 6.1, only six are found to be statistically significant at less than 10% probability level, (which we have considered as the safe level that allows for only 10% risk), which implies that confidence of making correct decision regarding the acceptance or rejection of the null hypothesis about the regression equation parameters is at least 90%. The quantitative variable, i.e. the annual income of the household to which the students belong is also found to have no influence on the learning achievement scores of the students. However, of these six variables which have significant influence upon the learning achievement score of the students, five variables have influences that are significant at less than 5 % probability level. Only one, which is β_9 that represents school's infrastructure at the level 'good', is significant at 6.2 % probability level. The six dummy variables that have significant influence on the dependent variable are

D_2 (academic qualification at least graduate), D_6 (main source of income: farming), D_8 (infrastructure: very good), D_9 (infrastructure: good), D_{12} (location: urban) and D_{15} [attendance: very good, (75% and above)].

The estimated values of the coefficients of these variables and their levels of significance (p's) are presented below:

$$\hat{\beta}_2 = 13, \quad p = 0.032$$

$$\hat{\beta}_6 = 24.906, \quad p = .000$$

$$\beta^{\wedge}_8 = 13.399, p = 0.019$$

$$\beta^{\wedge}_9 = 9.326, p = 0.062$$

$$\beta^{\wedge}_{12} = 18.468, p = .000$$

$$\beta^{\wedge}_{15} = 13.179, p = .032$$

As all the values of the coefficients are positive and statistically significant at less than 6.5% probability level, we can interpret the results in the following way:

The significance level of β_2 indicates that the academic qualification with graduation or above of the parents (father/mother) has significant influence upon their children's learning achievement score. But the estimated value of β_3 is not significant; this implies that variable, representing the parent's academic qualification - H.S or M.P, does not have significant influence on the dependent variable if compared to the effect of the benchmark characteristic, mentioned above. The value of $\beta^{\wedge}_2 = 13.17$ implies that the learning achievement score of the students whose parents' are at least graduates is higher than those of the students whose parents' education level is below secondary level by 13.17, on an average. Similarly, the value of $\beta^{\wedge}_6 = 24.91$ implies that the students whose parents' main source of income is farming achieve learning achievement score which is, on an average, higher by 24.91 than the learning achievement score of the students whose parents' income comes mainly from daily works on other farms, rickshaw pulling/auto/toto driving, working as domestic help etc, i.e., the persons having no regular source of reasonable amount of income.

$\beta^{\wedge}_8 = 13.399$. This implies that the learning achievement score of the students of the schools which have very good infrastructure facilities for the students is higher by 13.399 than the students of the schools with unsatisfactory infrastructure facilities for their students.

$\beta^{\wedge}_9 = 9.326$, the regression estimate of β_9 , the coefficient of the variable D_9 . As this coefficient is statistically significant at 6.2% probability level, it means that the learning achievement score of the students of the schools whose infrastructure is very good compared to the schools where infrastructure is not satisfactory is higher by 9.33 units.

$B^{\wedge}_{12} = 18.486$ (location : urban) . It is found to be significant at 0.000 probability level. This means that it is highly significant, so we reject the null hypothesis, $H_0: \beta_{12} = 0$ and accept the alternative hypothesis $H_1: \beta_{12} \neq 0$ This signifies that the learning achievement score of the students of the school situated in the urban area, is higher by 18.47 units over that of the students of the schools located in the rural area .Thus , location is found to have significant positive influence on the students' learning achievement score.

$\beta^{\wedge}_{15} = 13.179$ = the estimate of the coefficient of the variable, 'very good (75% and above) attendance. It is highly significant at 3.2 % probability level. This variable (very good attendance), therefore, exercises significant influence on the learning achievement score of the students, which is higher by 13.18 than the learning achievement score of the students whose attendance is below 60% (the attendance below 60% is the benchmark) , which is used for comparison with the higher levels of attendance .

From the Model 6.2 we observe that there is no statistically significant difference between the learning achievement score of the male students and female students since the hypothesis

$H_0: \beta_{12} = 0$ against the alternative hypothesis $H_1: \beta_{12} \neq 0$ is accepted. In other words, this regression model shows that gender does not have influence on the students' achievement score.

6.2 Determining the Variables Explaining the Variation in the LAS's of the Female Students

In section 6.1, we have studied the effects of different social, institutional, locational and economic variables on the learning achievement scores of the students – both male and female. In this section (6.2) we consider the LAS's of the female students and identify the factors that influence the LAS's of these female students. Does the same set of variables that significantly explain the variations in LAS's of the female students and if so, the levels of significance also the same? If not the same, which variables do have relatively higher levels of confidence of making the correct decisions about the relatively greater applied in the case of male or lower influence on the variables? This is explored in this section. The same exercise will be applied in the case of male students also, which will help us to identify the variables that have greater influence with same or higher level of confidence for the population parameters such as β_2, β_3, \dots etc for the LAS's of the female and male students considered separately.

Though from the aggregative model (all models) we do not find any difference between the LAS's of the male and female students, we would like to know if there are any variable(s) and their different categories have relatively greater influences on the learning achievement scores of the female students and those of the male students considered separately. When all the students are considered together, this question could not be answered. To identify the variables that have relatively greater influence on the learning achievement score of the female students vis – a – vis that of the male students, we use the same set of independent variables for both the female students and male students, excluding the gender variable, D_{13}

($D_{13} = 1$ if female and $= 0$ if male) that is included in model (6.1), to identify the variables that influence the LAS's of the female students more than the LAS's of the male students and vice versa.

We first consider the regression model for the **female students**. The model is given by

$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 D_{4i} + \beta_5 D_{5i} + \beta_6 D_{6i} + \beta_7 D_{7i} + \dots + \beta_{12} D_{12i} + \beta_{14} D_{14i} + \beta_{15} D_{15i} + \beta_{16} D_{16i} + \gamma_2 X_{2i} + u_i \quad (i = 1, 2, \dots, 76) \dots \dots \dots (6.3)$$

The estimated regression equation to determine the learning achievement score of the female students of the all schools considered in our sample is presented below:

$$y_i = 17.180 + 6.420 D_{2i} (\text{graduate}) + 6.871 D_{3i} (\text{H.S/M.P}) - 2.495 D_{4i} (\text{service}) + 4.273 D_{5i}$$

t=	1.397	0.597	1.233	-1.263	0.661
p=	0.167	0.553	0.222	0.794	0.511

$$(\text{business}) + 32.287 D_{6i} (\text{farming}) + 25.523 D_{7i} (\text{infra ex}) + 22.395 D_{8i} (\text{infra very good})$$

t=	4.553	2.431	2.959
p=	0.000	0.087	0.004

$$+ 13.554 D_{9i} (\text{infra good}) - 3.905 D_{10,i} (\text{private coaching}) + 8.224 D_{11,i} (\text{private school})$$

t=	2.035	-1.195	1.475
p=	0.046	0.237	0.145

$$+ 17.515 D_{12,i} (\text{urban school}) - 6.643 D_{14,i} (\text{caste GEN}) + 15.293 D_{15,i} (\text{attendance very good})$$

t =	3.168	-1.623	1.452
p=	0.002	0.110	0.152

$$+ 20.256 D_{16i} (\text{attendance good}) + 1.302 E-5 X_{2i} (\text{total household income}) \quad \dots \dots \dots (6.4)$$

t=	2.066	0.665
----	-------	-------

p= 0.043

0.509

$R^2 = 0.514$, $\bar{R}^2 = 0.392$, $DW = 1.870$, $F = 4.229$, $p = .000$, $D.F = N-K = 76-18 = 58$

The model gives a very good fit (R^2 being 0.514, $DW = 1.870$, $F = 4.229$, $p = .000$ (for the model as a whole).

$N = 154$, $df = 154-18=136$)

These results clearly testify the goodness of fit of the model. Therefore, the values of the coefficients and their relevance in influencing the regress and, i.e., the dependent variable, judged on the basis of levels of significance (p - values) are worth analyzing.

Inferential decisions

From the p - values of the estimates of the coefficients, we infer that the coefficient of neither D_{2i} nor D_{3i} is significant at even 10% probability level. This means that the academic qualifications of the parents irrespective of their levels of education, do not have any influence on the learning achievement score of the students, keeping the influence of all other variables constant. The level of significance given by the p values at which the null hypothesis can be accepted are 0.167 and 0.553, i.e. 16.7% and 55.3%, which are well above chosen level of 10%. Therefore, we accept our null hypotheses $H_0:\beta_2=0$ against $H_1:\beta_2\neq 0$ and $H_0:\beta_3=0$ against $H_1:\beta_3\neq 0$. However, compared to the estimates value of β_2 in the regression equation 6.4, the effect of D_2 has been found to be highly significant in the case of aggregate male and female students.

Variable: Source of income

Here we have considered three different main sources of income of the parents of the students, which represent three different characteristics of the dummy variable - source of income. These are service (government and private), business, farming and others

represented by three dummies – D_{4i} , D_{5i} and D_{6i} . The reason behind considering sources of income as an explanatory variable is that the attitude of the parents towards enriching the academic performance of their off-springs depends on the nature of their services or the main source of living, which gives them security, dignity, cognizance of their status in the society. The coefficients of D_{4i} , D_{5i} , D_{6i} are respectively -2.495, 4.273 and 32.287 respectively. Interestingly, mean of the learning achievements score of the students whose parents are employed in some private or government organization is lower by 2.495 than the mean score of those students whose parents are not even M.P. passed by 2.495 and that of the students whose parents whose main source of income is business is higher by 4.273 than that of the students whose parents are not even M.P. passed. But, however, these two coefficients are not statistically significant at our chosen level of significance, which is 10% on the contrary, the coefficient of D_{6i} , i.e. β_6 , is 32.287 and this coefficient is significant at .000 level of significance. This result can be interpreted in the following way:

The farmers' children could improve their school learning performance very significantly over the learning achievement score of those students whose parents work as casual labour, rickshaw puller, small businessman, or as workers in unorganized sectors. The amount of improvement for these students has been much greater than that of the students whose parents' main source of income is service or business. Some very likely reasons might be that they might have been encouraged by their fathers to concentrate on their studies so that they get good jobs in future and their future get secured and they do not have to do hard works like – farming and other works on their agricultural and farm lands to earn a minimum living. The sorry state of their parents' living might have inspired them to do well in the school examinations. On the other hand, the parents of the students who do business or work in the private or government offices may not bother about their children's education at such lower level as standard 1 to 4. They are more concerned about how their children would go to

school, what tiffin they would give to their children etc. They believe that when the time will come, they would do whatever is necessary for the wellbeing of their children .At this stage of education, the parents of these children are perhaps more concerned about the health and comfort of their children. So far as this variable is concerned , the behavior of this variable – source of income, farming, there is no influence on the LAS's of the male and female on the one hand , and female on the other .

Variable: Infrastructure of the school

The infrastructure index has been classified into three grades, namely, excellent, very good, good and almost not existent. These characteristics are represented by D_{7i} , D_{8i} and D_{9i} . The bench mark dummy here is infrastructure that is almost non-existent. Here the influence of the characteristics such as excellent (Ex), very good (VG) and good (G) is measured by the estimated values of D_{7i} , D_{8i} and D_{9i} , i.e. β_7 , β_8 and β_9 and their levels of significance respectively.

We see from (6.4) that the estimated values of β_7 , β_8 and β_9 are respectively 25.523, 22.395, and 13.554. All are positive and statistically significant at 1.8%, 0.4% and 4.6% respectively.

These results may be interpreted in the following way:

The excellent, very good, good infrastructures of the schools significantly improve the learning achievement score of the students of the schools by 25.533, 22.395 and 13.554 over that of the students of the schools where the infrastructure is almost not existent. All these coefficients (β_7 , β_8 and β_9) are all statistically significant. These results, therefore, very clearly point to the major role of infrastructure in comparison to many other factors considered in our study.

A comparison to the effect of school infrastructure which is of excellent grade on the LAS's of female students as obtained of the regression equation for female students, which is significant at 8.7 %, the effect of this variable for the female and male taken together is

highly significant ($p=22.6\%$). Thus, the female students benefit from the excellent school infrastructure which is not evident in the case of the students-both male and female. The aggregative data hide many valuable information which are clearly evident in the disaggregative data points.

Variable: Private coaching

Many students take private coaching from the tutors appointed by their parents or guardians with a view to improving their learning achievement score. But our results indicate that it does not play any role in improving the students' achievement score. Interestingly its influence is found to be negative, though it is not statistically significant ($\beta_{10} = -3.905$ and $p = 0.145$), therefore, it is of no value. This results contradicts our general perception that students receiving coaching by tutors do better than those who do not. How can one explain this paradox? Several reasons may be cited to explain this paradox. One may be that almost all the students' get home coaching, thus showing small variations in the LAS of the students caused by this variable. Another reason may be that the coach or the parents in general do not attach that much of importance to private coaching though they pay good amount to the private tutor. The tutors, though not all, but a good number of them may take it very casually. In either case, aggregative or disaggregative this variable plays innocuous role. In both the aggregative case of male and females, this variable does not play any significant role in influencing the LAS's of the students.

Variable: School management

By school management we mean public school and private school. The public schools are managed and financed by the government (mostly state government, at some places they are managed by local self-government like municipalities, panchayats etc.). The private schools, on the other hand, are managed by private entrepreneurs and financed by themselves. These

schools are, however, granted permission of the state government which sometimes imposes restrictions on school management, issuing guidelines for fixation of tuition fees etc.

As far as the role of management of the school is concerned, our results obtained from the estimated regression equation (6.4) show that the students' learning achievement scores do not significantly differ between private and public school. The value of the coefficient of school management variable D_{11i} is 8.224 and it is found to be significant at 14.5%. Therefore, the null hypothesis, $H_0: \beta_{11} = 0$ is accepted which means the effect of school management is not significant.

Variable: Location

As already discussed, the location variable refers to the area where the school under consideration is situated. It may be a rural or an urban area. The urban area is represented by D_{12i} , where $D_{12}=1$ if the school is in urban area, $D_{12}= 0$ otherwise and its coefficient is β_{12} . The value of β_{12} is 17.515 and it is significant at 0.02%. As $\beta_{12} = (\beta_1 + \beta_{12}) - \beta_1$, then β_{12} , measures the learning achievement score of the students of the schools located in urban areas, which is higher by 17.515 than the learning achievement score of the students of the schools in the rural areas. Thus, the variable 'location' is found to have played a very important role in influencing the learning achievement scores of the students of the urban schools. The effect of this variable does not vary between the aggregative and disaggregative sample students.

Variable: Caste

The dummy variable caste has two characteristics- general and others. This variable is represented by D_{13i} and its coefficient is β_{13} . The regression result shows that $H_0: \beta_{13} = 0$ is accepted and $H_1: \beta_{13} \neq 0$ is rejected. This means that there is no effect of caste on the learning

achievement score of the students. This applies to both the aggregative and disaggregative data only.

Variable: Attendance

Attendance of the students in their classes-the variable has been assigned two characteristics:

1) attendance 75% and above, represented by D_{15} with coefficient β_{15} and 2) attendance between 60% and 75%, which is represented by D_{16} with coefficient β_{16} . β_{16} is the differential slope of the D_{16} variable. Its value is 20.25 while the value of β_{15} , the coefficient of D_{15i} , is 15.29. From the equation (6.4), we see that $H_0:\beta_{15}=0$ against $H_1:\beta_{15}\neq 0$ is accepted while $H_0:\beta_{16}=0$, against $H_1:\beta_{16}\neq 0$ is rejected. These results, therefore, show that the learning achievement score of the students whose class attendance is 75% or more is not significantly different from that of the students whose attendance is below 60%. On the other hand, the learning achievement score of the students whose class attendance is between 60% to 75% is greater than that of those students whose class attendance is below 60%. The results in the first case that is for D_{15i} appears unconvincing, contradict our general expectation, while the second result appears all right.

Till now we have considered the dummy social variables that are supposed to have some effects on the learning achievement score of the students.

Now we consider one quantitative variable, that is the income of the parents or the head of the family. This variable, as our regression results show, does not have any significant effect on the students' learning achievement score.

To conclude this section we may make the following observations on the cause and effect relationship between the students' learning achievement and several qualitative and one quantitative variable, i.e. annual income of the parents or the head of the family of the students, on the basis of the regression results.

The variables that have significant influence on the female students' learning achievement score are, namely, 1) Farm income 2) infrastructure (all characteristics) 3) location of the school. Considering these variables/characteristics, the following variables are seen to have significant positive effect on the students' learning achievement score. These are: D_{6i} , D_{7i} , D_{8i} , D_{9i} , D_{12i} and D_{16i} . the only one variable is the parents academic qualification stands out as the only influential variable which makes the difference between the LAS's of the whole set of male and female students taken together and the LAS's of the female students. It significantly influences the LAS's of the students when we take them (male and female) together, but it does in the case of the female students. This means that the LAS's of the male students has been significantly influenced by their parents' academic qualifications. That the parents' academic qualification has not influenced LAS's of the female students may be interpreted as the effect of no change in the parents' attitude toward 'good' or 'quality' education of the female child.

6.4 We now develop and present below the regression model for **male students**. The objectives of developing this model is to identify the factors / variables that have significant influence on the dependent variable – the learning achievement score of the male students and to find out those variables which explain the differences in the variations in the dependent variables - the learning achievement scores of the female and male students.

In other words our objectives are just to check whether the same set of independent variables as in the case of female students, discussed in the preceding Section 6.3, can explain the variation in the dependent variable, Y_i and if so, how much the model explain it. If another set of variables, which may include some of the elements of the previous set, explain the variations in the learning achievement score of the male students what are they? Explain the variations in the learning achievement score of the male students what are they? Even if the same set as in the case of female students or just some elements in the set of female students

explain the variations in the LAS of the male students, are there any difference in the levels of significance of the coefficients of the common variables? So , what are these variables and whether the coefficients of these variables are significant and if significant, what are the probabilities at which they are significant and whether there are any variations in this levels of significance of the common variables? The model will identify these variables.

From the analysis of the model it will be also clear which variable exercise influence upon the dependent variable, and by how much amount be in comparison to the explanatory power of the individual variables in the model for female students as determined by the levels of significance at which the coefficients of the variables are significant.

$$\text{Model: } Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 D_{4i} + \beta_5 D_{5i} + \beta_6 D_{6i} + \beta_7 D_{7i} + \dots + \beta_{12} D_{12i} + \beta_{14} D_{14i} + \beta_{15} D_{15i} + \beta_{16} D_{16i} + \gamma_2 X_{2i} + u_i \quad (i = 1, 2, \dots, 78) \dots \dots \dots (6.5)$$

The estimated regression equation is given by (6.6)

$$Y_i^{\wedge} = 16.87 + 16.11 D_{2i} (\text{Graduate and above}) + 8.87 D_{3i} (\text{H.S/M.P}) + 19.89 D_{4i} (\text{Service}) -$$

t= 1.849	2.002	1.481	1.403
p= .069	.050	.144	.166

$$0.117 D_{5i} (\text{Business}) + 23.21 D_{6i} (\text{Farming}) + 5.416 D_{7i} (\text{Infraexcellent}) + 16.01 D_{8i} (\text{InfraV good}) +$$

t= -.022	3.952	0.519	2.011
p= -.983	.000	.606	.049

$$21.41 D_{9i} (\text{Infra good}) + 6.552 D_{10i} (\text{Private tution}) + 5.24 D_{11i} (\text{Private school}), +$$

t= 2.862	1.304	0.738
----------	-------	-------

P= .006 0.197 .463

26.97D_{12,i} (Urban school)+ 7.58D_{14,i} (Caste general) + 11.43D_{15,i}(Attendance V good)–

t= 4.48 1.679 1.50

P= .000 .098 .136

1.615D_{16,i}(Attendance good)- 3.69X_{2i}

t= -.236 -1.062

P= .814 .292

R²=.584; \hat{R}^2 =.484; D.W =1.463; F= 5.812; p=.000; df=N-K=(78-16)=62.....(6.6)

The following variables are found to have significant influence upon the dependent variable, Y_i, i.e. the learning achievement score of the male students. These variables are: D₂, D₆, D₈, D₉, D₁₂ and D₁₄. These notations represent the variables –D₂- academic qualification of the parents, which is graduate or higher level; D₆ – main source of income-farming; D₈ - infrastructure – with characteristic ‘very good’; D₉ – infrastructure – ‘good’, D₁₂ – location(urban), and D₁₄ – caste (general). All the estimated coefficients are positive and significant at less than 10% probability level. Out of these six variables as many as five variables have significant effect on the dependent variable at 5% or less than 5% probability level. The coefficient of just one variable-general caste-is significant at 9.8% probability level.

From a comparison of the results of this model (6.6) with (6.4), we make the following observations:

In model 6.4 where the dependent variable is learning achievement score of the female students, the independent variables that have statistically significant influences on the dependent variable, are namely,

D₆ (Farming main source of income); D₇ (excellent infrastructure); D₈ (very good infrastructure); D₉ (good infrastructure); D₁₂ (location-urban), and D₁₅ (attendance of the students).

The explanatory variables that are common to the estimated models (6.4) and (6.6) are D₆ (main source of income, farming), D₈ (very good infrastructure), D₉ (good infrastructure) and D₁₂ (location-urban). The other two variables D₂ (academic qualification-graduate and above) and D₁₄ (caste-general) have significant effect on the dependent variable in the case of the male students, while D₇ (excellent infrastructure) and D₁₆ (attendance-good (60%-75%)) are the variables having significant influence on the dependent variable in the case of the female students. The explanatory power of the variables common to both the equations are almost the same. The slight differences that exist in such cases do not matter much. But what is important is that while the caste factor appears as a significant factor (in the sense of having statistically significant effect ($p=.098$) on the dependent variable) in the case of male students, it, however, does not play that much significant role in influencing the learning achievement score of the female students ($p=.110$), which however we do not accept according to our specification of level of significance of 10%. The difference in the effects of the caste variable between female and male students on their respective LAS's may have arisen because the male students appear to be caste sensitive, but no so are the female students. The female students do not cherish any low caste feeling which the male students feel so, may be because of the frequent vulgar insinuation towards low caste students by the

general caste students. However, we can dismiss all these arguments justifying the difference between the effects of the caste on the LAS of the so called low caste and high caste, if we set the level of significance at 5% or less.

Table – 6.1:

Factors having significant impact on the LAS of the female and male students.

Female students			Male students		
Significant variables	Level of sig		Significant variables	Level of sig	
Main source of income farmer	0.000		Academic qualification of parents(graduate & above)	0.050	
School infrastructure excellent	0.087		Main source of income farming	0.000	
School infrastructure very good	0.004		School infrastructure very good	0.049	
School infrastructure good	0.046		School infrastructure good	0.006	
Location urban	0.002		Location urban	0.000	
School attendance of the students	0.043		Caste General	0.098	
R ² =.514, Adjusted R ² = 0.342, F=4.229, p=.000			R ² =.584, Adjusted R ² = 0.484, F=4.812, p=.000		

6.5 Factors explaining the difference in LAS the Rural and urban students

From the estimates regression equations 6.2, 6.4 and 6.6, we have seen that the location of the schools – rural or urban plays a very significant role in influencing the students’ learning achievement score. It is found that the learning achievement scores of the student of the urban schools is higher by around 18 to 26 units than the students of the schools located in rural areas. and the coefficients of the variable-urban location has been highly significant at less than 0.02% probability level. This special role of the variable prompts us to find the reason behind it. The reason can be unravelled by identifying the factors that cause improvement in the learning achievement score of the students of urban schools and the factors that are responsible for lower score of the students of the rural schools.

To find the factors responsible for difference between two locations – rural and urban, we have run two regression models using all the independent variables D_{2i} through $D_{16,i}$, barring D_{12} i.e. the location variable.

The estimated results of the regression equations for the LAS's of the students of urban and rural schools study are given below:

\hat{Y}_i = Estimated values of the dependent variable - learning achievement scores of the students of the schools (primary) in both the urban and rural locations.

Table 6.2

Results of Regressions of LAS's Students in Rural and Urban students

	Urban			Rural		
Model	coefficients	t	sig	coefficients	t	sig
constant	40.576	4.930	.000 (S)	34.246	2.134	0.037 (S)
Academic qualification -						
graduate and above (D2)	10.307	0.853	0.397	30.289	3.734	.000 (S)
H.S/M.P (D3)	6.185	0.787	0.434	13.193	2.883	.005 (S)
Source of income						
Service (D4)	-6.109	-	0.526	22.457	1.695	.095 (S)
		0.637				
Business (D5)	0.755	0.135	0.893	7.423	1.18	0.242
Farming (D6)	-	-		31.461	5.912	.000 (S)
Infrastructure -						
Excellent (D7)	16.258	1.906	.061 (S)	-	-	-
Very good (D8)	18.377	2.169	0.034 (S)	13.911	1.754	.085 (S)
Good (D9)	8.468	0.949	0.346	19.157	3.335	.001 (S)
Private coaching (D10)	-3.307	-	0.374	2.443	0.589	0.558
		0.896				
Organization (school management) (D11)	9.539	1.454	0.151	-8.311	-1.223	0.226
Gender (D13)	2.715	0.59	0.557	0.665	0.180	0.858
Caste (D14)	-9.337	-	0.085 (S)	9.312	2.225	.030 (S)
		1.751				
Attendance						
Very good (D15)	16.504	2.245	0.028 (S)	-9.894	-0.706	0.483
Good (D16)	9.821	1.732	0.088 (S)	-17.757	-1.234	0.222
Income of the household (X2i)	1.224 E5	0.68	0.499	1.695 E-5	0.447	0.656
			$R^2 = .485, \bar{R}^2 = .371, F = 4.237,$	$R^2 = .641, \bar{R}^2 = .559, F = 7.786,$		
			DW = 1.425, N=78, p = .000	D.W.=1.581, N=76, p = .000		

School Infrastructure Index:

In the case of the urban school students, the learning achievement score is influenced by only five factors, namely, infrastructure (two categories –excellent and very good) of the schools, caste and attendance (two categories –very good and good). All of them are statistically significant. The positive values indicate improvement in the learning achievement scores of the students and the negative values indicate fall in the achievement score. For example, the variables, namely excellent infrastructure and very good infrastructure of the schools improve the achievement score by 16.26 and 18.38 units respectively over the achievement score of the students of the Schools that have unsatisfactory infrastructural facilities.

It appears puzzling that the learning achievement scores of the students of the schools with excellent infrastructural facilities are higher by 16.26 than those of the students of the schools with unsatisfactory infrastructure but lower by 2.12 units than that achieved by the students of the schools with very good' infrastructure (18.38). This apparently unacceptable result can be explained by the fact that the “excellent” infrastructure includes those facilities which focus on the all-round development of the students, in addition to enhancing the learning achievement scores of the students. The cultural facilities that include sports, music, recitation, storytelling, general knowledge tests etc. offered by some schools aim at developing the overall intellect of the students. These students focus on improving their learning achievement score as well as on development of other mental skills and physical strength that are expected to stand them in good stead in their future pursuits. On the other hand, the students of the schools offering “very good” infrastructure focus mainly on the enhancement of their performance in the school exams, relatively much less on the development of other faculties of their own. This fact may be accepted as a cogent explanation for the difference in the extent of increase in the students learning scores between

two sets of students using “excellent” and “very good” infrastructural facilities offered by the schools.

Caste variable:

The caste variable is seen to have negative influence on the achievement score of the general caste people. The coefficient value of -9.337 means that the score of the general caste students is lower by 9.337 units than achievement score of the non-general caste people. This result appears unacceptable. However, as the significance level is very high, as high as 8.5%, well above 5% probability level, we may not accept this result, though we have taken up 10% as the maximum risk level as a safe one. However, here we accept the hypothesis that $H_0:\beta_{14}=0$ against the alternative hypothesis $H_1:\beta_{14}\neq 0$, do not reject H_0 and accept H_1 , because the acceptance of $H_1:\beta_{14}\neq 0$, i.e. statistically significant, defies our general conception that general caste students normally perform better than or as good as the non-general caste students. It is very unlikely to happen that the performance of the general caste student is worse than that of the non-general caste students.

Attendance of the students:

Attendance, both very good (75% and above) and good (60% - 75%), makes the achievement score of the students to go up by 16.504 and 9.821 units respectively over that of those students whose attendance is poor, i.e., less than 60%. In other words, the learning achievement score of the students of the schools which is greater by 16.50 units than that of the students with very poor attendance (less than 60%). Similarly, the students with good attendance (60% to 74%) achieve learning achievement score which is greater than that of the students with poor attendance by 9.82 units. These results very clearly show the importance of (higher percentages) of attendance in explaining the variations in the learning achievement score of the students.

Now we consider the LAS of the students in the rural areas. The factors that have significant influences on the LAS of the rural students are, namely 1) academic qualification of the parents 2) source of income: farming 3) school infrastructure (only 'very good' and 'good', and 4) caste. Of these five variables significantly influencing the LAS's of the students of rural schools, the three variables such as the academic qualifications of the parents (both the characteristics), and source of income-farming have insignificant influence on the LAS of the school students in the urban school. In rural areas as the parents of many school going children are economically poor, they do not have education, only few may have good education. Therefore they understand the importance of education and encourage their children to become serious and sincere in their studies, and they do well in the school exams. In the urban areas, on the other hand, almost all parents of the school going children are educated and therefore, this academic qualification variable with two characteristics do not have much variations in it. Therefore, its power of explaining the variation in the LAS's of their school-going children is almost nil. These explain the statistical insignificance of the academic qualifications variable explaining variations in LAS of the urban school children.

Variable: School infrastructure

The main cause of the better performance of the students of the schools situated in urban area appears to be the infrastructure facilities offered by the school authorities to the students in general. In many of the schools in the rural areas in urban areas the infrastructure is very good in quite a good number of schools in the rural areas it is not satisfactory.

In rural schools, no school enjoys excellent infrastructural facilities, though some of them can offer good infrastructural facilities. For them, its effect is statistically very significant (significant at .01% probability level).

Variable: Caste

In addition to these factors, in the rural schools, caste is an important factor. The caste difference in the schools in rural areas is deeply felt. In the urban areas, the students normally do not bother about caste factor, while in the rural areas the caste consciousness is quite prominent.

Variable: Attendance

Attendance is another important factor that largely accounts for the relatively higher learning achievement score of the students in urban area schools. In rural schools, none of the two characteristics of the attendance variable are significant. A glimpse at the right hand panel of the above table makes it clear to us. For example, attendance does not influence learning achievement score in the rural schools (none of the coefficients are significant). The levels of significance for two different ranges of attendance are 48.3% and 22.2% respectively, whereas in the case of urban schools they are 2.8% (for attendance 75% and above) and 8.8% for attendance of 60% to 70%.

There are a few other factors accounting for the differences in the achievement scores of the students of the schools in urban and rural areas. For example, level of education of the parents in the rural areas also matters a lot. Parents with higher education levels take special care of their school going children as compared to the low educated parents. Here in the rural areas the parents know each other, their caste, their education levels, income level etc. Therefore, the caste, education, income consciousness influence these parents in taking special care of their children. The highly educated parents might like to show the lesser educated parents their special dignity in the society.

6.6 Factors explaining the Difference in the LAS between General Caste and Non General Caste student

Since long the economists, sociologists and the government administrations and many scholars are worried about the slow progress of the development of the socially and economically under-privileged persons, mainly those belonging to SC, ST and OBC section of the population despite the fact several policy measures have been implemented to improve the socio-economic conditions of these people. Here in our present study we consider one particular aspect of socio-economic development of these people, that is, the education of their children and their academic performance. That is, do their children go to school regularly and if they go, how do they perform in school tests and examinations, and how and what hindrances and advantages do they face to make sure their educational development?

The questions whether the children of these under-privileged people go to schools on a regular basis and how they perform in the school exams and tests conducted from time to time by the school authority and the school teachers have been dealt with in the enrollment section of the Chapter 5 and the students' learning achievement scores (LAS) discussed in chapter 6.

Here in this section we try to identify the factors that act as hindrances and advantages or opportunities that pull down or pull up their learning achievement scores(LAS) of these two groups of students – general caste and non – general caste. For identification of these factors we develop a model where the LAS is the dependent variable and the same sets of variables – one set consists of dummy variables and other set of one qualitative variables – as have been considered in the above models involving different social and demographic aspects of the students such as male – female and rural and urban.

. This analysis will help us to identify the factors that need to be augmented or checked to improve the learning achievement scores of these two – socially privileged and socially disadvantaged – categories of primary school children.

6.6.1

Identification of Factors Determining LAS of General Caste Student

We first consider the case of general caste students. The model developed for this category of the students is given below. The variables have already been included in the models presented and discussed above.

$$y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 D_{4i} + \beta_5 D_{5i} + \beta_6 D_{6i} + \beta_7 D_{7i} + \beta_8 D_{8i} + \beta_9 D_{9i} + \beta_{10} D_{10i} + \beta_{11} D_{11i} + \beta_{12} D_{12i} + \beta_{13} D_{13i} + \beta_{15} D_{15i} + \beta_{16} D_{16i} + YX_{2i} + u_i \dots\dots\dots(6.7)$$

The estimated regression equation obtained by applying the least squares method to the data sets pertaining to this above variables is given below.

$$y_i = 37.087 + 13.711 D_{2i} (\text{graduate}) + 9.8981 D_{3i} (\text{H.S/M.P}) + 6.448 D_{4i} (\text{service}) + 15.357 D_{5i}$$

t=	3.271	2.030	1.983	.808	2.532
----	-------	-------	-------	------	-------

p=	0.002	0.046	.051	.423	0.014
----	-------	-------	------	------	-------

$$+ 22.132 D_{6i} (\text{farming}) + 26.685 D_{7i} (\text{infra ex}) + 19.192 D_{8i} (\text{infra very good})$$

t=	4.390	3.218	2.301
----	-------	-------	-------

p=	.000	.002	.024
----	------	------	------

$$\begin{aligned}
& + 18.989 D_{9i}(\text{infra good}) + 2.032 D_{10i}(\text{private coaching}) - 1.567 D_{11i}(\text{private school}) \\
& t= 3.335 \qquad \qquad \qquad 0.548 \qquad \qquad \qquad -0.215 \\
& p= .001 \qquad \qquad \qquad .586 \qquad \qquad \qquad .831 \\
& -7.011 D_{12i}(\text{Urban school}) - 4.852 D_{13i}(\text{Gender/Girl}) + 1.617 D_{15i}(\text{Attendance very good}) \\
& t = -1.133 \qquad \qquad \qquad -1.253 \qquad \qquad \qquad .154 \\
& p= .261 \qquad \qquad \qquad .215 \qquad \qquad \qquad .878 \\
& -6.898 D_{16i}(\text{Attendance good}) + 2.473 E-5 X_{2i}(\text{Annual household income}) \quad \text{----- (6.8)} \\
& t= -0.661 \qquad \qquad \qquad 1.514 \\
& p= 0.510 \qquad \qquad \qquad 0.135 \\
\mathbf{R^2 = .497, \bar{R}^2 = .387, F= 4.612, P = .000, DW= 1.786}
\end{aligned}$$

Y_2 = Learning achievement score of the general caste students of the schools under consideration.

$D_2, D_3, \dots, D_{13}, D_{15}, \dots, D_{16}$ and X_2 have been defined above . The regression results shows that the independent variables $D_2, D_3, D_5, D_6, D_7, D_8$ and D_9 have the significant bearing upon the learning achievement score of the general caste students. In other words, D_2 and D_3 representing the academic qualifications of the parents of the students under consideration - graduates and higher degree holders, have significant positive influence on the students' learning achievement score. The results show $\beta_2 \wedge = 13.721$ and $\beta_3 \wedge = 9.898$, which mean that the children of the parents with graduate and higher academic degrees get 13.721 units more and the children of the HS and MP pass certificate holders get 9.898 units more than the children of the parents with lower than M.P. certificate holders.

D_4, D_5 and D_6 represent different categories of sources of income namely, service, business and agricultural farming. All the sources are found to have significant positive influence on the learning achievement scores of the general cast students. In other words, incomes of the

general caste children, in general, irrespective of their sources, have significant influence on the learning achievement scores (LAS) of their children.

Infrastructure of the school is found to have exercise significant positive effect on the improvement in LAS, irrespective of the levels of infrastructure; i.e., whether it is excellent, very good or good ,over the LAS of the students of the schools with unsatisfactory infrastructural facilities. This result thus clearly establishes the role of infrastructure in the improvement of LAS of the general caste students. The infrastructure represented by its different categories, are denoted by D7, D8 and D9 – representing excellent, very good and good respectively. The results show that the LAS of the students of the school offering excellent, very good and good infrastructure are higher by 26.86, 19.19 and 18.99 units than the LAS of the students of those schools which have unsatisfactory infrastructure. None of the other independent variables are at all relevant in the sense that they do not have significant effect ($p > .10$) on the Learning achievement score(LAS) of the students belonging to the general caste .

At the end , we can very confidently infer that the variables such as academic qualifications of the parents of the students , incomes of the parents from farming and infrastructure of the schools at all levels (excellent, very good and good) play very significant role in the improvement of the Learning achievement score (LAS) of the students.

6.5.2 Determining the Factors Explaining the Differences between Variations in LAS's of Non-general Caste Students

Now we make an attempt to identify the factors that cause rise or fall in the Learning achievement scores (LAS) of school students belonging to the non – general castes like SC, ST and OBC. The same sets of variables are chosen as the most likely explanatory variables. Let us examine how many of them play significant role in explaining the variation in the LAS of the Non-general students.

The estimated regression equation is given below.

$$y_i = 12.826 + 26.818D_{2i}(\text{graduate}) + 16.946D_{3i}(\text{H.S/M.P}) - 11.06 D_{4i}(\text{service}) - 3.901D_{5i}$$

$$t = 1.463 \quad 2.297 \quad 2.615 \quad -0.888 \quad -0.755$$

$$p = 0.149 \quad 0.026 \quad 0.012 \quad 0.379 \quad 0.454$$

$$(\text{business}) + 42.312 D_{6i}(\text{farming}) + .981 D_{7i}(\text{infra ex}) + 11.142 D_{8i}(\text{infra very good})$$

$$t = 5.033 \quad .093 \quad 1.411$$

$$p = .000 \quad .926 \quad .164$$

$$+ .592 D_{9i}(\text{infra good}) - .771 D_{10i}(\text{private coaching}) + .092 D_{11i}(\text{private school})$$

$$t = .054 \quad -0.201 \quad .015$$

$$p = .957 \quad .842 \quad .988$$

$$+ 38.148 D_{12i}(\text{Urban school}) + 6.076 D_{13i}(\text{Gender/Girl}) + 18.399 D_{15i}(\text{attendance very good})$$

$$t = 7.197 \quad 1.527 \quad 2.739$$

$$p = .000 \quad .133 \quad .008$$

$$+ 7.556 D_{16i}(\text{attendance good}) - 2.643 \times 10^{-5} X_{2i}(\text{Annual household income}) \quad \text{----- (6.9)}$$

$$t = 1.348 \quad -0.637$$

$$p = .183 \quad .527$$

$$R^2 = .704, \quad \bar{R}^2 = .619, \quad F = 8.263, \quad P = .000, \quad DW = 1.278$$

From the results of the above regression equation the following variables are found to have very highly significant influence on the students' LAS. These variables are: academic qualification of the parents of the students (D_2 and D_3), source of income - farming (D_6), location (D_{12}), and D_{15} (school attendance of the students-very good)

In respect of academic qualification of the parents, the results show that the school going children of the parents with graduate or higher academic qualification achieve higher LAS by 26.82 units than the children of the parents with lower than MP certificate holders and the school going children of the parents with HS/MP certificates achieve higher LAS by 16.95 units (but lower than the children of the graduates or above) than the children of the parents with lower- than-the MP standard. This clearly demonstrates that the academic qualification of the parents of the students play a very significant role ($P < 0.03$)

The children of the parents using 'farming' as their main source of income perform much better than the children of the parents using very nominal types of work with no assured regular income like others earning mainly from service, business and farming.

Most of the parents of the SC, ST and OBC students earn their incomes mainly from farming. Very few of them do business or service. Therefore, the effect of the incomes from business or service on the LAS of the non-general students is not at all (statistically) significant ($p = 0.379$ for service holders and $p = 0.454$ for the businessmen).

The value of the coefficient of D12 (location Urban =1, otherwise=0) is 38.15 and this value is significant at less than 1% probability level. This signifies that the non-general students studying in the urban schools achieve much higher scores (higher by 38.15 units) than their counterparts in the rural areas. On the other hand, the LAS of the general caste students is not influenced by the location factor (see Equation 6.8).

School attendance of the attendance (D_{15}) plays an important role in improving the LAS of the non-general caste students, which is, however, found to have no impact on the LAS of the general students. This apparent puzzle may be explained by the fact while in the urban location of the schools, there is not variation in the percentage of student attendance of the SC, ST and OBC students, there is wide variation in the percentage of attendance of students

of the schools in the rural areas. We now summarize our findings in a tabular form on the factors explaining the variation in the LAS of the general and non-general caste students.

Table 6.3:

Factors having significant impact on the LAS of the general caste and non general caste students.

General Caste Students		Non-General Caste Students(SC+ST+OBC)	
Significant variables	Level of sig	Significant variables	Level of sig
Academic qualification graduate (D ₂)	0.046	Academic qualification graduate (D ₂)	0.026
Academic qualification Secondary and HS	0.051	Academic qualification Secondary and HS (D ₃)	0.012
Main source of income : business (D ₅)	0.014	Main source of income : farming (D ₆)	0.000
Main source of income : farming (D ₆)	0.000	Location of the school : urban (D ₁₂)	0.000
School infrastructure excellent (D ₇)	0.002	School attendance very good (70% and above)	
School infrastructure very good (D ₈)	0.024	(D ₁₅)0.008	
School infrastructure good (D ₉)	0.001		
R ² =.497, Adjusted R ² = 0.387, F=4.612, p=.000		R ² =.704, Adjusted R ² = 0.619, F=8.263, p=.000	

Common Factors: Academic qualification of the parents(graduate &above) , Academic qualification of the parents(MP/HS), Main source of Income (Farming)Infrastructure & Caste of the students

For General Caste Students: Infrastructure (excellent),Infrastructure (Very good), Infrastructure(good), Main source of Income (Business)

For Non-general Students: Location (urban), School Attendance

The estimated values of the coefficients (in figure) are in the parentheses

From Table (6.3) while infrastructure is, inter alia, an important factor influencing LAS of the general caste students, it has no impact on the non-general caste students. It happens mainly because these students either do not use the infrastructure facilities in the way and as much as their counterparts belonging to the general caste group do or they are not at all available in the rural schools where the percentage of the SC, ST and OBC students are relatively higher than that of the general caste students in the schools located in rural areas. The non-general caste students here in our study belong to both urban and rural areas. However in our case, these students come from rural areas. This clearly demonstrates that development of infrastructure of the school in the rural areas will undoubtedly improve the LAS of the non general caste students.

School attendance is also an important factor influencing the LAS of non general caste students. It happens because the school attendance factor has much variation for the non-general caste students while for the general caste students the variation in their percentage of attendance is not that wide as to affect the variation in their LAS. So the parents of the students not attending school classes must be given necessary helps and advice to send their children to schools. As regards the common factors affecting LAS's of the general caste and non-general caste students, academic qualifications of the parents of the non-general caste students have relatively much greater influence (26.82 and 19.95 against 13.71 and 9.90 of the general caste students) on the LAS's of the non-general caste students. Similarly, the income from the farming explains a relatively large variations in LAS of the students of the non-general caste students (43.31 against 22.13 for the general caste students) .

Variables Found Irrelevant

1. Impact of the school management

After scrutinizing the result presented in equations 6.2, 6.4, 6.6, 6.7, 6.8 and 6.9, it may be concluded that the school management (i.e. whether the schools are private or public) has, not any significant impact on the learning achievement scores of the students whether they are male or female, general caste or non-general caste students, and whether the schools are located in rural or urban areas.

2. Private coaching

Against our general belief that private coaching of the primary students enhances the students' learning achievement score, our regression results show unequivocally that it has no significant impact on any of the following categories of the students – male or female, urban or rural students, general or non-general caste students. The likely reasons have been discussed above.

3. Income

Income of the parents does not have any significant impact on the LAS of the students but classified into different categories as mentioned above. This variable may play an important role in the case of the students pursuing higher education. We used this variable as explanatory variable on the supposition that the parents which higher incomes may be more inclined to look after their children's performance in schools as they are logically expected to give their children higher education in future, and to be satisfied with good performance in school education from the very beginning. The unimportant role of the household's income in influencing the learning achievement scores of their school-going children may be reasoned that the parents, particularly parents belonging to middle and high income groups, spend a very small fraction of their income on the education of their children. Even the parents

earning more or less enough income to somehow maintain their families do not have to bother about spending on their primary school-going children's education. In most of the cases, particularly in the public schools, the school fees are minimal. In the private schools, though the school fees are quite high, the parents who send their children to these schools, can well bear these expenses quite comfortably.