# Statistical Analysis of Socio-Economic Variables Responsible for Wastage in Elementary Education in Rajasthan, India

**Gupta Manoj Kumar[1], Prasad Jagdish[1] and Verma J.P.[2]**
[1]Department of Statistics, University of Rajasthan, Jaipur, Rajasthan INDIA
[2]Department of Computer Science and Statistics, Laxmibai National University of Physical Education, Gwalior, Madhya Pradesh INDIA

## Abstract

*There are two goals of Sarva Shiksha Abhiyan (SSA) that universal enrollment and universal retention. Enrollment and Retention in elementary education system depends upon various educational and socio-economic variables. In this paper, a statistical analysis of enrollment and dropout rates those are factorized on various socio-economic variables referred to as factors here. The major findings revealed that engagement of children in agriculture work, poor economic condition of parents, engagement of children in grazing cattle and marriage of children in early age have influence over enrollment and dropout rates in Rajasthan. By controlling these four socio-economic variables, dropout rate can be decreased and enrollment rate can be increased to a great extent in Rajasthan. The data for enrollment and dropout rates and many other educational variables across Rajasthan for the year 2010 are considered.*

**Keywords:** Dropout, never enrollment, multiple regression, principal component analysis (PCA), multicollinearity, correlation matrix, eigen value, child tracking system.

## Introduction

Poverty and illiteracy are the two major evils for any society. Many researchers have already been proven that the poverty is the main cause of illiteracy[1] and others proven that illiteracy is one of the most important cause of poverty[2]. It is fact that literacy is generally related to reduction of poverty, economic growth and wealth creation in all over the world. High illiteracy areas correspond quite closely with high levels of poverty areas in the world[3]. Hence, the literacy is the key to unlocking the cycle of poverty[4]. The government of India has floated a large number of schemes since the independence to alleviate poverty but it has not made any real dent in the reduction of poverty. Conditional cash transfer programmes as in Mexico and Brazil where cash support is extended to the poor families conditional on children attending school, if adopted by India may result in reducing poverty and improving education that may help in sustaining the further growth[5]. It reveals that, poverty and illiteracy are dependent to each other. If one is eradicated from the society than other will automatically eradicate. If it is estimated that out of these two evils which can be eradicated by less efforts and resources, definitely illiteracy will be estimated.

Illiteracy can be eradicated by universal elementary education. The universal elementary education can be achieved by two major components namely universal enrollment and universal retention. First component (universal enrollment) of universal elementary education is directly related to out-of-school children and the second component (universal retention) of universal elementary education is directly related to in-school children. For this purpose, the children of the age group 6-14 years can be classified into two classes, viz those who are presently out-of-schools and those who are presently enrolled in schools. Hence, the aim of Sarva Shiksha Abhiyan (SSA) is to approach out of school children and enrolled them in schools as well as to retain the enrolled children in the schools. Non-enrollment (Never enrollment) and non-retention (Dropout) can be defined as-

Never Enrollment: Never enrollment means that school is available for the child to take admission but child could not reached to the school due to some socio-economic and educational factors.

Dropout: Dropout means that child enrolled in the school but leave school during any part of the academic year without completing the course in which he/she is studying. This is the pre-mature withdrawal of child from school at any stage before the completion of course for any reason, except death and transferring to another school.

The main challenge of universal elementary education seems universal enrollment than universal retention. To achieve the universal enrollment, it is to ensure that each and every child (of the age group 6-14 years) is enrolled in school[6]. Hence, in elementary education system, wastage due to never enrollment is very serious and must be controlled. Two methods which were used by researchers to find out dropout rate- one is cohort method and other is household census survey method[7]. In this paper, dropout rates are used which is calculated on the basis of Child Tracking System (CTS) household census survey method. Universal enrollment, retention and completion in elementary

education can only be achieved by improving quality of elementary education and financial conditions of society, especially for the lower section of the society[8]. In the Jharkhand state inadequate infrastructure and the lack of teachers affect the quality of teaching; poverty is responsible for the alarming rates of non-enrolment, dropouts and poor attendance of pupils[9].

Therefore, various variables are responsible for never enrollment and dropout in elementary education. These can be broadly classified in two classes - Educational Variables like number of schools, number of teachers, infrastructure available in schools etc and Socio-economic Variables like poverty, health, rural population, cast, religion, illiteracy in the society etc. Educational variables can be controlled by implementing different educational schemes by the Government like SSA but controlling of socio-economic variables is a very huge and difficult task for any government. Therefore, many schemes are required to implement to control these variables. In this paper, socio-economic variables are being considered so that some specific schemes can be emphasized. Educational variables are already studied by the authors[10]. Both educational and socio-economic variables are not studied simultaneously because the nature of available data is different.

## Research Methodology

**Sources of data:** Secondary data are used for 33 districts of Rajasthan for the year 2010. Various socio-economic variables which are described in the later sections, reduced into four components those are mainly responsible for never enrollment and dropout rates. Data were drawn from household census survey report conducted in the year 2010 by Rajasthan Elementary Education Council, Jaipur to develop a child tracking system for every child having the age between 6-14 years.

The data for enrollment and dropout rates and many other socio-economic variables across Rajasthan are considered for the year 2010 because of the following reasons: i. Since household census survey in Rajasthan was conducted in 2010, it was not conducted in subsequent years. ii. District-wise enrollment rate and dropout rates as published in CTS household census survey report are used because district-wise dropout rates are not available in DISE data and the children population of age of 6-14 year is required which is published in household census survey report.

**Variables under Study:** District-wise following variables based upon socio-economic factors as published in CTS household survey report 2010, are taken in this study[11].

Agriculture work: District-wise number of children those could not reach to the schools or dropped out from the schools due to their engagement in agriculture related works, are taken as one of the many explanatory variables.

Care of sibling: District-wise number of children those could not reach to the schools or dropped out from the schools due to their engagement in care of their siblings are taken as explanatory variable.

Grazing cattle: District-wise number of children those could not reach to the schools or dropped out from the schools due to their engagement in grazing cattle, are taken as explanatory variable.

Poor economic condition: District-wise number of children those could not reach to the schools or dropped out from the schools due to their parent's poor economic condition, are taken as explanatory variable.

Lack of educational facility: District-wise number of children those could not reach to the schools or dropped out from the schools due to lack of educational facilities in their area, are taken as explanatory variable.

Ignorance of guardian: District-wise number of children those could not reach to the schools or dropped out from the schools due to ignorance of their parents, are taken as explanatory variable.

Long illness: District-wise number of children those could not reach to the schools or dropped out from the schools due to long illness, are taken as explanatory variable.

Non friendly environment of school: District-wise number of children those could not reach to the schools or dropped out from the schools due to non friendly environment of their schools, are taken as explanatory variable.

Migration: District-wise number of children those could not reach to the schools or dropped out from the schools due to migration, are taken as explanatory variable.

Child marriage: District-wise number of children those could not reach to the schools or dropped out from the schools due to their marriage in early age, are taken as explanatory variable.

Regular absence: District-wise number of children those could not reach to the schools or dropped out from the schools due to their regular absence from the school, are taken as explanatory variable.

Homeless: District-wise number of children those could not reach to the schools or dropped out from the schools due to non-having of their home, are taken as explanatory variable.

Without adult protection: District-wise number of children those could not reach to the schools or dropped out from the schools due to without adult protection, are taken as explanatory variable.

## Results and Discussion

**Analysis:** In this paper, principal component analysis (PCA) and multiple regression techniques are applied by using SPSS to reduce the above mentioned variables into few components those are responsible for enrollment and dropout rates.

**Estimation of Dropout**: Preliminary Analysis: Determinant value of correlation matrix of all the variables under study was found 0.000000155. It should be less than the threshold value of 0.00001[12]. It means there is high multicollinearity exists between the variables. For improving the determinant value, two variables - care of sibling and non-friendly environment of school were eliminated from the study due to their high correlation with the variables agriculture work and long illness respectively. After removing these variables, the determinant value was found 0.0000279, which is greater than the threshold value. By this method, problem of high multicollinearity was removed and rest of the variables was moderately correlated with each other. The value of Kaiser-Meyer-Olkin (KMO) statistic was found 0.657, which is more than 0.5. It shows that used data is sufficient and no need to rethink about the variables under study. Bartlett's test was also done and p-value found highly significant ($p < .001$) therefore, there were some relationships between the variables are exists. By this analysis, it could be concluded that principal component analysis might be applied on these data.

Component Extraction: By using SPSS, four components extracted having eigenvalues greater than 1. Table-1 shows that all the four components explained 86.29 percent of the total variability. Before rotation, first component accounted for 34.07 percent of total variance than the remaining three components (25.93 percent, 17.08 percent and 9.21 percent respectively). However, after rotation method it accounted for only 25.25 percent of total variance compared to remaining three components (25.55 percent, 23.09 percent and 13.41 percent respectively). Rotation has the effect of optimizing the structure of component and one consequence for this data is that the relative importance of the four components is equalized[12].

Table-2 shows that it was assumed that each variable under study will explain by the underlying components with 100 % variability but 78.2 percent to 92.8 percent variability was explained by every variable in retained four components because some of the components are discarded after extraction.

Table-3 shows the orthogonal rotated component matrix of variables responsible for dropout which reflects the correlation coefficients between the variables under study and four components.

By using above matrix, the following variables were extracted from each component. Only those variables were considered for extraction which having the value of correlation coefficients more than 0.7.

**Table-1**
**Total Variance Explained for Dropout**

| Component | Initial Eigen values | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 3.747 | 34.066 | 34.066 | 3.747 | 34.066 | 34.066 | 2.777 | 25.247 | 25.247 |
| 2 | 2.852 | 25.927 | 59.993 | 2.852 | 25.927 | 59.993 | 2.700 | 24.549 | 49.796 |
| 3 | 1.879 | 17.080 | 77.073 | 1.879 | 17.080 | 77.073 | 2.539 | 23.085 | 72.881 |
| 4 | 1.014 | 9.214 | 86.287 | 1.014 | 9.214 | 86.287 | 1.475 | 13.406 | 86.287 |

**Table-2**
**Communalities of Variables Responsible for Dropout**

| Variables under Study | Initial | Extraction |
|---|---|---|
| Agriculture work | 1.000 | .782 |
| Grazing cattle | 1.000 | .909 |
| Poor economic condition | 1.000 | .842 |
| Lack of educational facility | 1.000 | .857 |
| Ignorance of guardian | 1.000 | .815 |
| Long illness | 1.000 | .869 |
| Migration | 1.000 | .905 |
| Child marriage | 1.000 | .918 |
| Regular absence | 1.000 | .928 |
| Homeless | 1.000 | .826 |
| Without adult protection | 1.000 | .842 |

**Table-3**
**Rotated Component Matrix of Variables Responsible for Dropout**

| Variables under Study | Component | | | |
|---|---|---|---|---|
| | **1** | **2** | **3** | **4** |
| Agriculture work | -.103 | -.054 | .872 | .087 |
| Grazing cattle | .932 | .000 | .127 | .155 |
| Poor economic condition | -.103 | .007 | .052 | .910 |
| Lack of educational facility | .281 | .044 | .844 | .252 |
| Ignorance of guardian | .409 | .362 | .348 | .629 |
| Long illness | .911 | .056 | -.145 | -.124 |
| Migration | -.372 | -.119 | .866 | -.054 |
| Child marriage | .023 | .939 | -.183 | .053 |
| Regular absence | .107 | .851 | .258 | .354 |
| Homeless | .182 | .879 | -.100 | -.096 |
| Without adult protection | .793 | .410 | -.213 | .000 |

**Table-4**
**Component-wise Extraction of Variables responsible for Dropout**

| Component-1 (Vulnerable Class) | Component-2 (Early Marriage and Nomadic Castes) | Component-3 (Unorganized Rural Labours) | Component-4 (Poverty) |
|---|---|---|---|
| Grazing cattle | Child marriage | Agriculture work | Poor economic condition |
| Long illness | Regular absence | Lack of educational facility | - |
| Without adult protection | Homeless | Migration | - |

The children who grazing cattle are also having long illness and without adult protection, hence the variable grazing cattle can be extracted from the first component. The children who are married in early age generally belong to nomadic castes in Rajasthan and due child marriage they do not attend the school regularly and drops out from their school. Hence, variable child marriage can be extracted from the second component. The children engaged in agriculture related works generally migrates from their native place to remote areas and could not avail the educational facility. Hence, the variable agriculture work can be extracted from the third component. The variable poor economic condition can be extracted from the fourth component because this is the only variable of fourth component. Further, these four variables also have maximum correlations with their respective components and independent to each other.

A model based on these four variables developed through multiple regression technique as below.

$\widehat{Dropout} = \alpha + \beta_1 \text{ (grazing cattle)} + \beta_2 \text{ (child marriage)}$
$+ \beta_3 \text{ (agriculture work)} + \beta_4 \text{ (poor economic condition)} + \varepsilon$

| Coefficients: | $\alpha$ =0.012 | $\beta_1$ =2.171 | $\beta_2$ =20.323 | $\beta_3$ =2.813 | $\beta_4$ =1.995 |
|---|---|---|---|---|---|
| p-value: | (0.016) | (0.013) | (0.002) | (0.000) | (0.014) |

R-Square: 0.786 and  Adjusted R-Square: 0.755

In relation to the dropout rate, the educational variables have also influence to dropout. So Adjusted R-Square: 0.755 is considered to explain the dropout.

**Estimation of Never Enrollment**: Preliminary Analysis: By using SPSS, determinant value of correlation matrix of all the variables under study was found 0.00000000109, which is less than the threshold value of 0.00001. For improving the determinant value, three variables - care of sibling, non-friendly environment of school and regular absence were eliminated from the study due to their high correlation with variables agriculture work, long illness and homeless respectively. After removing these variables, the determinant value of was found 0.0000195, which is greater than the threshold value of 0.00001. The value of KMO statistic was found 0.522, which is greater than 0.5. Bartlett's test was also found highly significant with *p value* < .001. By the preliminary analysis, it could be concluded that principal component analysis may be applied on these data.

Component Extraction: Three components extracted having eigenvalues greater than 1. Table-5 shows that all the three components explained 82.03 percent of the total variability. Before rotation, first component accounted for 37.87 percent of total variance than the remaining three components (25.21 percent, and 18.94 percent respectively). After rotation method it accounted for only 36.23 percent of total variance compared to remaining two components (23.54 percent and 22.26 percent respectively).

It was assumed that each variable under study will explain by the underlying components with 100 % variability but table-6 shows that 56.7 percent to 91.9 percent variability explained by the variable in retained three components because some of the components are discarded after extraction.

Table-7 shows the orthogonal rotated component matrix of variables responsible for never enrollment which reflects the correlation coefficients between the variables under study and three components.

**Table-5**
**Total Variance Explained for Never Enrollment**

| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 3.787 | 37.873 | 37.873 | 3.787 | 37.873 | 37.873 | 3.623 | 36.226 | 36.226 |
| 2 | 2.521 | 25.214 | 63.087 | 2.521 | 25.214 | 63.087 | 2.354 | 23.543 | 59.770 |
| 3 | 1.894 | 18.939 | 82.026 | 1.894 | 18.939 | 82.026 | 2.226 | 22.256 | 82.026 |

**Table-6**
**Communalities of Variables Responsible for Never Enrollment**

| Variables under Study | Initial | Extraction |
|---|---|---|
| Agriculture work | 1.000 | .871 |
| Grazing cattle | 1.000 | .919 |
| Poor economic condition | 1.000 | .667 |
| Lack of educational facility | 1.000 | .888 |
| Ignorance of guardian | 1.000 | .806 |
| Long illness | 1.000 | .871 |
| Migration | 1.000 | .833 |
| Child marriage | 1.000 | .909 |
| Homeless | 1.000 | .872 |
| Without adult protection | 1.000 | .567 |

**Table-7**
**Rotated Component Matrix of Variables Responsible for Never Enrollment**

| Variables under Study | Component | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Agriculture work | .902 | -.153 | -.186 |
| Grazing cattle | .158 | -.007 | .946 |
| Poor economic condition | .814 | .055 | .014 |
| Lack of educational facility | .795 | -.139 | .487 |
| Ignorance of guardian | .833 | .132 | .307 |
| Long illness | -.107 | .137 | .917 |
| Migration | .871 | -.210 | -.175 |
| Child marriage | .012 | .937 | -.176 |
| Homeless | -.013 | .931 | .075 |
| Without adult protection | -.158 | .696 | .240 |

**Table-8**
**Component-wise Extraction of Variables Responsible for Never Enrollment**

| Component-1 (Unorganized Rural Labours) | Component-2 (Early Marriage and Nomadic Castes) | Component-3 (Vulnerable Class) |
|---|---|---|
| Agriculture work | Child marriage | Grazing cattle |
| Poor economic condition | Homeless | Long illness |
| Lack of educational facility | - | - |
| Ignorance of guardian | - | - |
| Migration | - | - |

By using above matrix, the following variables were extracted from each component. Only those variables were considered for extraction which having the value of correlation coefficients more than 0.7.

The children engaged in agriculture related works generally migrates from their native place to remote areas and could not avail the educational facility. Due to their parent's poor economic condition and ignorance of their parents about education they have to do agriculture related labour work. Hence, the variable agriculture work can be extracted from the first component. The children who are married in early age generally belong to nomadic castes in Rajasthan. Hence, variable child marriage can be extracted from the second

component. The children who grazing cattle are also having long illness, hence the variable grazing cattle can be extracted from the third component. Further, these three variables also have maximum correlations with their respective components and independent to each other.

A model based on these three variables developed through multiple regression technique as below.

$$\widehat{\text{Never Enrollment}} = \alpha + \beta_1 \text{(agriculture work)} + \beta_2 \text{(child marriage)} + \beta_3 \text{(grazing cattle)} + \varepsilon$$

| Coefficients: | $\alpha$ =0.014 | $\beta_1$ =7.723 | $\beta_2$ =6.655 | $\beta_3$ =8.751 |
|---|---|---|---|---|
| p-value: | (0.002) | (0.000) | (0.067) | (0.000) |

R-Square: 0.881  and Adjusted R-Square: 0.869

In relation to the enrollment rate, the educational variables have also influence to never enrollment. So Adjusted R-Square: 0.869 is considered to explain the never enrollment.

**Findings:** If engagement of children in grazing cattle and agriculture related work, child marriage and poor economic condition tends to zero then dropout and never enrollment rates will decrease to 1.2 percent and 1.4 percent from average dropout rate 4.9 percent and never enrollment rate 5.8 respectively.

Hence, vulnerable class of children, early marriage and nomadic castes, unorganized rural labour class and poverty are the main factors, responsible for school dropout and never enrollment in Rajasthan.  Hence, for controlling the school dropout in Rajasthan, theses four socio-economic factors should be focused.

## Conclusion

It can be concluded by the results that engagement of children in agriculture work, engagement of children in grazing cattle, poor economic condition of parents and marriage of children in early age have influence over enrollment and dropout rates in Rajasthan. It indicates that Child Labour Act., Right to Education Act. and Child Marriage Act. are not being effectively implemented in Rajasthan state. If these three Acts are implemented effectively than school dropout and never enrollment due to engagement of children in agriculture work, engagement of children in grazing cattle, poor economic condition of parents and marriage of children in early age will

certainly decreased to a great extent in Rajasthan. By controlling only these four variables, cent percent enrollment rate and zero dropout rate can not be achieved because there are some educational variables which are also responsible for enrollment and dropout in elementary education in Rajasthan.

## References

1. Mukherjee D., Reducing out-of-school children in India, National University of Educational Planning and Administration, Delhi **(2011)**

2. Basumatary R., School dropout across Indian state and UTs: An econometric study, *Int. Res. J. Social Sci.*, Vol. **1(4)**, 28-35 **(2012)**

3. UNESCO www.unesco.org **(2013)**

4. Bokova and Bush, Literacy is key to unlocking the cycle of poverty, at http://www.chron.com/opinion/outlook/article/Literacy-is-key-to-unlocking-the-cycle-of-poverty-3848564. php **(2012)**

5. Sajjad H., Iqbal M., Siddiqui M.A. and Siddiqui L., Socio-Economic determinants of primary school dropout: evidence from south east Delhi, India, *European Journal of Social Sciences*, **30(3)**, 391-399 **(2012)**

6. Jayachandran U., How high are dropout rates in India?, *Economic and Political Weekly*, **XLII(11)**, discussion **(2007)**

7. Biswal K., An introduction to child tracking system and true cohort analysis in basic education, NUEPA, New Delhi **(2008)**

8. Sikdar S. and Mukherjee A.N., Enrollment and Dropout rate in School Education, *Economic and Political Weekly*, **XLVII (1)**, 27-31 **(2012)**

9. Kumar R. and Das S., Primary education in Jharkhand, *Economic and Political Weekly*, **XXXIX (11),** 1172-1178 **(2004)**

10. Gupta M.K. and Prasad J., Statistical Analysis of Educational Variables Responsible for Wastage in Elementary Education in Rajasthan, sent for publication to *Indian Journal of Public Administration for Publication*

11. Rajasthan Portal-SSA www.rajssa.nic.in **(2013)**

12. Field A., Discovering Statistics Using SPSS, 2nd edu, SAGE, London **(2005)**