

EDUCATIONAL INEQUALITIES ACROSS SOCIAL GROUPS IN INDIA

by

Tenzin Yindok
BA Economics, Simon Fraser University

PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ECONOMICS

In the
Department of Economics

© Tenzin Yindok
SIMON FRASER UNIVERSITY
Fall 2009

All rights reserved. However, in accordance with the *Copyright Act of Canada*, this work may be reproduced, without authorization, under the conditions for *Fair Dealing*. Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.

APPROVAL

Name: Tenzin Yindok
Degree: Master of Arts
Title of Project: Educational Inequalities Across Social Groups in India

Examining Committee:

Chair: **Christoph Luelfesmann**
Professor, Department of Economics

Alexander Karaivanov
Senior Supervisor
Associate Professor, Department of Economics

Brian Krauth
Supervisor
Associate Professor, Department of Economics

Stephen Easton
Internal Examiner
Professor, Department of Economics

Date Defended/Approved: December 11, 2009



SIMON FRASER UNIVERSITY
LIBRARY

Declaration of Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the "Institutional Repository" link of the SFU Library website <www.lib.sfu.ca> at: <<http://ir.lib.sfu.ca/handle/1892/112>>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada

ABSTRACT

School enrolment and educational attainment of primary and secondary school-going children in India are estimated by social group, using the India Human Development Survey, 2005. The main findings of this paper are that Muslims as a social group has by far the lowest probability of enrolment and attainment of education as measured by current enrolment and completed years of education, followed by Scheduled Tribes and Scheduled Castes. The 'social group effect' is smaller in magnitude for enrolment of children of primary schooling age than for secondary school-aged children; bigger in magnitude for educational attainment of female as compared to male children.

Keywords: schooling India; enrolment; caste system; social groups India; developing countries

ACKNOWLEDGEMENTS

I would like to express my deep and sincere gratitude to my supervisor Professor Alex Karaivanov, Simon Fraser University, for his overall guidance and advice during the preparation of this paper.

My thanks are due for Professor Brian Krauth and Professor Stephen Easton of Simon Fraser University for their insightful comments and detailed review of my paper.

Finally, I would like to acknowledge my classmate Raged Anwar for, among other things, his invaluable help with formatting in Microsoft Word.

TABLE OF CONTENTS

Approval	ii
Abstract	iii
Acknowledgements	iv
Table of Contents	v
Introduction	1
Literature Review	4
Model	8
Data	13
Regression Results	15
1.1. Enrolment	16
1.1.1 Enrolment of children aged 6-14	16
1.1.2 Enrolment of children aged 15-18	18
1.1.3 Enrolment of children aged 15-18 with at least 7 years of schooling	20
1.2. Educational Attainment	21
1.3. Robustness	22
Discussion and Conclusions	24
Appendices	29
Appendix A	29
Table A. Enrolment, Children aged 6-14	30
Table B. Enrolment, Children aged 15-18	31

Table C. Enrolment conditional on 7 years of schooling, age 15-18	32
Table D. Partial Effects	33
Table E. Enrolment Gaps across Social Groups	34
Table F. Educational Attainment, 17-18 Cohort by Gender	36
Appendix B	38
Descriptive Statistics of Variables	38
Selected Statistics by Social Group	39
Bibliography	40

INTRODUCTION

Research on schooling in India has largely concluded that although there has been a rapid rise in enrolment rates over the years¹, the quality of schooling is extremely poor in terms of administration, school infrastructure and learning levels². In light of the fact that the Indian society is stratified across religious and caste lines, certain social groups have long been identified as being more vulnerable to these deficiencies.

The purpose of this paper is to analyze the channels through which the socio-religious association of a household exhibits itself in terms of the schooling of its children. In particular, my research questions of interest are:

- 1) How does school enrolment and attainment differ by caste and religion in India?
- 2) Does the relationship persist after controlling for family socioeconomic status and location effects?

The Indian population is 80 percent Hindu and 13 percent Muslim; the remaining is constituted by Christians, Sikhs, Jains and Buddhists. Additionally, among the Hindus, the four primary castes are Brahmins, Kshatriya, Vaishya and Shudras. These hierarchical divisions were formed historically on the basis of

¹The World Bank

²See Glewwe & Kremer (2006)

occupation; Brahmins were priests and teachers, Kshatriyas were warriors and rulers, Vaishyas traders and Shudras servants. These groups are further divided into sub-castes, which are still relevant today in terms of social association and marriage patterns.

'Dalits' were left outside of the caste system for engaging in occupations considered too lowly (janitors, cremators, removers of animal skin from carcasses and so on). They were called the 'untouchables' due to a practice that prohibited Hindus from being in direct physical contact with them, leading to strict social rituals of exclusion. Although untouchability is illegal in India, social discrimination against the dalits still exist in many avenues of social and professional life³. Adivasi is a term used to refer to the tribal communities in India that inhabit in geographically isolated regions. Dalits and Adivasis are respectively called Scheduled Castes (SC) and Scheduled Tribes (ST) in the Indian Constitution. The prevalence of the caste system and the identification of Scheduled Caste and Scheduled Tribe as lagging social groups have led to constitutional provisions to mitigate the negative implications of their status. Affirmative action for SC and ST are highly prominent in both legislature and educational institutions in the form of reservations.

On the other hand, Muslims have historically been the significant 'other' to the Hindu majority. A vast number of Muslim middle-class families migrated out of the country during the India-Pakistan partition in 1947, leaving behind the

³Dalits could be precluded from such things as entry into public temples, use of public wells and common burial grounds.

poorer Muslims (Desai et al, Forthcoming). Social discrimination and lower economic status has led to lower access to formal sector jobs and education.

Individuals belonging to groups that were historically discriminated on religious and caste grounds have had little access to education and preferred occupational categories. This has led to even less acquisition of education, creating a loop of lower income, education and life capabilities. The main findings of this paper are that Muslims as a social group has by far the lowest probability of enrolment and attainment of education as measured by current enrolment and completed years of education, followed by Scheduled Tribes and Scheduled Castes. The 'social group effect' is smaller in magnitude for enrolment of children of primary schooling age than for secondary school-aged children; bigger in magnitude for educational attainment of female as compared to male children.

LITERATURE REVIEW

Research on educational inequalities across social groups has been motivated by the emphasis placed on education for the economic and social progress of individuals and households. In the US, such research has focused largely on the white-black gap in schooling outcomes. For example, Barro (1987) emphasizes the importance of family socioeconomic status (SES) indicators in the incidence of dropping out from high school, and finds that differences in educational attainment between whites and blacks wash out when SES factors are accounted for. Mare (1980) contends that parental SES effects are less severe for transition between grades for children in later school years. The emerging picture is that different educational outcomes for different social groups are largely due to 'endowment' factors such as economic status and parental education, and that once children manage to attain a certain level of schooling, SES factors matter much less.

Almost all studies done on schooling choices in the Indian context include dummies for SC and ST (Muslim dummies are common too) in their analyses even though their primary focus could be in something else like wealth or state effects. Almost all of the studies find significant negative coefficients on the dummies of SC, ST and Muslim. Although the magnitudes of coefficients on various social groups are quite different from the results of this paper on the account that studies were done for different periods using different data sets and

explanatory variables, they are qualitatively similar in suggesting a disadvantage for the lagging groups.

Filmer and Pritchett (1998) run a probit regression of primary school enrolment using the National Family Health Survey of 1992-93, where they include a single dummy for SC and ST. They find that being a SC/ST means 4.7 percentage points less likelihood of being in primary school, after controlling for gender, age, household and community variables (the presence of primary, middle, secondary schools and health facilities in the village). Duraisamy (2000) uses the 1994 National level survey data of the NCAER on rural India in sixteen states to conclude that children from SC and ST households are less likely to attend school and more likely to work than others, and that Muslims families are less likely to send their children to school than other religious groups. The probability of child work-participation is significantly lower for Hindu and Christian families. The author also alludes to the fact that SCs and STs do not appear to differentiate based on gender, whereas Hindus and Muslims seem less keen on enrolling girls in school. Dreze and Kingdon (2001) analyze the determinants of school participation in rural North India using the PROBE⁴ survey on 4,400 children in five Indian states. They conclude that SC, ST children are less likely to go to school than children in general castes, after controlling for household variables.⁵ The Muslim dummy, although negative, tests insignificant, which is in

⁴PROBE stands for the Public Report on Basic Education.

⁵ The paper additionally refers to the finding in Kingdon (1998) where with data from 1995, conditional on school enrolment, years of education attained by SC children were no lower than those of others

stark contrast to the findings of this paper (highly significant and largely negative).

Barooah and Iyer (2005) examine whether and to what extent the enrolment of children at school in India is influenced by religious or caste norms using an NCAER survey of 1994 on 33,000 rural households in 16 Indian states. They find that on average, 54 and 52 percent of the difference in enrolment rates between Hindu and Muslim boys and girls – and 37 and 27 percent of the difference between Hindu and Dalit⁶ boy and girls- was ‘community-determined’⁷. Their finding that the Hindu-Muslim enrolment gap is not very different between genders is consistent with the results in this paper.

Enrolment rates across social groups have evolved overtime. Desai and Kulkarni (2008) analyze the changes in educational inequalities in India in 1980s and 1990s using the National Sample Surveys from different time periods. They focus on educational transitions and uses social groups, age, gender, urban dummy, region, household size and per capita household expenditure as explanatory variables. They find that Muslims underwent substantial improvements in school entry, but Dalits and Adivasis experienced stronger progress. The likelihood of primary school completion actually declined for Muslims even though enrolment increased. They also suggest that while Dalits and Adivasis, benefiting from positive discrimination, have managed to narrow educational gaps in primary education, this has not been true for Muslims, who

⁶In their paper, dalits refer to both SC and ST.

⁷ ‘Community-determined’ is used to refer to the enrolment gap attributable to social groups after controlling for the rest of the variables.

did not receive a preferential treatment. After controlling for age, residence and household size, the most recent (1999-2000) predicted probabilities of primary enrolment are 71.8, 79.8, 75.9, 85.8 percent for Muslim, Dalit, Adivasi and upper caste Hindus males respectively (equivalent figures for girls are 61.7, 73.3, 71 and 81.8 percent). The predicted probabilities of completing secondary school conditional on completing middle school for male (female) are 58.5 (59.7), 59.8 (60.3), 53.5 (54.5), 66(70.1) percent for Dalit, Adivasi, Muslim and upper caste Hindus respectively.

MODEL

Schooling choices such as enrolment and dropping out are made by individuals and/or households by comparing the costs and benefits of education. We would expect to see 'social-group effects' if the costs and benefits of acquiring education were different for households *because* they belong to a particular caste or religion.⁸The purpose of the following empirical analysis would be to distinguish group-determined differences from secularly-determined differences in enrolment and mean years of schooling.

A general measurement issue presents itself when examining determinants of years of completed schooling from retrospective survey data. A child's completed years of schooling is known only many years after he or she first enrolls in school. If we take years of completed schooling from a sample of young children, we create censoring bias (Holmes, 1997) since we do not distinguish between currently enrolled children and those who have terminated their schooling. That is, when we take the current level of schooling of a child who has not completed his/her schooling, we are censoring his/her years of schooling. This bias is not corrected by simply including age as a dependent variable since two children of the same age are treated identically in the estimation, regardless of their current enrolment in school. Taking completed

⁸A more detailed speculation on why costs and benefits of acquiring education could be different for households of different social groups is included in the 'discussion and conclusions' section.

years of schooling of adults on the other hand runs into the problem of not being able to measure the environment each lived in when the schooling decisions were made. Therefore, to study the determinants of years of completed schooling, I restrict the area of study to the cohort of 17 and 18 year olds in the sample. Amount of completed education is top-coded at twelve years to study the effect of social group association on school-level educational attainment. It can be reasonably expected that by the age of 17 or 18, a child would have completed most of its school-level education, and that the children of that age have not left their parental homes. Formally, years of schooling (S) are modeled as a function of social-group association⁹ and set of control variables (X_k);

$$S = \beta_0 + \beta_3 \text{High_Caste} + \beta_4 \text{Other_Backward_Castes} + \beta_5 \text{SC} + \beta_6 \text{ST} + \beta_7 \text{Muslim} + \beta_8 \text{Other} + \beta_k X_k + \varepsilon$$

Let us take the choice of enrolling in school. The outcome of this discrete choice can be viewed as a reflection of a latent regression. For instance, assume that a household makes a marginal benefit/ marginal cost calculation based on utilities achieved by enrolling or not enrolling a child. We can model the difference between the benefit and cost as an unobserved variable Y^* such that;

$$Y^* = \beta_0 + \beta_1 \text{Age}^{10} + \beta_2 \text{Female} + \beta_3 \text{High_Caste} + \beta_4 \text{OBC} + \beta_5 \text{SC} + \beta_6 \text{ST} + \beta_7 \text{Muslim} + \beta_8 \text{Other} + \beta_k X_k + \varepsilon$$

⁹The social groups considered are Brahmin, High Caste (HC), Other Backward Castes (OBC), Scheduled Castes (SC), Scheduled Tribes (ST) and Muslim, with Brahmin being the base category. The paper (and the data set) follows the Census of India classification and terminology of religion and caste groups. Social group classification is discussed in the Data section.

¹⁰ Current enrolment of children aged 6 to 14) include age and age-squared.

X_k is the vector of the remaining variables. We do not observe the net benefit of enrolment, only whether the child is enrolled or not (Y). Our observation is whether the child is currently enrolled or not;

$Y =$ currently enrolled if $Y^* > 0$

$Y =$ not enrolled currently if $Y^* \leq 0$

To estimate this function, the distribution of the error term in the equation for Y^* is assumed to be normal with mean zero. The enrolment probabilities can be derived as;

$$\begin{aligned}
 P(Y=1|X) &= P(Y^*>0 |X) \\
 &= P(\varepsilon > -\beta_0 - \beta_1 \text{Age} - \beta_2 \text{Female} - \beta_3 \text{High_Caste} - \beta_4 \text{OBC} - \beta_5 \text{SC} - \beta_6 \text{ST} \\
 &\quad - \beta_7 \text{Muslim} - \beta_8 \text{Other} - \beta_k X_k |X) \\
 &= 1 - \Phi(-\beta_0 - \beta_1 \text{Age} - \beta_2 \text{Female} - \beta_3 \text{High_Caste} - \beta_4 \text{OBC} - \beta_5 \text{SC} - \beta_6 \text{ST} \\
 &\quad - \beta_7 \text{Muslim} - \beta_8 \text{Other} - \beta_k X_k |X) \\
 &= \Phi(\beta_0 + \beta_1 \text{Age} + \beta_2 \text{Female} + \beta_3 \text{High_Caste} + \beta_4 \text{OBC} + \beta_5 \text{SC} + \beta_6 \text{ST} \\
 &\quad + \beta_7 \text{Muslim} + \beta_8 \text{Other} + \beta_k X_k)
 \end{aligned}$$

where Φ is the cumulative standard normal distribution function.

The control variables, X_k , include a measure of household economic status (household monthly consumption per capita), years of schooling of the highest educated adult in the household, a dummy variable for living in a rural area and state dummies for each of the thirty-two states in India.

If families are credit-constrained, current economic resources of the household may influence the family's capacity to invest in the child's schooling. The economic resources of a household are reflected in both its total income and consumption. Although it is not immediately clear which of the two measures should be used, lifetime wealth could be expected to be reflected better in consumption spending due to consumption smoothing. Per capita monthly consumption could also be more accurately reported than annual household income. Therefore, consumption is used as a measure of the economic status of the household in the original regressions. Robustness checks are done with income as an alternative measure later on.

The level of education of the best-educated adult in the family could account for better home learning and reflect 'taste' for education. It could also mirror higher earning potential of the households. Being in a rural region could affect the cost and quality of school services available to a child. Being in a rural area may mean fewer schools, and may also mean higher opportunity costs due to immediate alternative of farm employment or child labor needs at home. State dummies are included to capture the administrative and policy-related costs and benefits of schooling, which might differ across states as education is to a large extent a state matter in India.

Since the base category 'Brahmin' is the caste that is conventionally considered to be in the most advantageous position, we would expect the signs on the remaining caste/religion variables to be negative. Simple summary statistics on the data stratified by social groups show that Muslims, SC and ST

have the lowest enrolment rates, income and consumption and adult education¹¹

The testable hypothesis of the paper is whether, even after controlling for correlates of social group association like economic status and household education, the negative signs on the coefficients of the social groups remain. Since more disadvantaged social groups also have lower average income and household adult education, we would expect the magnitude of the coefficients on social groups to diminish as we add more control variables.

Omitted variable bias in the estimates of social group effects could arise from the fact that there can be components in the error term that are correlated to belonging to a particular social group. For example, if some social groups form geographical enclaves within cities or villages, and quality of schooling and access to facilities vary accordingly, then omission of variables such as distance to school and school quality will create bias in the estimates of SRC effects. Such variables are not available by household in the data set used for this paper, and as a result, 'social group effects' are not controlled for school and other related community variables like civic amenities and infrastructure. If school quality and availability of related infrastructure are poorer for the lagging communities, the coefficients on social group dummies are expected to be higher than when such variables are controlled for.

¹¹ See Appendix B

DATA

The data for this paper comes from the India Human Development Survey (IHDS) 2005, a nationally representative survey of 41,554 households in 1,503 villages and 971 urban neighborhoods across India. The survey was conducted by the National Council of Applied Economic Research in Delhi and the University of Maryland.

The analysis is broken down into two categories: enrolment (where the dependent variable is whether the child is currently enrolled or not, named *ED4* in the data set) and attainment (the dependent variable is standard years of schooling, *ED5*). Primary school in India typically consists of grades 1-5, while middle school consists of grades 6-8 with minor regional variations by state. The typical age for completing of primary and middle school is 14 years. For enrolment, children of primary school age (6-14) and of secondary school age (15-18) will be analyzed in separate regressions.

The primary variables of interest are the socio-religious groups or communities, henceforth referred to as social groups or SRCs. The survey categorizes households into eight groups in the variable *GROUPS8*; Brahmins, high castes, other backward castes (OBC), Scheduled Castes (SC), Scheduled Tribes (ST), Sikhs or Jains and Christians¹². For brevity, Sikhs, Jains and Christians are categorized as 'Other' in this analysis. In the full sample, 33

¹²The households were asked the religion and caste of the head of the household.

percent of individuals are from OBC households, 20 percent SC, 16 percent high castes, 13 percent Muslims, 8 percent ST and about 5.6 percent Brahmins. Sikhs, Jains, Christians and others constitute the remaining 3 percent.¹³

Household economic resources are measured by the log of household monthly per capita consumption (*COPC*) spending on things like food items, utilities, rent, entertainment and services. The sample mean of monthly per capita consumption spending is Rs. 853. In addition to the age (*RO5*) and gender of the child (*RO3*), years of schooling of the most educated adult in the household (*HHED5ADULT*) is also controlled for (mean adult education is 7.8 years). A variable indicating whether the household is in a Census 2001 urban or rural area *URBAN* (66 percent of households in the full sample are rural), and state dummies (*STATEID*) for the 34 states in the country are also included. The largest state in the sample is Uttar Pradesh, followed by Karnataka, Maharashtra, Madhya Pradesh, Rajasthan and West Bengal.¹⁴

¹³See Appendix B, Table 2.

¹⁴See Appendix B, Table 1 for statistical summaries of the explanatory variables.

REGRESSION RESULTS

The analysis is broken down into enrolment (of children aged 6-14 and 15-18 respectively) and attainment (of age 17-18 cohort). For brevity, only coefficients on the social groups and how they change with the gradual addition of control variables will be discussed. The rest of the coefficients are reported in the regression tables. Probit regressions are run on the probability that a child is currently enrolled. For these regressions, the discussion of coefficient estimates of the social groups is cast in terms of the partial effect for an 'average' observation- that is, the effect of belonging to a social group on the predicted probability of current enrolment of an average child. For example, the partial effect of belonging to a Muslim household is;

$$\Phi (\beta_0 + \beta_1 \text{Age} + \beta_2 \text{Female} + \beta_7 [\text{Muslim} = 1] + \beta_k X_k) - \Phi (\beta_0 + \beta_1 \text{Age} + \beta_2 \text{Female} + \beta_k X_k)$$

All variables other than SRCs are evaluated at their mean in the full sample of children aged 6-14. To analyze years of schooling, I run OLS regressions on social groups and the rest of control variables. The implications of the results from various regressions are discussed in the final section of the paper.

1.1 ENROLMENT

1.1.1 Enrolment of children aged 6-14

The total sample size of children in this age bracket is 39,232, out of which 94.7 percent are currently enrolled. To avoid reporting error, individuals with years of schooling higher than age minus three were dropped. Out of the total sample, 53.8 percent are boys with an enrolment rate of 95.15 percent (94.29 for girls). *Observed* enrolment rates for the social groups in the sample are in the table below;

Table 1: Observed Enrolment Rate by Social Groups for Children Aged 6-14

Social Groups	Observed Enrolment Rate
Brahmin	98.71%
High caste	97.22%
OBC	95.48%
Dalit	93.58%
Adivasi	92.70%
Muslim	91.10%
Other	97.85%
Total	94.74%

Table A¹⁵ reports the regression coefficients and Table D.1 reports the partial effects of each social group. We start with a baseline regression of enrolment of age (with age-squared), gender and social groups in Model 1. The social group 'other' is insignificantly different to Brahmins. The most disadvantaged social groups in terms of current enrolment probability compared to Brahmins are Muslims, SC and ST with partial effects of -6.8, -3.6 and -6

¹⁵ Appendix A

percentage points (pp) respectively¹⁶. Being a girl makes you only 0.7 percentage points less likely to enroll. Interaction terms of gender with each social group in the model test insignificant. The linear term in age enters the model positively and the non-linear term enters negatively, implying that enrolment probability rises initially with age, and declines after reaching a peak.

Adding log of monthly per capita consumption, as a proxy for the household's economic resources Model 2 reduces the partial effects of social group association on enrolment by 1 pp each for SC and Muslims and by 2.7 pp for Scheduled Tribes. The fact that much more of the ST enrolment gap is explained away by consumption reflects that the lower enrolment among Adivasis, when compared to other groups, is due to their lower economic status.

When I add adult education in Model 3, the effects of being from Muslim, SC and ST households falls to -4.3, -1.5 and -2 percent respectively. Model 4 introduces the rural dummy, which is found to be insignificant. The effect of being in a rural area in a naïve regression (without controlling for household characteristics other than social groups) is around -1 pp, but is explained away once household consumption is controlled for. The rural dummy becomes significant again when state dummies are included.

In Model 5, after controlling for age, gender, economic status, adult education and location (with state dummies), Muslims are -3.5 percentage points less probable to enroll children of age 6-14 as compared to Brahmins. The

¹⁶Equivalent figures for high caste and other back castes are -1 and -2.6 percentage points. They become insignificantly different from Brahmins after controlling for adult education in model 3.

equivalent figures for Scheduled Caste and Scheduled Tribe households are -1.4 and -1.7 respectively.

1.1.2 Enrolment of Children aged 15-18

Out of the total 14906 children in the sample of this age group, 57.5 percent are enrolled now. Individuals with years of schooling higher than 11 years and with schooling higher than age minus three were dropped. Male children constitute 54 percent of the sample with enrolment rate of 60.3 percent (54.2 for girls). *Observed* enrolment rates for social groups are in the table below;

Table 2: Observed Enrolment Rate by Social Groups for Children Aged 6-14

Social Groups	Observed Enrolment Rate
Brahmin	76.82%
High caste	68.14%
OBC	56.67%
Dalit	53.54%
Adivasi	53.48%
Muslim	45.73%
Other	72.13%
Total	57.50%

There is a more pronounced observed gap between the different groups in the enrolment of children who are in the secondary education age-bracket. In Table B¹⁷, Model 1 regresses enrolment on age¹⁸, gender and social groups. It can be seen that compared to Brahmins, predicted probability of current enrolment is 30 pp lower for Muslims. The commensurate figures for ST, SC,

¹⁷ Appendix A: Table B reports coefficients, Table D.2 reports partial effects.

¹⁸ Age-squared is not included since the age range is small.

OBC and high caste children are 23, 22, 17 and 7 respectively.¹⁹ Girls are around 8.7 pp less likely to be enrolled in school, independent of their social groups since the interaction terms of gender and social groups are insignificant.

Around 5-8 percentage points of the SRC effects are explained away by difference in consumption per capita across households. For ST households, 14 pp of the initial partial effect is explained away by consumption, suggesting that a sizable part of the reason they are lagging behind is again because of their poorer economic condition.

Accounting for difference in adult education makes every SRC effect insignificantly different from Brahmins, except for Muslims. This implies an intergenerational disadvantage in schooling across SRCs. It also brings forth the strong favorable impact of parent or household-adult education. After the inclusion of rural and state dummies, a Muslim household is still -19 percentage points behind Brahmins.²⁰

The implication from these results is that social group association, after controlling for adult education is insignificant in determining school enrolment of older children, except for Muslims. For Muslims, the social group effect for older children is substantially higher than that for primary school-going children (-19 pp vs. -3.5 pp).

¹⁹Difference between 'other' and Brahmins is insignificant throughout.

²⁰In a naïve regression (without controls for consumption and adult education), being in a rural area is associated with 10 percent less probability of being enrolled in secondary education. After controlling for consumption, the effect reduces to about 3 percent less probability. However, once differences in adult education are accounted for, the rural dummy coefficient is both insignificant and close to zero.

1.1.3 Enrolment of children aged 15-18 with at least seven years of schooling

Some of the non-enrolled children in the 15-18 age-group might have dropped out in primary school. I next investigate the determinants of enrolment in secondary education itself. For this purpose, I focus on children with at least seven years of schooling. The sample size is reduced to 9,725, of which 71.7 percent are currently enrolled. 54.28 percent of the sample is male with 73.8 percent current enrolment; girls have an enrolment rate of 69.2 percent. Conditional on completing a certain amount of education, we might expect community-based gaps in current enrolment to be less severe due to selection effects.

As usual, we start with a naïve regression of current enrolment on age, gender and the social groups in Model 1 of Table C²¹. SRC effects are -7.2, -14, -17, -14 and -23 percentage points for high caste, OBC, SC, ST and Muslims respectively. Being in a rural area has a -3.8 percentage points' effect on enrolment probability, which becomes insignificant after controlling for household consumption. Notably, controlling for household consumption makes the ST effect insignificant, but doesn't explain more than 1 to 2 percentage points of the other SRC effects. Adding the state dummies makes being a high caste insignificant, however, even after full set of control variables, OBC, SC and Muslims have significant enrolment gaps (-5.5, -7 and -19.5 percentage points respectively). This indicates that when we focus on dropping-out decisions *in* secondary school – restricting the sample space to children with at least seven

²¹ Appendix A: Table C reports coefficients, Table D.3 reports partial effects.

years of completed schooling implies that the non-enrolled children have dropped out after finishing seven years of schooling – social groups have a higher impact on enrolment decisions. Differences between other social groups (e.g. HC and OBC, etc) are noted in Appendix A, Table E. The findings are that while the difference between Muslim and every other social group is significant, differences among almost all of the other groups are insignificant.

1.2 EDUCATIONAL ATTAINMENT

In the sample of 9433 seventeen and eighteen year olds, the mean years of schooling is 7.6 years. Fifty-two percent of the children in the sample are male, with mean years of schooling equal to 8 (the value for females is 7 years).

Observed average years of schooling across different social groups are in the following table; Christians, Sikhs, Jains and Brahmins have the highest mean years of schooling, while ST, Muslims and SC have the least.

Table 3: Observed: Mean Years of Schooling by Social Groups for 17-18 Cohort

Social Groups	Mean Years of Schooling
Brahmin	9.63
High caste	9.16
OBC	7.77
Dalit	6.73
Adivasi	6.08
Muslim	6.27
Other	9.84
Total	

Interactions of gender with Muslim, SC, ST and OBC are significantly different from the base category of male Brahmins. For this reason, the remaining regressions are run separately for male and female²².

For the cohort under study, the differences in mean years of schooling across social groups among boys reduce as we account for differences in endowment and education. After accounting for adult education, the difference between Brahmins and every social group other than Muslim become insignificant. For boys from Muslim households, the average years of school education is 1.4 years less than that of boys from Brahmin households after controlling for household consumption, adult education and location. The differences in years of schooling among female children across the social groups are higher. After controlling for consumption, adult education and location, a Muslim girl acquires 2.2 less years of education compared to a Brahmin girl. The commensurate figure for OBC, SC and ST are 1.28, 1.85 and 2 years respectively. As seen with probability of enrolment, Muslims are yet again the lowest ranked in terms of years of schooling acquired by its children in the past decade, followed by ST, SC and OBC.

1.3 ROBUSTNESS

Controlling household economic status with log annual income divided by the square-root of household size gives very similar coefficients on social groups in terms of sign and magnitude. Controlling for the full set with income as the measure of economic status, the persistent effects are -1.6, -2.3 and 3.9

²² Appendix A, Table F reports the regression results.

percentage points for OBC, SC, ST and Muslim respectively in the 6-14 enrolment regression (the respective figures with consumption as the control are -1.4, -1.7 and -3.5 respectively). In the 14-18 enrolment regression, as in the baseline specification, the coefficients for all social groups except for Muslims are insignificant. The partial effect for Muslims is similar to one from the original specification where consumption is used to control for economic status; namely -19.95 pp. In the attainment regression; SC, ST and Muslim remain significant after accounting for all the control variables with income as measure of economic status for male. The coefficients are -0.89, -1.02 and -1.67 (in the original specification, only Muslim was significant at -1.4). For female, all social groups test significantly different from Brahmin; -0.85, -1.48, -2.16, -2.42, -2.48 and -0.83 for high caste, OC, SC, ST, Muslim and 'other' respectively (comparable figures from the original specification are -1.28, -1.85, -2 and -2.2 for OBC, SC, ST and Muslim).

Therefore, the social group effects are fairly robust to the choice between annual income and consumption as controls for economic status in terms of sign and magnitude. At most, some social groups like high caste and 'other' become significant in regressions where it wasn't before.

DISCUSSION AND CONCLUSIONS

The analysis in this paper indicates that across the major socio-religious communities in India, enrolment probability and years of completed education differ significantly and persistently after controlling for family SES indicators and location. Notably, it is evident that children from Muslim households fare much worse in every avenue of schooling outcomes discussed above.

The magnitude of the coefficients on the social groups diminishes as more variables are controlled for, which could be construed to imply that social group effect is 'explained away' by correlates of social group association. It should however be noted that parental education is endogenous in the sense that their years of schooling is determined by their socio-religious association. Alternatively one can think of child-outcomes in a setting where the households they are born into are treated as random. An objective interpretation of the 'social group' coefficients remaining after a full set of control variables would be to regard them as enrolment or educational attainment gaps across social groups that are unexplained by the differences in economic status, household adult education and location. Subject to this caveat (of language and interpretation), following are the remaining conclusions from this paper.

The magnitude of the 'SRC-determined' gap is much lower for children of primary school-age, compared to older school children. Social group association stops being a significant factor after accounting for differences in adult education

for enrolment probabilities of older school children, except for Muslims. However, once we restrict the sample space to those who have completed at least seven years of schooling, we again find significant social group effects for OBC and SC in addition to Muslims. The finding that the social group effects are bigger for secondary school enrolment is a different result from that of studies done on black-white gap in the United States, where family SES factors mattered less in the later stages of schooling. Usually, as children complete more years in school, family and SES indicators matter less because those that remain in school are likely to be smarter and more motivated and would continue in school regardless of their backgrounds. However, this does not seem to be the case in India. Restricting the sample space to those who have completed at least seven years of school enables us to focus on the effect of social group on dropping-out decisions *in* secondary school, which is evidently bigger than in earlier stages of schooling. Lower labor market returns to education for members of lower castes and Muslims (as discussed later in this section) could be playing a more important role in discouraging further schooling at the stage of secondary education.

Another aspect that the results point to is the diversity across the lagging social groups in terms of how different family background correlates of caste and religion affect school participation. For children in the 6-14 age bracket, high caste and OBCs have lower enrolment rates that become insignificant after controlling for expenditure. Lower economic status explains a relatively bigger proportion of why children from ST households have lower enrolment probability

compared to other castes and religions. Conditional on having the same level of adult education in the household, enrolment of secondary school-age children from OBC, SC and ST households are insignificantly different from Brahmins. This suggests that there is a much brighter outlook for future generations of these households if current levels of schooling are improved. However, Muslim households do not fall in the same category. Higher adult education in the household does reduce the enrolment gap, but a Muslim child is still 19 pp less likely to be enrolled in school than an observationally equivalent Brahmin child between ages 15 and 18. On the bright side, children from Scheduled Tribe households with higher economic status are as likely to be currently enrolled as an observationally equivalent Brahmin after finishing seven years of schooling.²³

The use of terms like 'social group effect' in economics begs clearer thinking in terms of what are the true underlying causes of such effects. Two overarching explanations forwarded most frequently for this are culture and discrimination. Barooah and Iyer (2005) and Dreze and Kingdon (1999) respectively refer to the coefficients on SC, ST and Muslims that are unexplained by differences in household economic resources and location as 'cultural effect' and 'intrinsic disadvantage'. Cultural effect encapsulates the importance placed on education that could be 'community-specific', the status of women in that particular social group and the psychological barriers faced by children in

²³Desai and Kulkarni (2008) state that social exclusion pattern for dalits and adivasis may be quite different. Even though adivasis could face some discrimination, there is no religious basis for prejudice. When they move into urban areas, gain better income and have generationally been in an advantageous position, they may be less likely to face discrimination. A dalit however is one regardless of economic status; the Hindu society continues to see them as polluted and unacceptable.

attending school. Calling these community-specific 'culture' might be misleading since such things as the importance placed on education and the psychological barriers faced by children in attending school by different social groups are not exogenously determined. For instance, the expected return to education for being a dalit or a Muslim could be lower if they face systematic discrimination in the labor market. The Sachar Committee Report on Muslims in India (Government of India, 2006) reports that the perceived returns from education could be lower for Muslims as many do not see education as necessarily translating into formal employment (p 15). There is a perception of discrimination in securing salaried jobs, and Muslims are severely underrepresented in public sector employment.²⁴ The proportion of Muslims in government service in India was about 2 percent in 2002 (Barooah & Iyer, 2005). Similarly, the costs (tangible and/or psychological) of schooling faced by the disadvantaged groups could be different if schools and teachers discriminate against dalits and Muslims.²⁵

Other factors that systematically affect both the benefit and cost of schooling for these groups include the structure of formal schooling system and curricula. Lack of Urdu medium schools and teaching in Urdu²⁶, lack of tribal language instruction in primary schools and an overtly Hindu content and perspective on historical events have alienated many Muslim and tribal

²⁴In contrast to SC, ST and even OBC, Muslims have not been beneficiaries of affirmative action policies such as job reservations in the public sector.

²⁵Reference to differential treatment from teachers towards dalits, especially in villages, and how Muslim identity affects access to schooling have been made by Barooah & Iyer and Desai & Kulkarni. Malik (1999) reports a high degree of discrimination against dalit children in schools by other children and faculty. The village primary school might be located in a part of village where upper caste Hindus live.

²⁶Urdu is regarded as the primary language of Muslims (and non-Muslims) in North India, where it is spoken by 40 million people (Barooah & Iyer).

households (Barooah & Iyer, 2005) (Desai & Kulkarni, 2008). Yet another explanation that is anecdotally referred to in the Sachar Committee report is the ghettoization of Muslims in urban areas, especially in communally sensitive cities. Such enclaves are easy targets for neglect by municipalities which lead to inferior access to quality health facilities, schools and other civil amenities.

Enrolment gaps underestimate inequality in schooling since enrolment does not necessarily translate into regular attendance and more importantly, does not ensure satisfactory learning levels.²⁷ More data and research on how access to schools, schooling experience per se and labor market practices differ across social groups in the future for a more precise understanding of the processes through which socio-religious association lead to educational inequalities.

²⁷Adams, Desai et al (2008) find that reading and arithmetic skill levels are lower for the OBC, SC, ST and Muslims even after controlling for enrolment, grade completion, demography, SES and location.

APPENDICES

Appendix A Regression Results

This appendix presents the regression results.

Table A. Enrolment, Children aged 6-14²⁸
 Dependent Variable: Indicator Variable – whether currently enrolled

	Model 1	Model 2	Model 3	Model 4	Model 5
Age	0.3546** (0.0567)	0.3768** (0.0574)	0.4163** (0.0574)	0.4165** (0.0575)	0.4128** (0.0588)
Age-squared	-0.0255** (0.0026)	-0.0269** (0.0027)	-0.0290** (0.0027)	-0.0290** (0.0027)	-0.0292** (0.0027)
Female	-0.0845* (0.0364)	-0.0831* (0.0365)	-0.1026** (0.0376)	-0.1027** (0.0377)	-0.1142** (0.0368)
High Caste	-0.3166* (0.1430)	-0.3355* (0.1454)	-0.2197 (0.1527)	-0.2168 (0.1529)	-0.2443 (0.1574)
OBC	-0.5491** (0.1379)	-0.4495** (0.1402)	-0.2719 (0.1468)	-0.2721 (0.1472)	-0.2872 (0.1512)
SC	-0.6589** (0.1369)	-0.5317** (0.1403)	-0.3136* (0.1477)	-0.3121* (0.1480)	-0.3488* (0.1532)
ST	-0.8736** (0.1452)	-0.6181** (0.1497)	-0.3951* (0.1560)	-0.3963* (0.1564)	-0.4098* (0.1612)
Muslim	-0.9297** (0.1424)	-0.8513** (0.1445)	-0.6396** (0.1538)	-0.6310** (0.1547)	-0.6336** (0.1558)
Sikh/ Jain/ Christian	0.0036 (0.1746)	-0.0915 (0.1804)	0.0064 (0.1816)	0.0010 (0.1856)	-0.2027 (0.2023)
Log consumption pc		0.4075** (0.0335)	0.2947** (0.0358)	0.3029** (0.0360)	0.3174** (0.0402)
Adult education			0.0525** (0.0045)	0.0531** (0.0046)	0.0517** (0.0045)
Rural				0.0524 (0.0411)	0.1157** (0.0407)
State Dummies	No	No	No	No	Yes
Constant	1.5854** (0.3182)	-1.0999** (0.3746)	-1.0352** (0.3798)	-1.1335** (0.3789)	0.6096 (0.5079)
Pseudo RS	0.1260	0.1496	0.1719	0.1721	0.1916
LL	-7425.3	-7222.4	-7032.9	-7031.4	-6865.6
Observations	39232	39211	39211	39211	39211

²⁸Enrolment is estimated with probit regressions (coefficients reported are parameter estimates and not marginal effects). Figures in parenthesis immediately below are standard errors adjusted for clusters at the household level. (**) indicate significance at 1 percent, (*) indicate significance at 5 percent.

Table B. Enrolment, Children aged 15-18²⁹
 Dependent Variable: Indicator Variable – whether currently enrolled

	Model 1	Model 2	Model 3	Model 4	Model 5
Age	-0.3468** (0.0141)	-0.3692** (0.0140)	-0.3902** (0.0144)	-0.3902** (0.0145)	-0.3990** (0.0150)
Female	-0.2214* (0.0321)	-0.2239* (0.0323)	-0.2711** (0.0327)	-0.2711** (0.0327)	-0.2832** (0.0341)
High Caste	-0.2032* (0.1001)	-0.1990* (0.1011)	-0.0757 (0.1043)	-0.0754 (0.1041)	-0.0188 (0.1039)
OBC	-0.4648** (0.0948)	-0.3196** (0.0960)	-0.1475 (0.0996)	-0.1474 (0.0995)	-0.0718 (0.1006)
SC	-0.5866** (0.1012)	-0.4023** (0.1042)	-0.1603 (0.1095)	-0.1601 (0.1095)	-0.1315 (0.1035)
ST	-0.6354** (0.1088)	-0.3182** (0.1108)	-0.0744 (0.1168)	-0.0745 (0.1169)	-0.0076 (0.1230)
Muslim	-0.7971** (0.0985)	-0.6616** (0.1002)	-0.4735** (0.1040)	-0.4727** (0.1035)	-0.5072** (0.1055)
Sikh/ Jain/ Christian	0.0633 (0.1256)	-0.0145 (0.1459)	0.0893 (0.1461)	0.0897 (0.1458)	-0.0176 (0.1422)
Log consumption pc		0.5199** (0.0298)	0.4116** (0.0310)	0.4124** (0.0315)	0.3991** (0.0339)
Adult education			0.0598** (0.0038)	0.0598** (0.0038)	0.0589** (0.0040)
Rural				0.0041 (0.0348)	0.0392** (0.0375)
State Dummies	No	No	No	No	Yes
Constant	6.4035** (0.2631)	3.2954** (0.3257)	3.7763** (0.3383)	3.7684** (0.3357)	4.5936** (0.3670)
Pseudo RS	0.0910	0.1291	0.1562	0.1562	0.1757
LL	-9322.6	-8924.7	-8647.3	-8647.3	-8433.6
Observations	14906	14893	14893	14893	14868

²⁹Enrolment is estimated with probit regressions (coefficients reported are parameter estimates and not marginal effects). Figures immediately below each parameter estimate are p-values. (**) indicate significance at 1 percent, (*) indicate significance at 5 percent.

Table C. Enrolment conditional on 7 years of schooling, age 15-18³⁰
 Dependent Variable: Indicator Variable – whether currently enrolled

	Model 1	Model 2	Model 3	Model 4	Model 5
Age	-0.3917** (0.0169)	-0.4072** (0.0171)	-0.4163** (0.0175)	-0.4165** (0.0174)	-0.4314** (0.0181)
Female	-0.2163* (0.0391)	-0.2160* (0.0397)	-0.2467** (0.0402)	-0.2452** (0.0403)	-0.2628** (0.0397)
High Caste	-0.2691** (0.1024)	-0.2785** (0.1002)	-0.2128* (0.1007)	-0.2097* (0.1008)	-0.1664 (0.1019)
OBC	-0.4858** (0.0958)	-0.3715** (0.0944)	-0.3871** (0.0953)	-0.2860** (0.0953)	-0.1957* (0.0969)
SC	-0.5704** (0.1035)	-0.4141** (0.1030)	-0.2924** (0.1058)	-0.2906** (0.1060)	-0.2405* (0.1050)
ST	-0.4798** (0.1265)	-0.2261 (0.1266)	-0.0914 (0.1307)	-0.0926 (0.1307)	-0.0184 (0.1462)
Muslim	-0.7266** (0.1055)	-0.6441** (0.1042)	-0.5510** (0.1044)	-0.5445** (0.1047)	-0.5998** (0.1084)
Sikh/ Jain/ Christian	-0.2025 (0.1342)	-0.2716 (0.1535)	-0.2104 (0.1525)	-0.2087 (0.1524)	-0.2595 (0.1508)
Log consumption pc		0.4857** (0.0384)	0.4291** (0.0400)	0.4368** (0.0408)	0.4263** (0.0431)
Adult education			0.0347** (0.0048)	0.0352** (0.0048)	0.0339** (0.0050)
Rural				0.0298 (0.0420)	0.0323** (0.0443)
State Dummies	No	No	No	No	Yes
Constant	7.5641** (0.2965)	4.5593** (0.3834)	4.7341** (0.3915)	4.6545** (0.3968)	5.4051** (0.4337)
Pseudo RS	0.0994	0.1327	0.1423	0.1424	0.1655
LL	-5357.1	-5153.1	-5096.3	-5095.6	-4958.3
Observations	9725	9714	9714	9714	9714

³⁰Enrolment is estimated with probit regressions (coefficients reported are parameter estimates and not marginal effects). Figures immediately below each parameter estimate are p-values. (**) indicate significance at 1 percent, (*) indicate significance at 5 percent.

Table D. Partial Effects³¹

1. Enrolment of Children Aged 6-14

	Model 1	Model 2	Model 3	Model 4	Model 5
High Caste	-1.16*	-1.30*	-0.96	-0.95	-0.86
OBC	-2.64**	-1.99**	-1.25	-1.27	-1.07
SC	-3.60**	-2.59**	-1.51	-1.52	-1.39*
ST	-6.04**	-3.32**	-2.09	-2.12*	-1.76*
Muslim	-6.83**	-5.91**	-4.39**	-4.33**	-3.52**
Sikh/ Jain/ Christian	0.01	-0.27	0.02	0.00	-0.68

2. Enrolment of Children Aged 15-18

	Model 1	Model 2	Model 3	Model 4	Model 5
High Caste	-7.10*	-7.24*	-2.86	-2.85	-0.68
OBC	-17.10**	-11.88**	-5.64	-5.83	-2.71
SC	-21.94**	-15.13**	-6.13	-6.13	-4.97
ST	-23.89**	-11.82**	-2.81	-2.82	0.30
Muslim	-30.30**	-25.44**	-18.56**	-18.53**	-19.95**
Sikh/ Jain/ Christian	2.05	-0.51	3.28	3.29	-0.89

3. Enrolment of Children Aged 15-18 with At Least 7 Years of Schooling

	Model 1	Model 2	Model 3	Model 4	Model 5
High Caste	-7.20**	-7.68**	-6.08*	-6.00*	-4.66
OBC	-14.18**	-10.65**	-8.44**	-8.43**	-5.54*
SC	-17.16**	-12.06**	-8.61**	-8.58**	-6.94*
ST	-13.98**	-6.10	-2.48	-2.52	0.48
Muslim	-22.95**	-20.27**	-17.69**	-17.48**	-19.51**
Sikh/ Jain/ Christian	-5.25	-7.47	-6.00	-5.97	-7.54

³¹ Partial Effects are the difference in predicted probabilities between each social group and the base category Brahmin, with non-social group control variables evaluated at their sample mean. Figures are reported as percentage points. (**) indicate significance at 1 percent and (*) at 5 percent.

Table E. Enrolment Gaps across Social Groups ³²

1. Children Aged 6-14

	OBC	SC	ST	Muslim	Other
HC	0.0332 (0.0661)	0.0894 (0.0649)	0.1533 (0.0851)	0.3719** (0.0733)	-0.0596 (0.1493)
OBC		0.0562 (0.0509)	0.1201 (0.0750)	0.3387** (0.0612)	-0.0928 (0.1466)
SC			0.0639 (0.0733)	0.2825** (0.0585)	-0.1490 (0.1420)
ST				0.2185** (0.0842)	-0.2130 (0.1564)
Muslim					-0.4315** (0.1490)

³² Figures reported are the difference in the parameter estimates between groups. For example, the first cell (HC, OBC) is $\hat{\beta}_{hc} - \hat{\beta}_{obc}$ and the value in parenthesis immediately underneath is the standard error of the difference. (**) implies significance at 1 percent and (*) at 5 percent. The parameter estimates are from regressions that are controlled for age and gender of the child, household monthly per capita consumption, adult education, rural dummy and state dummies.

2. Children Aged 15-18

	OBC	SC	ST	Muslim	Other
HC	0.0534 (0.0537)	0.1120 (0.0584)	-0.0263 (0.0921)	0.4902** (0.0631)	0.0585 (0.0526)
OBC	-	0.0585 (0.0526)	-0.0798 (0.0833)	0.4367** (0.0552)	-0.0480 (0.1097)
SC		-	-0.1384 (0.0889)	0.3781** (0.0595)	-0.1066 (0.1145)
ST			-	0.5166** (0.0902)	0.0317 (0.1336)
Muslim				-	-0.4848** (0.1160)

3. Children Aged 15-18 with At Least 7 Years of Schooling

	OBC	SC	ST	Muslim	Other
HC	0.0293 (0.0622)	0.0741 (0.0738)	-0.1848 (0.1254)	0.04334** (0.0805)	0.0931 (0.1313)
OBC		0.0448 (0.0625)	-0.2141 (0.1160)	0.4041** (0.0731)	0.0638 (0.1269)
SC			-0.2589* (0.1252)	0.3593** (0.0813)	0.0190 (0.1315)
ST				0.6182** (0.1296)	0.2780 (0.1679)
Muslim					-0.3402* (0.1364)

Table F. Educational Attainment, 17-18 Cohort by Gender³³ :

Male

Dependent Variable: Years of Completed Schooling

	Model 1	Model 2	Model 3	Model 4
High caste	-0.1015 (0.3607)	-0.0310 (0.3102)	0.3706 (0.3211)	0.0047 (0.3192)
OBC	-1.3022** (0.3572)	-0.6431* (0.3104)	0.1849 (0.3221)	-0.0748 (0.3270)
SC	-2.3775** (0.3819)	-1.5009** (0.3387)	-0.3661 (0.3583)	-0.6902 (0.3563)
ST	-2.8103** (0.4359)	-1.4740** (0.4045)	-0.3150 (0.4112)	-0.7546 (0.4078)
Muslim	-2.8835** (0.3782)	-2.0967** (0.3373)	-1.2034** (0.3508)	-1.4062** (0.3508)
Sikh/Jain/Christ	0.6779 (0.3788)	0.0593** (0.3449)	0.2899 (0.3415)	-0.2915 (0.3605)
Log consumption pc		1.8721** (0.1115)	1.2742** (0.1151)	1.1461** (0.1248)
Adult education			0.2627** (0.0146)	0.2619** (0.0143)
Rural				-0.0591 (0.1293)
State dummies	No	No	No	Yes
Constant	9.2983** (0.3291)	-3.4849** (0.8705)	-2.1973* (0.8634)	0.7639 (0.9461)
R2	0.0782	0.1754	0.2841	0.3312
Observations	4915	4909	4909	4909

³³Attainment is estimated with OLS. Figures immediately below each parameter estimate in parentheses are standard errors adjusted for clusters at the household level. (**) indicate significance at 1 percent, (*) indicate significance at 5 percent.

Female

Dependent Variable: Years of Completed Schooling

	Model 1	Model 2	Model 3	Model 4
High caste	-0.8793* (0.3647)	-0.8307 (0.4390)	-0.1956 (0.3745)	-0.8552 (0.3867)
OBC	-2.9785** (0.3572)	-2.0419* (0.4426)	0.8557* (0.3713)	-1.2816** (0.3915)
SC	-4.0453** (0.3810)	-2.8582** (0.4734)	-1.2345** (0.4110)	-1.8545** (0.4240)
ST	-4.5040** (0.4092)	-2.6521** (0.5043)	-1.1413** (0.4418)	-1.9983** (0.4618)
Muslim	-4.1997** (0.3755)	-3.2109** (0.4655)	-1.6020** (0.3885)	-2.2259** (0.4051)
Sikh/Jain/Christ	0.4194 (0.3764)	0.5775** (0.5435)	1.1602* (0.4566)	-0.7099 (0.4360)
Log consumption pc		2.3269** (0.1348)	1.4306** (0.1255)	1.0242** (0.1247)
Adult education			0.3903** (0.0180)	0.3516** (0.0185)
Rural				-0.9763** (0.1389)
State dummies	No	No	No	Yes
Constant	9.6157** (0.3117)	-6.2627** (1.0519)	-4.5029** (0.8907)	2.4573** (0.9088)
R2	0.1036	0.2117	0.3690	0.4549
Observations	4518	4511	4511	4511

Appendix B Descriptive Statistics

A. Descriptive Statistics of Variables³⁴

Variable	Age 6-14	Age 15-18	Age 17-18	Full Sample
# Obs	39232	14906	9433	213060
Enrolled Now ⁱ	0.95	0.58	-	-
Mean Years of schooling	-	-	7.57 (3.91)	4.65 (4.69)
Age	10.13 (2.60)	16.40 (1.15)	-	27.33 (19.33)
Female ⁱ	0.47	0.46	0.48	0.49
Brahmin ⁱ	0.05	0.05	0.05	0.05
High Caste ⁱ	0.15	0.15	0.16	0.16
OBC ⁱ	0.33	0.33	0.33	0.33
SC ⁱ	0.21	0.21	0.20	0.20
ST ⁱ	0.08	0.08	0.09	0.08
Muslim ⁱ	0.14	0.14	0.15	0.13
Other ⁱ	0.03	0.03	0.03	0.03
Monthly per capita consumption	760.59 (685.86)	847.17 (736.98)	884.52 (820.08)	854.62 (862.75)
Adult education	7.06 (4.86)	7.19 (4.77)	7.24 (5.00)	7.83 (5.01)
Rural ⁱ	0.68	0.66	0.65	0.66

³⁴ Reported figures are means (and standard deviations in parenthesis below) of continuous variables, and proportions in the case of indicator variables. 'i' indicates that the variable is an indicator.

B. Selected Statistics by Social Group (in the Full Sample)

Social Group	Per capita consumption	Mean household annual income	Mean years of schooling	Mean household adult education	Proportion rural
Brahmin	1273.26	92373.64	7.17	11.52	0.51
High Caste	1154.92	86424.58	6.18	9.94	0.57
OBC	796.97	53755.14	4.59	7.86	0.70
SC	690.52	43492.35	3.68	6.36	0.72
ST	538.27	44154.53	3.29	5.52	0.86
Muslim	763.49	56820.74	3.74	6.48	0.54
Sikh/Jain	1433.35	133626.8	6.67	10.61	0.59
Christian	1334.44	89331.79	7.38	10.72	0.57
Total	855.04	60736.60	4.68	7.83	0.66

BIBLIOGRAPHY

- Barro, S., & Kolstad, A. (1987). Who drops out of high school? Findings from high school and beyond (Report 87-397). U.S. Department of Education/ National Center for Education Statistics. Washington, DC.
- Borooah, V.K. & Iyer, S. (2005). Vidya, Veda and Varna: the influence of religion and caste on education in rural India. *Journal of Development Studies*, 41(8), 1369-1404
- Desai, S. & Kulkarni, V. (2008). Changing educational inequalities in India in the context of affirmative action. *Demography* 45(2), 245-270
- Desai, S., Cecily Darden Adams, & Amaresh Dubey. Forthcoming. Segmented Schooling: Inequalities in Primary Education. In *Blocked by Caste: Discrimination and Social Exclusion in Modern India*, edited by K. Newman and S. Thorat. New Delhi: Oxford University Press
- Dreze, J., & Kingdon, G. (2001). School participation in rural India. *Review of Development Economics*, 5(1), 1-24
- Duraisamy, M. (2000). Child schooling and child work in India. IIAT, Madras, India. Presented at *the Eighth World Congress of the Econometrics Society*, Seattle, Washington (August 11-16)
- Filmer D., & Pritchett, L. (1999) Educational enrolment and attainment in India: household wealth, gender, village and state effects', *Journal of Educational Planning and Administration* Vol. 13(2)
- Glewwe, P. & Kremer, M. (2006). Schools, teachers, and educational outcomes in developing countries. In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland: 943-1017
- Holmes, J. (1997). Measuring the determinants of school completion in Pakistan: analysis of censoring and selection. *Economic Growth Center Discussion paper No. 794*. New Haven, CT: Yale University.

- Kremer, M., Karthik Muralidharan, Nazmul Chaudhury, Jeffrey Hammer and F. Halsey Rogers. (2005). Teacher Absence in India. *Journal of the European Economic Association*, 3, 658-67
- Mare, Robert D. (1980). Social background and school continuation decisions. *Journal of American Statistical Association*, 75(370), 295-305
- Malik, B. (1999). Untouchability and Dalit Women's Oppression. *Economic and Political Weekly*, 6 February
- Government of India. (2006). Social, Economic and Educational Status of the Muslim Community in India. New Delhi: Government of India
- Desai, Sonalde, Reeve Vanneman, and National Council of Applied Economic Research, New Delhi. India Human Development Survey (IHDS), 2005 [Computer file]. ICPSR22626-v5. Ann Arbor, MI; Inter-university Consortium for Political and Social Research [distributor], 2009-06-22. doi:10.3886/ICPSR22626