Effect of Socio-Economic Factors on Literacy in India

A Cross-Sectional Study

ECON F342: Applied Econometrics

Under the supervision of
Dr. Rishi Kumar

NAME	ID
Sriram Karthik	2021B3A73033H
Om Bhuva	2021B3A81920H
Ashish C	2021B3A73029H
Niveditha G	2021B3A73030H
Keshav Saraiya	2021B3A72631H
Satya Shreyas Kusumanchi	2021B3A70723H
Harshith Borundia	2021B3A70727H
Karthikeya Kuchibhatla	2021B3A31024H
Sudarsan Mohanchander	2021B3A70558H
Mugdha Gupta	2021B3A72724H

Abstract

The report proposes a model analyzing the role of social capital, economic factors, and demographic variables in predicting the literacy of an individual. In the Indian context, we explore a few unique factors such as caste categories, marriage expenditure of household, ownership of vehicle etc., apart from the more straightforward variables such as income, household expenditure and individual data such as sex, income category, residence type among others. We see a very loose correlation to social factors such untouchability practices and communal conflict on literacy rates. But factors such as sex, caste categories and class of the household gives us an idea of how literacy rate is distributed among the said groups.

Contents

SECTION I: INTRODUCTION

Literacy is an important aspect for the development of an individual and the economy as a whole. Literacy rates have often been used as one of the dimensions to quantize overall human development, also being one of the 3 dimensions for calculating the UN Human Development Index. Making provisions for increasing literacy rates in a country is important for poverty and lifestyle upliftment. It is a core question in development economics. Importance of education is evident in literacy rates of developed countries versus developing and underdeveloped countries. According to 2021 World Bank data, countries like Greenland, Luxembourg, Norway report a 100% literacy rate. While moving towards the Sub Saharan region, we observe countries like Niger, South Sudan, Somalia, and Guinea report a very poor <40% literacy rate. Education is a necessity even for agricultural development and human capital. One can also observe how crime rates, malnourishment, and child morbidity rates are different between the two aforementioned groups of countries. India reported a 72.2% literacy rate in 2021.

Many obstacles arise in the improvement of literacy. We observe gender, regional and cultural disparities. Poverty, untouchability, child nutrition, unwillingness to educate girl children, child labor prevalence etc. are big hurdles in literacy. Communal conflict between classes, untouchability practices, physical disabilities from malnourishment etc. are discouraging and contribute to pupils dropping out of schooling programs, or skipping education. Coverage of government policies, trust of the rural population in government schools also affect these numbers. IHDS 2 (2011) sample survey suggests 41.1% of the households have some to hardly any confidence in government schools. On the other hand, private schooling (71.6% confidence) expenses are not viable for the poor and lower middle class population. Schemes like midday meals have been introduced to encourage children to attend schools and have proven to be successful and also deal with problems of child malnutrition.

SECTION II: LITERATURE REVIEW

II.i Paper 1

"Gender Inequality in Education and Kinship Norms in India" by Anu Rammohan & Patrick Vu

About the Paper:
Anu Rammohan and Patrick Vu (2017) investigate a critical social issue in India: **gender inequality in education**. The results provide quantitative evidence of different economic and sociocultural factors contributing to the persistent gap in educational attainment between boys and girls in India.

Use Case:
Understanding the interplay between socioeconomic factors and cultural norms influencing gender disparity in education can inform policy interventions to promote equal access to quality education for girls and boys.

Methodology:
Rammohan and Vu (2017) might utilize **established frameworks** for analyzing education inequality. They likely employ **statistical methods** on the district-level data to examine the **correlations** between factors like:

- **Socio Economic development:** This study might explore how wealthier districts fare better in bridging the gender gap in education.
- **Cultural factors:** This research might delve into the impact of **kinship norms** like **patrilocal exogamy** (where wives move to their husbands' families) on girls' education, potentially revealing negative consequences for their schooling compared to boys.
- **Geographic variations:** This study might highlight **regional disparities** in gender education, potentially observing wider gaps in northern India, as previous research suggests.

Conclusion:
Their findings include:

- Addressing cultural norms that hinder girls' education.
- Implementing policies promoting girls' education in economically disadvantaged regions.
- Tailoring strategies to address the specific challenges faced in different regions of India.

By providing **evidence-based insights**, this research can contribute significantly to the ongoing efforts to achieve gender equality in education in India.

II.ii Paper 2

"Disparities in earnings and education in India" by Geetha Rani

About the Paper:
P. Geetha Rani explores the link between education and income in India, analyzing how gender, religion, caste, and location impact earnings. It utilizes data from the Indian Human Development Survey to examine how these factors influence returns on different education levels. It highlights existing disparities and the need for policy interventions to promote equal opportunities.

Use Case:
It can be used to understand how factors like education level, location, and caste influence income in India. This knowledge can inform policies to reduce and improve access to quality education for all and to reduce income inequality.

Methodology:
Broadly, there are two methods of estimating the rates of return to education-

- **Full accounting method**
- **Mincer wage equations**

Among the two, discounting actual net age-earnings profiles is the most appropriate method of predicting the returns to education as it is considered the most critical part of the individual's early earning history. As a complete method, the internal rate of return approach requires far more data. Much of this information is not readily available to

researchers. Hence, it is less commonly applied. When used, it is applied to estimate the social rate of return by incorporating the social costs of education into the total cost calculation.

Conclusion:
The disparities in earnings and education in India persist due to various socioeconomic factors such as caste, gender, and regional inequalities. Addressing these disparities requires comprehensive policies focusing on improving access to quality education, promoting skill development, and creating equitable opportunities in the workforce. Additionally, efforts to combat discrimination and promote social inclusion are essential for fostering a more equitable society.

II.iii Paper 3

"Private versus social returns to human capital: Education and economic growth in India"

by M Schündeln & J Playforth

About the Paper:

- Investigates the micro-macro paradox in education and economic growth in India.
- Hypothesizes a disparity between private and social returns to education among government sector employees.
- Analyzes 40 years of state-level data from India.

Use Case:

- Focuses on India's significant educational growth juxtaposed with slow economic growth before reforms.
- Examines the role of the government sector in employing educated individuals.
- Seeks to understand how the large public sector in India influences the impact of education on economic development.

Methodology:

1. Data Collection:

- Utilizes state-level data from India spanning 40 years, primarily from regular censuses.

- Includes education levels, government sector employment, and economic growth indicators.

2. Analytical Approach:

- Employs cross-state regressions to control for regional differences.
- It uses a modified Cobb-Douglas function to account for unproductive human capital in the government sector.
- Incorporates interaction terms to assess the moderating effect of government size.

3. Estimation Strategy:

- Utilizes dynamic panel estimators to address endogeneity concerns.
- Conducts robustness checks with alternative measures of capital and GDP sector exclusions.

4. Results Interpretation:

- Analyzes coefficients and significance levels to understand education and government size impacts on growth.
- Discusses findings' implications for policy and understanding economic dynamics.

5. Limitations and Future Research:

- Considers data limitations and India-specific factors.
- Suggests future research directions.

The methodology involves empirical analysis to understand how education and public sector dynamics influence economic growth in India, providing insights for policymakers and researchers.

Conclusion:

- Finds a negative link between government size and the effectiveness of education in promoting economic growth.
- Identifies a positive link between education and growth in states with smaller public sectors.
- Highlights the absorption of skilled individuals by the public sector, with limited contributions to economic growth.
- Discusses implications for individual and national benefits of education.
- Acknowledges limitations related to data availability and India-specific factors.

II.iv Paper 4

“Education and caste in India”

by Chandra Pal Singh Chauhan

About the Paper:

1. Introduction to Caste Dynamics in India:

- Indian society is characterized by caste hierarchy, religious affiliation, linguistic diversity, and regional loyalty.
- Focus on caste-based stratification due to its historical significance and contemporary political dimensions.
- Major social categories: Forward Castes (FCs), Scheduled Castes and Scheduled Tribes (SC/ST), and Other Backward Castes (OBCs).

2. Educational Empowerment and Affirmative Action:

- Education is a vital tool for upward social mobility and reducing socioeconomic disparities.
- Government initiatives include reservations in educational institutions and jobs, scholarships, and free education.
- Despite affirmative action efforts, lower castes still face challenges in educational attainment and socioeconomic advancement.

3. Paper's Objectives and Research Questions:

- Analyzes the impact of affirmative action policies on educational and socioeconomic promotion.
- Key questions include the validity of caste as a measure of educational backwardness and the effectiveness of reservation policies.

Use Case:

1. Historical Roots of Caste System:

- Historical origins are traced to the varna system with four major occupational classes: Brahmins, Kshatriyas, vaishyas, and Shudras.
- The gradual evolution from varnas to thousands of castes and sub-castes led to complex social stratification.

2. Constitutional Safeguards and Policy Framework:

- Indian Constitution provides reservations for SCs, STs, and OBCs to address historical injustices and social inequalities.
- Reservations aim to promote equality and social justice but face challenges in implementation and effectiveness.

3. Progress and Challenges in Education:

- There is some progress in primary education enrolment for SCs/STs but stagnation at secondary and higher levels.
- Due to data limitations and evolving social dynamics, OBC education status remains to be clarified.

Methodology:

1. Research Framework:

- Analysis of government reports, scholarly literature, and socioeconomic indicators to assess reservation policy impact.
- Critically evaluates existing data on educational representation and caste dynamics.

2. Data Collection and Analysis:

- Utilizes historical records, constitutional provisions, and empirical studies to understand caste system evolution and policy implications.
- Examines literacy rates, enrolment patterns, and dropout rates to gauge educational progress and challenges.

3. Policy Evaluation and Recommendations:

- Critically assesses the effectiveness of reservation policies in promoting educational equity and socioeconomic mobility.
- Proposes alternative criteria for identifying disadvantaged groups and suggests policy reforms based on economic considerations.

Conclusion:

1. Summary of Findings:

- Despite affirmative action measures, genuine empowerment of lower castes remains limited.
- Reservation policies face criticism for perpetuating caste consciousness and benefiting only a privileged few.
- Need for data-driven, economically focused interventions to address root causes of educational inequality.

2. Recommendations for Policy Reform:

- Shift from caste-based to economically based criteria for identifying disadvantaged groups.
- Focus on quality primary and secondary education, especially in rural and tribal areas.
- Regular review and revision of reservation policies to ensure inclusivity and effectiveness.

3. Call for Comprehensive Approach:

- Emphasizes the importance of multi-dimensional strategies for promoting social justice and equality.
- To address systemic inequalities, policymakers must prioritize investments in education, poverty alleviation, and social cohesion.

In conclusion, the paper comprehensively analyzes the complex relationship between caste, education, and affirmative action in India. Critically examining historical contexts, policy frameworks, and empirical data sheds light on the challenges and opportunities in promoting educational equity and social justice in a diverse and stratified society.

II.v Paper 5

"Inequalities in Demand and Access to Early Childhood Education in India"
by Saikat Ghosh

About the Paper:

India has one of the world's largest provisions for free, public early childhood education, yet almost half of the children lack access, and the reasons for this disparity are not fully understood. The objective of the paper is to investigate factors affecting demand and access to preschool in India, particularly focusing on parental decisions influenced by socio-economic factors and awareness.

Use Case:

- **Importance of ECE:** Acknowledged as crucial for building human capital, especially in developing countries where children are vulnerable to poverty and disadvantage.
- **Global Perspective:** ECE is seen as a tool for sustainable development, offering long-term benefits for individuals and society.
- **India's Scenario:** Despite significant investment in public preschool programs, many children still lack access, highlighting the need for understanding and addressing inequalities.

Methodology:

- **Data Source:** Survey data collected from 1373 households in West Bengal, India, in 2015.
- **Sampling:** Purposive and multi-stage sampling to ensure variation in socio-economic status, with one district representing higher socio-economic status and another representing lower status.
- **Data Collection:** Paper-based questionnaires administered by fieldworkers, focusing on parental reasons for preschool attendance decisions.
- **Data Analysis:** Utilized both quantitative (descriptive statistics, Probit modeling) and qualitative methods to explore factors influencing preschool attendance.

Conclusion:

- **Unequal Access:** Preschool opportunities in India are unevenly distributed across regions and socio-economic groups.
- **Demand-Side Factors:** Lack of awareness about the benefits of ECE, especially among parents with lower education levels, contributes to unequal access.
- **Supply-Side Factors:** While not extensively explored, factors like resource availability in disadvantaged areas likely contribute to the disparity.
- **Socio-economic Status (SES):** Parental SES, particularly education level, emerges as the primary driver of unequal access.
- **Recommendations:** Calls for raising awareness about ECE benefits, implementing stronger policies, advocating for universal access, and addressing regional and socio-economic disparities.

Overall, the paper underscores the need for concerted efforts to address the complex issue of unequal access to early childhood education in India, emphasizing the importance of both demand-side interventions and supply-side improvements

SECTION III: DATA AND METHODOLOGY

The Cross sectional data used here provided by IHDS2 DS1 containing a sample size of 204,569 individuals and DS2 containing 42,152 households. The relevant variables were cleaned for null values after merging the two datasets.

The dependent variable here assumes the value 1, for literate candidates, and 0 for illiterate. Thus, the method of evaluation is a linear probability model. The regression coefficients reflect the extent to which a one-unit change in an independent variable influences the probability of the dependent variable. For the qualitative factors, each category is given a dummy variable and the respective coefficients show the change in probability for the same.

Quantitative variables like consumption expense of the household, and household income have been normalized by using logarithm as a unit change in a variable such as income will not reflect much on the dependent variable. For the variables using log, a small arbitrary number (10e-10) to keep the values in the allowed domain.

Variables used:

VARIABLE - ED2 (Dependent Variable)
Description - Education: Literacy
Reason - ED2 is the dependent binary variable for which we are running the regression. It takes value 0 if the individual is illiterate and value 1 if the individual is literate.

VARIABLE - log_INCEARN
Description - Annual household wage and salary earnings with bonuses
Reason - To study the variation in the educational attainment of an individual depending upon the income of the household. It is expected that the education level will increase with an increase in the household income.

VARIABLE - log_COTOAL

Description - Total household consumption expenditure
Reason - To study whether an increase in the household consumption expenditure will lead to an increase in the educational attainment of an individual in the household.

VARIABLE - log_CO35
Description - School/College fees: Total value Rs
Reason - To study the variation in the educational attainment of an individual depending upon the fees charged by various institutes.

VARIABLE - STATEID
Description - State code - It is a qualitative variable including the states and union territories of India ranging from 1 to 35. 1 represents Jammu & Kashmir, 2 represents Himachal Pradesh and so on.
Reason - To study literacy distribution among states. The variable has been introduced to control for the state variable.

VARIABLE - RO3
Description - Sex
Reason - To study whether the sex of an individual determines the level of educational attainment of the individual. This also gives us an idea about disparity between the two genders.

VARIABLE - ID11
Description - Religion - It is a qualitative variable ranging from 1 to 9. 1 represents Hindu, 2 represents Muslim and so on.
Reason - To study whether the religion of an individual is affecting the educational attainment of the individual.

VARIABLE - ID13
Description - Caste Category - It is a qualitative variable ranging from 1 to 6. 1 represents Brahmin, 2 represents other General Caste except Brahmin and so on.
Reason -
To study whether the caste of an individual is affecting the educational attainment of the individual.

VARIABLE - CGVEHICLE
Description - It is a binary variable that checks if the household owns any vehicle or not. It takes the value 0 if the household does not own a vehicle and 1 if it owns a vehicle.
Reason - To study whether there is a correlation between the educational attainment of an individual and the ownership of a vehicle by the household from which the individual belongs.

VARIABLE - URBAN2011
Description - Variable depicts the Urban residence from census 2011 rural is given value 0 and urban is given value 1.
Reason - To study whether there is a relationship between the educational attainment of an individual and the type of habitat/locality where the individual belongs.

VARIABLE - CG3
Description - Variable depicts whether the household is Poor/ Middle/Comfortable category. Poor is given value 1, Middle 2 and Comfortable 3.
Reason - To study if there is a relationship between the financial situation of the household,the individual is in and the educational attainment of the individual.

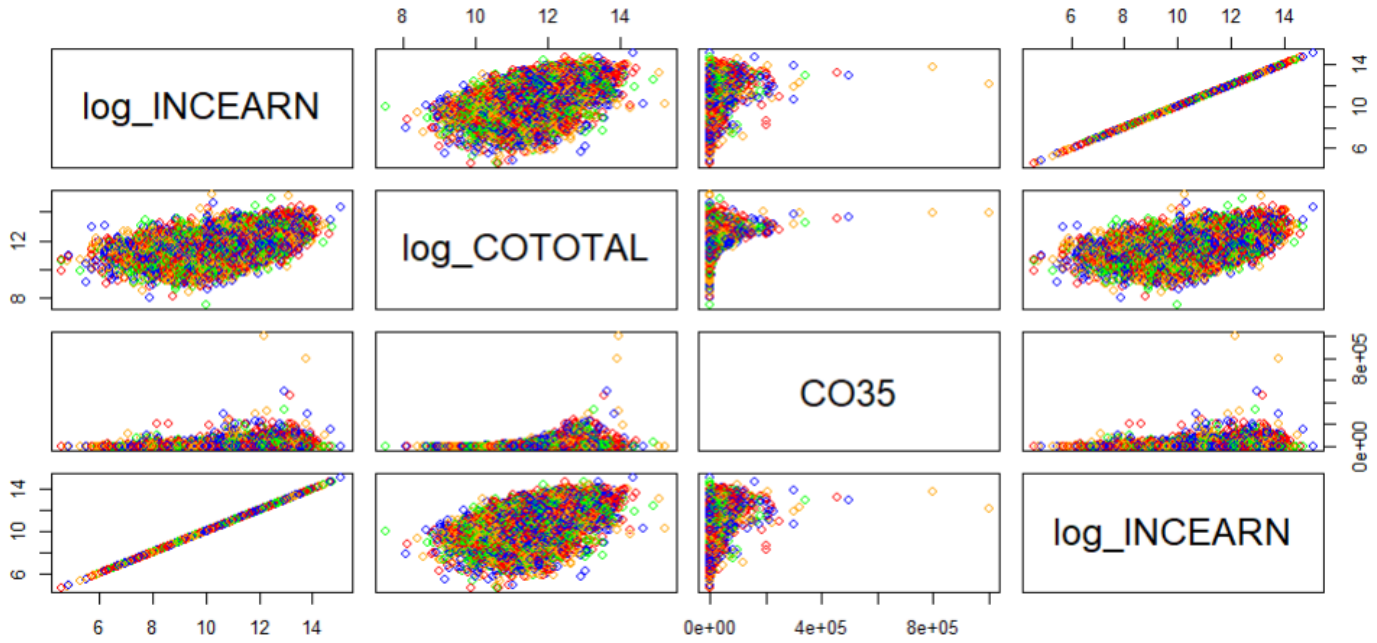
VARIABLE - DB8A
Description - Apply to bank - The variable takes 3 values. 1 for those who did not apply, 2 for those who did not get it and 3 for those who were able to get it.
Reason - To study whether the educational attainment of an individual is related to his ability to apply at a bank. It is expected that with an increased ability to apply at a bank the education level increases.

VARIABLE - TR3
Description - The variable takes 3 values. 1 if there is a lot of conflict, 2 if there is some conflict and 3 if there is not much conflict.
Reason - To study how does the educational attainment of an individual varies with the level of conflict prevalent in the place the individual is residing.

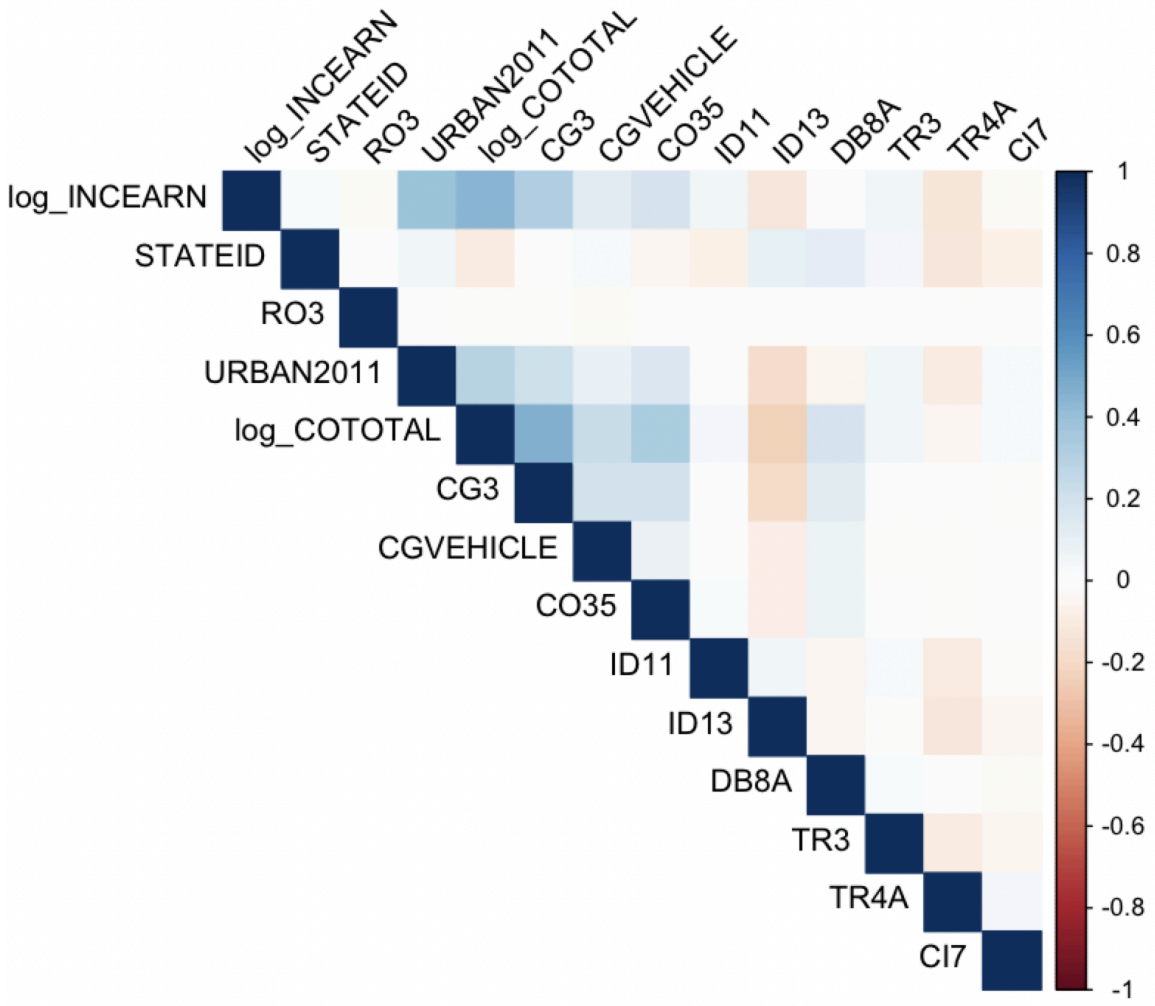
VARIABLE - TR4A
Description - It is a binary variable that depicts whether the individual practices untouchability. It takes the value 0 if the individual does not practice and takes the value 1 if the individual practices.
Reason - To study if there is a correlation between the educational attainment of the individual and the practice of untouchability.

VARIABLE - CI7
Description - Variable depicts the confidence the individual has on government schools to provide good education. It takes the value 1 for “A great deal of confidence”, takes 2 for “Only some confidence” and 3 for “Hardly any confidence at all”.
Reason - To study whether the confidence the individual has on the government to provide good education has a relationship with the individual’s educational attainment.

GRAPH MATRIX



CORRELATION MATRIX:



SECTION IV: REGRESSION

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-1.8737e+00	1.5962e-01	-11.7385	< 2.2e-16	***
log_INCEARN	5.8402e-03	6.1500e-03	0.9496	0.3422964	
factor(STATEID)2	6.5445e-01	5.8266e-02	11.2322	< 2.2e-16	***
factor(STATEID)3	1.6830e-01	6.1777e-02	2.7244	0.0064420	**
factor(STATEID)4	3.8958e-01	2.4321e-01	1.6018	0.1091979	
factor(STATEID)5	2.1913e-01	7.2255e-02	3.0328	0.0024232	**
factor(STATEID)6	9.4586e-02	5.4561e-02	1.7336	0.0829942	.
factor(STATEID)7	1.8194e-01	6.4914e-02	2.8029	0.0050650	**
factor(STATEID)8	-4.1255e-01	5.0671e-02	-8.1417	3.897e-16	***
factor(STATEID)9	-9.5505e-02	5.0200e-02	-1.9025	0.0571096	.
factor(STATEID)10	-1.9667e-01	5.6372e-02	-3.4887	0.0004853	***
factor(STATEID)11	5.7934e-01	1.4885e-01	3.8920	9.942e-05	***
factor(STATEID)12	5.4864e-01	1.4171e-01	3.8717	0.0001081	***
factor(STATEID)13	8.2847e-01	1.9058e-01	4.3470	1.380e-05	***
factor(STATEID)14	1.7994e-01	1.4276e-01	1.2604	0.2075241	
factor(STATEID)15	1.6657e+00	2.2859e-01	7.2866	3.179e-13	***
factor(STATEID)16	9.6392e-01	1.0247e-01	9.4065	< 2.2e-16	***
factor(STATEID)17	9.9105e-01	1.3608e-01	7.2829	3.268e-13	***
factor(STATEID)18	6.8275e-01	6.1439e-02	11.1127	< 2.2e-16	***
factor(STATEID)19	4.2062e-01	5.3387e-02	7.8786	3.310e-15	***
factor(STATEID)20	1.9759e-01	6.1109e-02	3.2334	0.0012234	**
factor(STATEID)21	2.5952e-01	5.4996e-02	4.7188	2.372e-06	***
factor(STATEID)22	2.4675e-01	5.7379e-02	4.3004	1.705e-05	***
factor(STATEID)23	1.8561e-01	5.2049e-02	3.5661	0.0003623	***
factor(STATEID)24	4.7575e-02	5.4584e-02	0.8716	0.3834289	
factor(STATEID)25	5.4239e-01	1.8017e-01	3.0105	0.0026081	**
factor(STATEID)26	3.3020e-01	1.4997e-01	2.2018	0.0276789	*
factor(STATEID)27	4.1652e-01	5.2589e-02	7.9203	2.369e-15	***
factor(STATEID)28	-1.9052e-01	5.3583e-02	-3.5556	0.0003771	***
factor(STATEID)29	1.2543e-01	5.0538e-02	2.4818	0.0130718	*
factor(STATEID)30	1.1884e+00	1.5002e-01	7.9212	2.353e-15	***
factor(STATEID)32	1.2312e+00	6.5979e-02	18.6603	< 2.2e-16	***
factor(STATEID)33	2.1813e-01	5.7296e-02	3.8070	0.0001406	***
factor(STATEID)34	6.4724e-01	1.7377e-01	3.7247	0.0001956	***
factor(RO3)2	-7.1887e-01	1.1923e-02	-60.2934	< 2.2e-16	***
factor(URBAN2011)1	5.3701e-01	1.5400e-02	34.8706	< 2.2e-16	***
log_COTOTAL	2.5317e-01	1.2820e-02	19.7487	< 2.2e-16	***

```
factor(CG3)2          3.1200e-01  1.3706e-02  22.7642 < 2.2e-16 ***
factor(CG3)3          4.8123e-01  3.4600e-02  13.9081 < 2.2e-16 ***
factor(CGVEHICLE)1    2.0101e-01  1.3896e-02  14.4649 < 2.2e-16 ***
CO35                  2.4442e-05  1.5634e-06  15.6342 < 2.2e-16 ***
factor(ID11)2        -4.3895e-01  1.9920e-02 -22.0352 < 2.2e-16 ***
factor(ID11)3         3.7784e-01  5.3814e-02   7.0212 2.200e-12 ***
factor(ID11)4         2.9997e-02  5.7229e-02   0.5242 0.6001664
factor(ID11)5         3.3242e-01  7.8629e-02   4.2277 2.361e-05 ***
factor(ID11)6         6.5689e-02  1.7227e-01   0.3813 0.7029673
factor(ID11)7         4.5855e-02  8.2913e-02   0.5530 0.5802309
factor(ID11)8         5.7367e-01  2.0156e-01   2.8462 0.0044251 **
factor(ID11)9        -2.3897e-01  5.4691e-01  -0.4369 0.6621561
factor(ID13)2        -4.0938e-01  3.9204e-02 -10.4425 < 2.2e-16 ***
factor(ID13)3        -5.6689e-01  3.7599e-02 -15.0774 < 2.2e-16 ***
factor(ID13)4        -7.6288e-01  3.8229e-02 -19.9555 < 2.2e-16 ***
factor(ID13)5        -9.1045e-01  4.1657e-02 -21.8561 < 2.2e-16 ***
factor(ID13)6        -6.5708e-01  6.8967e-02  -9.5274 < 2.2e-16 ***
factor(DB8A)2        -1.5831e-02  3.4767e-02  -0.4553 0.6488721
factor(DB8A)3         9.7808e-02  1.5988e-02   6.1174 9.510e-10 ***
factor(TR3)2         -2.3974e-02  2.2838e-02  -1.0497 0.2938334
factor(TR3)3         -9.5661e-03  2.1885e-02  -0.4371 0.6620313
factor(TR4A)1        -1.0667e-01  1.5905e-02  -6.7070 1.987e-11 ***
factor(CI7)2         -6.8411e-02  1.3252e-02  -5.1623 2.440e-07 ***
factor(CI7)3         -2.1101e-02  2.2224e-02  -0.9495 0.3423900
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

As the maximum of our independent variables are categorical or qualitative, the model takes each category as a dummy variable and provides a dummy variable for each category.

F statistic = 181.5781
P value very close to 0, model significant as a whole
R-squared:

INTERPRETATION:

- Looking at the coefficient for RO3, we see a significant gap in literacy between males and females. This points towards the gender gap in education. Discouraging female literacy is evident here.
- We also observe that for Non-General caste categories (ID13), and for Poor category (CG3) we see how caste and class have an effect on literacy rates. This points towards class differences and social capital, and how this affects distribution of literacy between classes. Religion (ID11) also shows us the distribution among religions.
- Opposed to our hypothesis that untouchability and conflict between classes in villages will show an impact on the dependent variable, the significance is low, and a clear interpretation cannot be made.
- There is a positive correlation of around 0.3 between the literacy rate and the income of the household and literacy rate and whether the individual stays in an urban area or not.
- This suggests that literacy rates tend to be higher among individuals residing in urban areas compared to rural areas. It indicates a potential association between urbanization and literacy levels.
- There is a positive correlation of around 0.5 between literacy and whether the individual is from a household having a good financial situation or not. It suggests that financial stability may facilitate access to education or enhance the likelihood of literacy development.
- The positive correlation between literacy and household expenditure suggests that households with higher literacy rates tend to have higher levels of expenditure. It could imply that education or literacy is associated with greater economic activity or consumption patterns.

SECTION V: DIAGNOSTICS

Test for heteroskedasticity:
Breusch-Pagan test

Heteroskedasticity refers to the situation where the variance of the errors (residuals) in a regression model is not constant across observations.

```
> bptest(mod2)

        studentized Breusch-Pagan test

data:  mod2
BP = 9518.1, df = 70, p-value < 2.2e-16
```

The p-value is rejected which implies that **heteroskedasticity exists** in the model.
To counter this issue, we have applied **Robust Linear Regression**.

Multicollinearity:

The presence of multicollinearity may be ascertained by computing the Variation Inflation Factor for each variable. Typically, a VIF exceeding 10 indicates the presence of multicollinearity amongst the variables. On conducting the test,

	GVIF	Df	GVIF^(1/(2*Df))
log_INCEARN	1.580152	1	1.257041
log_MP4	1.216237	1	1.102831
factor(STATEID)	9.059811	32	1.035035
factor(RO3)	1.031089	1	1.015426
factor(URBAN2011)	1.374306	1	1.172308
log_COTOTAL	1.897231	1	1.377400
factor(CG3)	1.486537	2	1.104190
factor(ED6)	1.425072	16	1.011131
factor(CGVEHICLE)	1.265955	1	1.125146
log_CO35	1.201232	1	1.096007
factor(ID11)	3.778547	8	1.086633
factor(ID13)	2.064974	5	1.075205

In light of the VIF values of less than 10 for all variables, it is reasonable to assert with confidence the absence of multicollinearity amongst the said variables.

SECTION VI: CONCLUSION

In light of the aforementioned empirical inquiry, our endeavor was to explore the relation between distinct variables amongst Indian individuals. Drawing on the tenets of regression analysis performed on cross-sectional data, we demonstrated how socio economic factors have affected literacy rate on various social groups under various circumstances and drew several inferences. These results are aligned with extant literature on the subject that also postulated analogous relationships between the said variables and the eventual diagnosis. Similar effects were also seen in the 5 papers read. That being said, we must underscore several limitations that warrant attention. We utilized a rudimentary cross-sectional dataset that precludes us from accounting for several non-quantitative variables which might have affected literacy.

SECTION VII: REFERENCES:

Zua, Biale,Literacy: Gateway to a World of Exploits, International Journal of Education and Literacy Studies, v9 n1 p96-104 Jan 2021

Aggarwal, Y. 1997.*Small schools: issues in policy and planning*. New Delhi, National Institute of Educational Planning and administration. (NIEPA occasional paper, no. 23).

Agrawal, T. (2011). *Returns to education in India: Some recent evidence*. Mumbai: Indira Gandhi Institute of Development Research. Retrieved from http://www.igidr.ac.in/pdf/publication/WP-2011-017.pdf

Geetha Rani, P. (2014). Equity in the distribution of India's government subsidies on education. *International Journal of Education and Economic Development*, *5*, 1–39.

Dighe A. (1991). Women and literacy: Some policy considerations. Indian Journal of Adult Education (New Delhi), 52(1&2), January–March, April–June, 58.

Greene, W. (2011). *Econometric analysis* (7th ed.). New Delhi: Pearson Education in South Asia.
Sinha S. (1981). Level of educational development among the SCs of Bihar. *M.Phil. Dissertation*, Jawaharlal Nehru University, New Delhi, India.

Thangaraj M. (1995). Analysis of literacy by social groups in India. *Indian Journal of Adult Education* (New Delhi), 56(4), 12.

Chauhan , C.P.S. 2004 . Modern Indian education: Policies, progress and problems , New Delhi, India : Kanishka Publishers .