The London School of Economics and Political Science

Economics of Social, Gender, and Income Inequalities

Sutanuka Roy

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy, London, March 2018

Declaration

I certify that the thesis I have presented for examination for the MRes/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of about 33,000 words.

Statement of Conjoint Work

I confirm that Chapter 2 is jointly co-authored with Dr. H.F. Tam, all authors contributed equally to the project. Chapter 3 is jointly co-authored with Dr. Prakarsh Singh, all authors contributed equally to the project.

Statement of use of third party for editorial help

I can confirm that my thesis was copy edited for conventions of language, spelling and grammar by LSE proof reading service.

Acknowledgement

I am deeply grateful to Prof. Tim Besley for guiding my PhD and being my teacher and mentor. I am very thankful to Prof. Maitreesh Ghatak, Dr. Alessandro Tarozzi and Dr. Gharad Bryan for their supervision. I would like to thank Prof. Herakles Polemarchakis for teaching me economics and for supporting me throughout my PhD.

I would like to thank my coauthors: Dr. H.F. Tam and Dr. Prakarsh Singh. I am thankful for the feedback I have received from Dr. Guy Michaels, Prof. John List, Prof. Alwyn Young, Dr. Marcia Schafgans. Special thanks go to Mark WiIbor and Charlotte Knights for their kindness and help. I would like to thank all my research assistants, college principals for making the experiment possible. I would like to thank my friends who made my research years at the LSE enjoyable. Finally, I am grateful to my parents for supporting me.

Abstract

The thesis contains three chapters. The first chapter reports on the first large-scale randomized field experiment involving legally-recognized minorities to examine the causal effects of providing performance-based financial incentives based on social or income disadvantage on high stakes university test scores. The results are that the average test scores of the whole cohort goes down by .14 standard deviations when financial incentives were provided by income disadvantage while there is no effect on the test scores when financial incentives were provided by social disadvantage or when financial incentives were provided to all students. The chapter provides evidence of academic non-cooperation when financial incentives are offered by income status and no evidence of such peer effects when prize incentives are given by social disadvantage.

The second chapter, which is a joint work with Dr. H.F.Tam, studies the impact of matrimonial laws introduced by the British in British provinces in colonial India during 1800s and early 1900s. Exploiting quasi-random variations of districts that were former British Provinces within each post-independent Indian states, we find that females have 5% lower chances of marrying under the current legal age, and 1.6% higher chance of attending school at 10-16 years old in regions that were formerly British Provinces. Furthermore, using historical Census of India 1901-1931 on marriage status of population between 0-15 years at district level, the chapter estimates the impact of Child Marriage abolition Act (1931) on child marriages in colonial India.

The third chapter uses a large-scale novel panel dataset (2005-14) on schools from the Indian state of Assam to test for the impact of violent conflict on female student's enrolment ratios. We find that a doubling of average killings in a district-year leads to a 13 per cent drop in girl's enrolment rate with school fixed effects.

Contents

[1]This chapter was written jointly with Dr. H.F. Tam.
[2]This paper was coauthored with Dr. Prakarsh Singh

List of Figures

List of Tables

Chapter 1

Field Study of Preferential Incentives

This paper reports on the first large-scale randomized field experiment (14,190 undergraduate students) involving legally-recognized minorities to examine the causal effects of providing performance based financial incentives to disadvantaged students on high stakes university test scores. Two definitions of being disadvantaged are examined separately: 1) income disadvantage, and 2) social disadvantage of belonging to minority groups, i.e., the lower caste groups. The paper aims to measure the impact of two types of affirmative action policies on the disadvantaged groups that the policies target and on the excluded relatively advantaged peers. When only poor students were given the opportunity to win the prize incentives, the average test scores of the whole cohort decreased by .14 standard deviations. There is also a negative spillover effect on the test scores of the nonpoor peers who are excluded from the opportunity to win the prize incentives. Mechanisms of academic non-co-operation as a response to preferential policies are tested. The paper provides evidence of social tension and consequent non-cooperation among peers when only poor students are incentivized and the majority of the peers, who happen not to be poor, are excluded.

1.1 Introduction

According to the All India Survey on Higher Education (2011-2012), 48.3% of all students enrolled in higher education in India come from a historically discriminated minority background, i.e., lower-caste groups. Lower castes are the legally recognized minorities in India which include Scheduled Castes (SC), Scheduled Tribes (ST) and

Other Backward Caste (OBC) groups.[1] Despite high enrolment rate of lower caste students in Indian universities, there exists a persistent educational attainment gap between the lower caste and upper caste students, comparable to the black and white student achievement gap in the U.S (Desai and Kulkarni (2008)).

In light of the above problem, common solutions sought by the policymakers in India are affirmative action policies which aim to provide incentives for students from disadvantaged economic and social backgrounds. These schemes are merit cum means scholarship schemes, which are provided in most universities in India and are presented in the form of tournament incentives among disadvantaged students. Tournament incentives are schemes where prizes are awarded on the basis of relative performance of the students, for example, a prize for the highest scorer in an exam is a tournament incentive scheme. Merit cum means schemes are tournament prize schemes where students from disadvantaged background compete only among themselves for a prize. Disadvantage is variously defined by the policymakers. Some schemes are given by income status where poor students from both minority and non-minority background compete and some schemes are given by minority status, i.e., lower caste status, where the lower caste students from both low and high income households compete.

As of 2009, the Indian Government introduced national merit cum means scholarships, each worth about USD 152 per year for roughly three years, to be provided to 82,000 undergraduate students from disadvantaged backgrounds. The objective on the part of the government for introducing such tournament incentives is to use limited government resources for selecting and providing financial assistance to the ablest candidates among disadvantaged groups: "To promote qualitative education amongst SC students, by providing full financial support for pursuing studies beyond 12th class (...). The general selection criteria among the eligible candidates of any institution must be the merit.(Ministry of Justice and Social Empowerment, 2017)" Besides, the motivation provided by the policy makers, there is an economic argument for the need to incentivize students from weaker socio-economic groups. The economic justification for providing incentives to disadvantaged groups is that market frictions (such as perceived low returns to education and discrimination) beyond liquidity constraints are known to affect human capital investment as well as the labor market outcomes of discriminated social and economic groups in developing countries (Attanasio et al. (2011); Carneiro et al. (2011); Jensen (2010); Banerjee

[1] According to Census of India, 2011, 16.6% of the population are Scheduled Castes, and 8.6% of the population are Scheduled Tribes.

et al. (2007); Hanna and Linden (2012)). Therefore, it is possible that students who are financially disadvantaged or students who are socially disadvantaged or students who are handicapped both socially and economically are not optimally investing in education. While there is an economic motivation to incentivize students from disadvantaged backgrounds, little is known about how the beneficiaries and the non beneficiaries respond to such incentives and how it impacts their test scores.

This chapter examines the effectiveness of the above merit cum means tournament scholarship schemes, which are implemented nationwide to provide a level playing field to the poor or historically discriminated minorities. Two definitions of being disadvantaged are examined: 1) financial disadvantage 2) social disadvantage of belonging to minority groups, i.e., the lower caste groups. We incentivize the poor students from all social backgrounds to explore the incentive effects of tournaments on the financially disadvantaged students. To examine the impacts on socially disadvantaged, we incentivize the lower caste groups from all income background. The novel contribution of the paper is that it provides evidence on the impact of preferential policies on the excluded advantaged social groups in an environment where both advantaged and disadvantaged social groups work together. Since most affirmative action policies are implemented in contexts of tournaments, our findings on the effectiveness of merit cum means tournament incentives could potentially be relevant for many other affirmative action policies (Hickman (2013); Calsamiglia et al. (2013); Sheremeta (2015); Fryer and Loury (2005)).

To investigate the impact of the merit cum means scholarship policy, in this paper the author implemented a large-scale randomized field experiment involving more than 14,000 students, where prize incentives for the poor, the lower caste groups i.e, Scheduled Castes, Scheduled Tribes and Other Backward Castes, and prize incentives for all were randomized across 470 cohorts in Indian universities. To the best of our knowledge, this is the first paper to randomize actual government affirmative action policy, involving legally recognized minorities. From now on we will refer to the prize incentives for all as *untargeted schemes*, prize incentives for the poor as *income or economically targeted schemes* and prize incentives for the lower castes as *socially targeted schemes*. Under untargeted tournaments, students from all backgrounds compete. Under income targeted tournaments, upper and lower caste students from low income households compete and under socially targeted tournaments, rich and poor lower caste students compete for the prizes.

The tournament incentives in the experiment were given out on the basis of the performance in high stakes university exams. Our treatments change both the size and the demographic composition of the contests. The incentives that were tested included untargeted incentives (i.e., prize incentives for all), where two financial rewards were given to students in a cohort. The first monetary prize of USD 62 was awarded to the student with the highest grade point average (GPA) in the cohort, and the second monetary award of USD 54 went to the student in the top 40% with the highest GPA improvement in the cohort. On average, a cohort had about 50 students. Thus, when prize incentives were provided for all, on average 50 students competed for the two monetary prizes. The second treatment arm was socially-targeted incentives (i.e., prize incentives only for the lower castes), where the same two financial incentives were only given to the lower caste students within a cohort: the first to the highest scorer from the lower castes in the cohort, and the second to the student who was in the top 40% among the lower caste students in the cohort and who showed the highest improvement in scores when compared with previous university test scores. On average, 50% of the students within a cohort belong to the lower caste groups. Thus, when prize incentives for only the lower castes were provided, 25 students on average competed with each other for the two prizes. The third treatment was income targeting (i.e. prize incentives only for the poor), where the same two financial incentives were given to students with an annual family income of USD 2,340 or less. Approximately 10% of all students within a cohort were students with an annual family income of USD 2,340 or less. Therefore, when prize incentives are provided only for the poor, on average, only 5 students, competed for the two prizes. The control group comprised of students in cohorts with no incentives.

In order to capture the effects of targeting (i.e., providing incentives only to the lower castes or only to the poor students) on non-beneficiary peers, incentives were randomized across cohorts, where a cohort was defined as an entire group of students studying the same degree, in the same degree year, and at the same college (e.g. the entire first year of undergraduate students studying economics at a particular educational institution is a cohort). The spillover effects on non-beneficiaries were estimated by comparing the means of the advantaged students in a cohort whose disadvantaged peers were incentivized with advantaged students in a cohort whose disadvantaged classmates were not incentivized.

Tournament incentives are known to generate high powered incentives for exerting effort among participants, particularly if contestants have similar abilities and socio-economic backgrounds (Sheremeta (2015)). However, such incentives can result in unintended impacts on non-beneficiaries. The psychological literature on affirmative action policies presents evidence of a backlash by non-beneficiaries, who may feel unfairly excluded by affirmative action policies (Shteynberg et al. (2011); Harrison et al. (2006); Lynch (1989)).

Our paper provides the first causal evidence of a negative impact of targeted tournament incentives. It also provides first evidence of negative spillover effects on the peers within the cohort who were excluded from the opportunity to win the prize incentives. Surprisingly, the negative effects were statistically and economically significant only in income targeted tournaments, where poor students were provided with monetary incentives. Under income targeted tournaments, 10% of the cohort became eligible to compete for the monetary incentives. The mean deterioration in scores in entire cohorts where only poor students were incentivized was .14 standard deviation (SDs) of student performance, compared to the cohort with no incentives. Socially-targeted incentives, i.e., prize incentives only for the lower castes, incentivized, on average, 50% of the cohort and produced no average effect on test scores, as was the case for untargeted pure merit tournaments. In Kremer et al. (2009), tournament incentives for girls increased test scores of girls by .19 SDs and also increased the performance of boys. However, the study indicates that the attendance of teachers improved due to school-based student incentives, which can potentially explain the positive spillover on boys. In our case of student incentives, institutional channels were not present, due to a high rate of student absenteeism and poor quality teaching. The parameter that was relevant in our setting was peer interaction among colleagues, which was probably less salient among primary school students in incentivized Kenyan schools. In Bertrand et al. (2010), there is suggestive evidence of a negative impact of affirmative action policies on the labor market outcomes of non-beneficiaries, but the limited sample size prevented the authors from drawing such conclusions.

We investigated whether non-beneficiary peers disapprove of preferential incentives. Non-beneficiaries may find it unfair to be excluded from a prize competition, and they may become upset and demotivated and refuse to cooperate with weaker students who are included in the tournament. The demotivation effect of being excluded from an opportunity can negatively impact on their test scores, while non-cooperation with low ability students who are given an opportunity can affect the performance of disadvantaged students, who, on average, have very low test scores and who are likely not to be familiar with the education production function (Hanushek et al. (2008); Krueger (1999)).

An attitude survey administered post-announcement of prize incentives and prior to university exams delivers evidence of tensions among cohorts that had income-targeted tournaments. We further investigated whether resourceful and close friendship ties, which are unlikely to dissolve due to one-time incentives, could have a beneficial effect on academically-weak students. Using pre-intervention network data, we find that low ability students who sorted themselves into friendship networks with high ability students prior to our intervention experienced a positive effect on test scores under income-targeted incentives. This suggests that the friendship networks of disadvantaged groups can potentially play a role in targeted incentives. To investigate further whether friendship networks play a role because colleagues who are not close friends refuse to cooperate, we analyzed post-intervention network data collected from a sub-sample. We provide evidence that income targeting caused poor students to seek support from their closest friends who are also low ability students like themselves. However, since the response rate of the post-intervention network data was very low, the results on changing peer interaction should be interpreted with caution.

This paper has several implications for policy. The most striking finding in our paper is that the academic performance of students excluded from the tournament competition significantly decreased relative to no tournaments. More importantly, the targeted incentive, which excluded the largest proportion of students, had the largest negative effect. This result has important implications for affirmative policies in weak institutions. In developing economies, where institutions are weak, policies that are designed to aid highly disadvantaged groups will exclude large sections of resourceful social groups, who may find such policies unfair (Patilder Reservation Agitation, 2017; Jat Quota Protests, 2017; Anti-Reservation Student Protests, 2006) and therefore refuse to cooperate with weaker demographic groups. This can have a potentially a negative impact on the output of all social groups.

We further find evidence that students with resourceful friends respond positively to incentives. This implies that the negative effect of preferential policies can potentially be mitigated by institutional support. This paper provides support for examining the role of institutional support in neutralizing the disincentive effects of incentive schemes designed to help the needy.

The negative effects on performance as a result of economic targeting are not in line with the empirical findings in the literature on performance-based incentives (Burgess et al. (2016); Fryer Jr et al. (2015); Levitt et al. (2011); Kremer et al. (2009), Angrist et al. (2009); Glewwe and Kremer (2006); Van der Klaauw (2002);

Hartog and Oosterbeek (1988)). It may be that negative peer effect parameters did not operate in the contexts in which previous studies were conducted.

This paper relates to the literature on experimental economics, that tests the link between affirmative action policies and effort and performance. Examples include Schotter and Weigelt (1992), where asymmetry was exogenously imposed by researchers assigning cost functions to subjects, and Bracha et al. (2015), who focused on gender-based asymmetry in quantitative problem solving. Calsamiglia et al. (2013) conducted a field experiment in which 10 to 13 year old children competed Sudoku puzzles in a pair-wise (one-to-one) tournament, with asymmetry stemming from previous exposure to Sudoku. Hickman (2013) created disadvantaged groups by making school children from different grades compete with each other, with school children from lower grades being viewed as the disadvantaged group. Our paper makes a step forward by examining the incentive effects of affirmative action policies by randomizing preferential treatments over legally recognized, discriminated against minorities. In the real world, disadvantaged groups can experience severe resource constraints when compared to non-minorities. Additionally, the minority groups may also have nonminority friends or colleagues. Pro-minority policies may antagonize nonminority colleagues. By involving minorities in the experiment, the current study complements the above literature by capturing the actual constraints that could impact the outcome of the minorities. This can also help explain why the results of the current paper differ from the experiments noted above, which found positive effects on the performance of disadvantaged groups

The paper makes a unique contribution to the literature concerning tournaments (Sheremeta (2015)). We find that tournament incentives can affect on the outcomes of individuals who are excluded from opportunities to compete and be recognized for their performance. This is particularly relevant in the setting of equal opportunity laws, which primarily arise in the form of tournaments where weaker social groups have a handicap.

The work also contributes to the literature on peer effects in education (Burns et al. (2015); Lavy and Sand (2016); Lavy et al. (2012); Carrell et al. (2009); Hoxby (2000); Sacerdote (2001)). The literature on peer effects have shown that social interactions among peers matter for academic performance. Our paper contributes to the literature by examining the causal impact of financial incentives on test scores, given that peer effects are at work.

The organization of the paper is as follows. Sections 2 and 3 describe the experiment details and the data, while Section 4 describes the model. Section 5 outlines the estimation strategy, Section 6 presents the main findings, and Section 7 summaries the heterogeneous effects of income targeting. Section 8 reviews the robustness checks and Section 9 concludes the paper.

1.2 Background and Field Experiment Details

The experiment was conducted in the states of Assam and Meghalaya. Assam is an extremely poor state with a state GDP of approx. USD 781 (Directorate of Economics and Statistics, Government of Assam, 2016) and 32% of the population below the poverty line (Ministry of Development of North Eastern Region, Government of India, 2011-2012), relative to the national average of 22%. In Meghalaya, 16% of the population are below the poverty line (Ministry of Development of North Eastern Region, Government of India, 2011-2012). The gross enrolment ratios in higher education in Assam and Meghalaya are 12.8 and 17.3, respectively (NITI Aayog, 2011-2012) compared to the national average of 24.3 (NITI Aayog, 2011-2012).

The experiment was conducted in collegiate universities. Most of the state and central government universities in India are collegiate universities. A collegiate university is a university in which the governing authority and functions are divided between a central administration and various constituent colleges. The constituent colleges are not merely halls of residences but have their own management, student unions, and academic faculties, and they require different cut-off scores for admissions. All students in the colleges take centralized university exams every six months, which are marked anonymously by external examiners organized by the university. The mark sheets and degrees are awarded by the university. There are about 760 universities in India, with almost 40,000 affiliated colleges. Each university has, on average, about 200 to 560 affiliated colleges. In India, about 42% of the universities are either central universities or state public universities (AISHE Report, Ministry of Human Resources, 2016). Assam has a total of six state and central government universities, of which four are included in this study's sample. In Meghalaya, there are five universities in total, of which the largest university is included in the sample.

All public colleges in India are legally required to implement affirmative action policies in the form of caste reservation quotas. Since 1950, the Indian constitution prohibits discrimination based on caste and has put affirmative action measures in place to protect members of disadvantaged groups, such as Scheduled Castes and Scheduled Tribes. In 1990, provisions were made to include Other Backward Castes (OBCs) in the caste-based reservation system. Reservation is a form of quota-based affirmative action governed by constitutional and statutory laws, as well as by local rules and regulations. The quota system sets aside a proportion of all possible positions for members of a specific group. Those not belonging to the designated communities can only compete for the remaining positions, while members of the designated communities can compete for all positions (both reserved and open). Enrolment in colleges and universities and vacancies in public institutions have a total of 49.5% constitutional reservations. Meghalaya, being a tribal minority state, has reservations totaling approximately 90% for tribal minorities.

The public higher-education institutions in India are highly subsidized, are of high quality, and have almost no tuition fees; however, supply is limited and competition for admissions into public collegiate universities is extremely high. Entry into higher educational institutions in India is based solely on the test scores of standardized school-leaving exams. Students belonging to lower castes are given preference in college admissions by having lower cut-off entry score requirements. The difference in cut-off scores for admission between those given preferences and the upper castes is substantial (Krishna and Tarasov (2016); Bertrand et al. (2010)).

One key feature of the research design is that only students who study major subjects were considered for the experiment. Students who have majors take a fixed set of compulsory exams, which is common for all students studying that major. The classification of a degree for a major student is determined by their performance in the university exams for their major subject. The students studying the same major in the same degree year spend many lecture hours together in college because they take a common set of courses.

1.2.1 Randomization

To begin the field experiment, the standard protocol was followed. Firstly, permission for the experiment was secured from the university vice chancellors and college principals in the States of Assam and Meghalaya. A list of 11 public colleges were randomly selected using a random number generator from a list of over 40 colleges, where permission was granted, and that were in the town of Shillong, Hailakandi and Guwahati, where the author had accommodation.

The experiment included 14,190 college undergraduate students in all years. There were 10 non-engineering colleges and 1 engineering college.

All non engineering colleges have on average two faculties: faculties of arts, faculty of science. Arts and sciences faculties on average are subdivided into 8-9 departments for example, faculty of arts has department of sociology, department of political science and so on and faculty of science has department of physics, department of chemistry and so on. Each department of all non-engineering colleges has 3 cohorts: 1st year students, 2nd year students and 3rd year students.

The engineering college is divided into 10 faculties. The faculties in the engineering college are not subdivided into departments. Each faculty has 4 cohorts such as 1st year students, 2nd year students, 3rd year students, 4th year students.

Our unit of randomization was a cohort of students enrolled on degree D under faculty f studying in the year t in college C at university U. For example, a class of students studying history major during their first year of a degree program at St Johns College, which is affiliated with the University of U, is our unit of randomization.

Simple stratified random sampling was used to randomize the treatment incentives. The sample was stratified by the following covariates (wherever possible): i) university ii) college iii) faculty, such as faculty of arts or faculty of sciences iv) degree year of study.[2] To avoid potential inference problems stemming from the way the sample was stratified, we will do our analysis using randomization inference test that is robust to the problem of computing correct clustered robust standard errors. Additionally, we will divide our sample by colleges that were stratified similarly (e.g. by dropping engineering college) and test the significance of our results using different levels of clustering and report the clustered robust standard errors.

Within each strata, cohorts were randomized into three treatments and control groups using random number generators. A fixed rule of having equal proportion of treatment units in every strata was implemented. Each strata of non engineering colleges had on average of about 8-9 departments. Students were compared across departments but within the same degree year and within the same faculty and within the same college. Most of the non engineering colleges have arts and science faculties. Some non engineering colleges have arts, science and commerce faculties. Some non engineering colleges have only arts faculty. Two non engineering colleges have only first and second year students. In total, non engineering colleges make

[2]The engineering college could not be further stratified at degree year level because the faculties were not subdivided into departments to enable stratification by degree year. In engineering college, faculty and departments are synonymous. Due to potential inference problem that might arise due to the inclusion of this college, we will show our analysis are robust after entirely dropping the engineering college from the analysis. We will further do the randomization inference test analysis which is a method that is robust to inference issues related to stratification designs, cluster size and strata size.

up 49 strata which includes 430 cohorts. Non engineering college have on average (approximately) 5 strata per college. Engineering college make up 10 strata, with each strata having 4 cohorts, which adds up to 40 cohorts. We have in total 470 cohorts divided into 59 strata.

Fig. 1.10.1 displays an example of stratification in a college that has both arts and science faculties. The figure shows how the treatments were randomized within arts and science departments for every degree year of study.

In our design, we have 49 strata with between 8-9 cohorts in each stratum, on average. Each of the 49 strata will have 8 or 9 cohorts to assign to treatments. The remaining 10 strata will have 4 cohorts to assign into treatments in each of the strata. To think about variation in treatment assignment that determines the power of the experiment, one needs to think about how many possible ways treatments can be assigned (Fisher (1937)). For 49 strata, each stratum had on average about 8 cohorts to randomize into 3 treatments and 1 control group. Within each stratum, one has to compute the possible combinations of treatment assignment, where each group will consist of two cohorts. This is 2520 for each stratum. So, in total 49 strata, there are 123480 possible ways of assigning treatments. For each of the remaining 10 strata, there are only 4 cohorts to be assigned to 4 treatments. This amounts to additional 240 possible ways of assigning treatments. In total, the variations in the treatment variable are 123,720.

Ideally, to increase power, one would like to stratify the sample on the basis of variables that predict test scores such as baseline test scores, the proportion of poor students, etc. Due to unavailability of data at the design stage, no pre-treatment analysis was possible. Moreover, the spillover effects on the non-incentivised students found in the experiment could not be predicted beforehand. This is because the literature on the theories of tournaments does not make any predictions on the non-contestants of the tournament. More importantly, we could not predict that the proportion of poor students whose annual family income would be USD 2,340 and below would be as low as about 10 percent on average among cohorts where administrative data on income were collected by the colleges before the intervention. Therefore any pre-treatment analysis of spillover effects was not possible.

We have 106 cohorts in each treatment arm and 152 cohorts in the control groups. The treatment group with untargeted financial incentives is comprised of 3,298 students, the group with a caste-based treatment has 3,369 students, and the group with income-based treatment has 3,114 students. The control group has 4,409 students.

Our randomization implies that students studying in the same university, in the same college, in the same faculty and in the same year of their degree are compared. This further implies that sociology major students in their freshman year studying in college C, affiliated to university U, and treated with incentives are compared to history major students in their freshman year studying in college C, affiliated to university U, who are in the control group, because both fall under the faculty of humanities. But our randomization allows for the possibility that if a sociology major student in the freshman year in college C affiliated to university U is treated with incentives, it may happen that a sociology major student in freshman year in college K affiliated to the same university U is in the control group. Table A1 provides a balance check as to whether subjects jointly predict treatments. There are 47 subjects in total.[3].

Since colleges affiliated to university U are different in terms of management practices, quality of faculty, and quality of admitted students, it was not possible to compare economic major students studying in college C with economics major students studying in college K, even if both colleges are affiliated to university U. Due to a lack of credible measures to quantify the differences among colleges and due to the availability of few colleges, it was not possible to group colleges in the stratification along characteristics, and then stratify by subjects within the pooled colleges.

An alternative method of randomization is to perform a comparison between the students studying the same major, who are in the same year of their degree and at the same college. However, this would imply randomization among students studying at the same college and taking the same major exams. This would cause spillovers in study partnerships, violating our identifying assumptions to measure the spillover effects of incentives.

Our stratification can have problems of identification. To deal with the problem

[3]One concern about our randomization is that we are comparing students who have selected to major in different subjects. Two factors are important to note: 1) admissions into a specific degree course is not entirely driven by the choice of the candidate. Admission is dependent on the candidate's school leaving exam scores and the cut-off scores required by a college department. If we observe a student studying physics, it does not imply that this subject was her first choice; and 2) during the pre-college education years (e.g. high school studies) students may choose to study arts, science and commerce courses. Students who select arts and commerce courses in their pre-college years cannot then study any science or engineering subjects at college. Arts and commerce students are restricted to choices in arts subjects and their final degree partially depends on their choice of subject and the ongoing cut-off scores of the colleges. Thus, stratifying candidates between arts and science and engineering degrees is likely to induce a balance along the unobservable that is correlated with subject preferences.

of identification, in this section we provide balance tests of key observables. Later, while describing the estimation strategy we will explain how our design solves key issues such as teacher bias in grading test scores and grading on a curve problem. Table 1.11.1 provides a balance test of covariates, that includes pre-intervention university test scores, of the full sample of 11 colleges using regression specification with treatments and stratifying dummies (Duflo et al. (2007)). The standard deviation of each of the covariates are also specified for each of the treatment and control groups. Additionally, we report in Table B1 a placebo check where we examine treatment effects on university test scores *before our intervention*. Since freshman year students did not take any university exams by then, we only report pre-intervention university test scores of the students of higher years (such as second year students, third year students, so on). We find no effect on the pre-intervention university test scores of the students. In column (2), we report treatment effects on the university test scores of *the same students after our intervention*. We find a negative treatment effect of income targeting on post-intervention test scores. Furthermore, we use randomisation inference test to solve the issues of identification. Randomisation inference is robust to outliers (Young (2015)). Therefore, any imbalance between the treated and control groups would not bias the results.

The way we stratified our sample could also lead to problems of statistical inference. We use both conventional econometrics and randomization inference methods to establish our results.

Assuming 470 clusters are sufficiently many, conventional econometrics suggests that clustered treatment assignment can be accounted for using cluster robust standard errors (Cameron and Miller (2015)). Similarly, the problems that stratification can pose for estimation can be solved using strata fixed effects (Bruhn and McKenzie (2009); Duflo et al. (2007)). In the robustness check, we try to address potential issues regarding inter cluster spatial correlations.

However, to remove all concerns about the correct specification of standard errors that would lead to valid inference, we use Fisher's randomization inference test (Fisher (1937)) to estimate the treatment effect of incentives. Randomization Inference method tests the sharp null of no treatment effect i.e. the test which examines whether the treatment was completely irrelevant.[4] Fisherian randomized inference

[4]The sharp null hypothesis is particularly relevant for estimating the effect on the test scores of the non incentivized students who were exposed to their peers receiving the incentives i.e. the spillover effect. The non incentivized peers can be expected to have no effect on their test scores. The sharp null of no effect on test scores is most relevant in this case.

test is exact with a known distribution regardless of the sample size or the covariance structure of the errors. It does not rely on the asymptotic properties of the estimators. Problems with inference due to non standard techniques of stratification or varying cluster size or strata size or outliers can be solved using randomisation inference test (Young (2015)). In conventional econometric methods, the statistical distribution of the estimated treatment effects is assumed to be emerging from the random draw of the disturbance and the regressor terms from a population distribution. In contrast, in randomization inference the key thought experiment is that, given the experimental sample, the only random element determining the realization of outcomes is the randomized allocation of treatment. Indeed in RCTs, the researcher randomizes the treatment which determines , if any , the differences in outcomes. In RCTs, the researcher knows exactly how the randomization was carried out, and randomization inference uses this information to assess whether observed outcomes in a given sample are likely to have been observed by chance even if treatment had no effect. To compute randomization inference p-values, we only need to replicate the original random assignment method and determine the distribution of the test statistic under the null hypothesis, through resampling. We report p values of treatment effect coefficient of two randomization inference test: 1) randomization inference-c and 2) randomization inference-t using Alwyn Young's randcmd code (Young (2015)). Any randomization test is exact. The choice of test statistic is based on power rather than on sample size. We will report results using both methods for our key findings.

1.2.2 Treatment

The financial incentives were randomized across three-treatment arms. The first treatment arm received untargeted incentives, in which all students studying the same major, in the same year of their degree program, and at the same college were considered for financial rewards. The second treatment arm received socially-targeted incentives, for which only lower-caste students among a cohort of students that study the same major, are in the same year of their degree, and are at the same college were eligible for financial rewards. The third treatment arm received economically-targeted financial incentives, in that students from households with an average annual income of 150,000 rupees (approximately equivalent to USD 2,340 in PPP terms) or less among a cohort of students studying the same major, in the same year of their degree program, and at the same college were eligible for monetary prizes. The control group consisted of the entire cohort of students studying the same major, in the same year of their degree program, and at the same college who received no financial incentives.

There were two types of financial incentives in each of the treatment arms: 1) one prize was for the student who scored the highest in the major subject among all the classmates within the cohort; and 2) the second prize was for the student who was in the top 40% of the cohort and had shown the highest improvement in scores in the major subject, compared to their previous university exam scores in the same subject. These two types of monetary prizes were administered to all treatment groups. Under the lower-caste-based incentives, the student who scored the highest among the lower castes within the cohort and the student in the top 40% percent among the lower castes within the cohort who had shown the highest improvement in scores compared to previous university exams were awarded prize money. Similarly, under the income-based incentives, students from low-income backgrounds, defined by an annual family income of USD 2,340 or less, competed only amongst themselves. The highest scoring student among these poor students was awarded the prize, along with the student who had shown the highest improvement in test scores and came in the top 40% among the poor students in the cohort. Financial incentives were only given to the major students according to their performance in major exams.

Students were automatically eligible for the incentives if they were enroled in college. They did not have to apply for the incentives. All treatment groups and control groups contained students from freshman years and from higher years. Students from freshman years only had one prize per cohort because there were no baseline university exams available for them to award the second incentive prize. Students from higher years had two prizes per cohort.

Using the baseline survey, financial incentives in each city were awarded according to one month's rent for the students in the private sector. In the colleges in Assam, the financial incentive was about USD 62 for the highest scorer and USD 54 for the student showing the highest improvement in performance. In Meghalaya, the financial incentive was 5,500 rupees (USD 84) for the highest scoring student and 5,000 rupees (USD 78) for the student that showed the highest improvement in performance. This is because the cost of living in Meghalaya is higher. Nine out of eleven colleges are from Assam and only two colleges are from Meghalaya. The financial incentive given approximates to about 6-8% of state GDP per capita. The colleges in the sample, on average, offer merit and merit-cum-means scholarships of USD 11 annually for every department, such as the department of mathematics, department of economics, etc. The monthly rent for college accommodation averages out to approximately USD 23.

Both caste- and income-based incentives were offered using administrative records collected by the college at the time of student admission. Due to affirmative action policies, all public colleges maintain detailed caste records of their students. Colleges maintain income records of those students whose parents' annual incomes are 300,000 rupees (equivalent to USD 4,680) or less. However, the income records of second- and third-year students were missing for ten colleges because they did not maintain income records in previous years. Since colleges are not legally required to maintain records of income, most colleges do not maintain student's income records. However, in 2016, the government of Assam urged colleges to maintain detailed income records of poor students who are eligible for educational subsidies by the government. Therefore, we obtained income records of freshman-year students from most of the colleges. In severely resource-constrained colleges, most of these records have been lost. In such colleges, the students were asked to submit government-issued income certificates to the heads of their departments. However, this would potentially cause endogeneity in this study's treatment. Consequently, to examine the spillover effects of income tournaments on test scores, we only examined the sample for which administrative records were maintained before our intervention.

In each cohort, students who were eligible for prize incentives are called contestants, and students who were ineligible but their colleagues within the cohort were eligible are non-contestants who were exposed to the tournament treatment. We estimated the average effect of the treatment by comparing the average means of the test scores of the treated and the control cohorts. To estimate the effects on groups that were directly incentivized, we compared the mean performance of the incentivized groups in the treated and control cohorts. To estimate the spillover, we compared the mean performance of the non-contestants in the treated cohort that were exposed to the treatment with the non-contestants in the control cohort.

1.2.3 Timing

Fig. 1.10.2 summarizes the timeline of the experiment. In 2016, between July and mid-September, a baseline survey was conducted in eight colleges across three cities. The baseline survey consisted of data pertaining to the parent's occupations, preferences for quota policies, student friendship networks, and information on student living expenses and monthly scholarships. For logistical reasons, the baseline survey of the remaining three colleges could not be completed by mid-September. From the end of September to the first week of October, students were informed about the monetary rewards being given out. The information was provided by each college principal using a college notification and was also announced in class by the respective heads of departments. Students were also told that limited funding was available and that the funding was allocated to departments using a lottery. Additionally, we went to all the treatment and control groups and collected information, using surveys, regarding their attitudes towards targeted funding and their perspectives on winning untargeted and targeted monetary rewards.

The colleges remained closed from the second week of October until the first week of November. The exams took place between mid-November to the first week of December across eleven colleges from four universities. Another post-intervention network survey was administered in january in one college. The exam results were published between the end of December and the second week of March. The monetary rewards were administered at each of the colleges when the exam results were given to the colleges by the respective universities.

1.3 Data

1.3.1 Administrative data

GPA

We used administrative data on university exam scores from the university database as our primary outcome of interest. College degrees are graded based on performance in the high-stakes university exams, and the test scores of university exams are used to apply for both private and public-sector jobs after graduation. These test scores are also required for entry into master's programs after graduation.

The semester exams taken by the students determine the GPA scores of the students, which range from 0 to 10. All students from each college in the sample take one university exam each semester. Each university provides an aggregate GPA score for the major subjects of each student, on a scale of 0 to 10. Two universities did not provide GPAs on a scale of 0 to 10. Thus, we changed the score to a percentage score and scaled it by 10.

The baseline scores of the students (i.e. the GPA of the students in previous university exams taken before the intervention) were only available for those not in their freshman year of study. Additionally, the baseline data from two colleges both affiliated with the same university could not be retrieved due to administrative changes in the post-experimental period. The total number of students for whom baseline scores were available is 6,643. The mean baseline scores were 6.05. Both the proportion of students for whom baseline data was not available and the baseline scores of students for whom the baseline test scores data were available were balanced along the observables across the treatment and control groups, as reported in Table 1.11.1. The pairwise tests across treatment groups are shown in Table A2.

The dropout rate of these high-stake university exams is about 8%, and these students will need to repeat all their exams the following semester. The control group has 397 of these students, the income-targeting treatment group has 259, the caste-targeted incentive group has 292, and the untargeted incentives group contains 244. A randomization check is shown in Table 1.11.1 and Table A2, which highlights that the treatment groups are neither pairwise nor jointly different from the control group regarding dropout rates. Baseline data for students only in the higher years of study is available. In the attrition sample, approximately 427 students are not in their freshman year, from which only the baseline records of 235 students are available. The mean of the baseline scores of the attrition students is 3.89, and the scores range from 0 to 9. About 60% of students in the attrition sample belong to lower castes; however, the attrition rates of both the lower and upper castes are balanced across the treatment arms.

Data on Caste and Income

Every college maintains records on student's castes using government-issued caste certificates at the point of student admission. We used administrative data to identify caste. Data on castes was available for 11,226 out of 14,190 students using college administrative records. Students whose caste data was not available were treated as upper caste.

In contrast, data regarding income was only collected by a subsample of colleges. Data on income was available for 6,542 students using administrative data. Poor students who received prize money were identified using administrative data.

In colleges where income records were not maintained, the head teachers of each department were asked to collect the income records of their students. Very few records were collected by the head teachers. Some students directly submitted their records to the research team. Poor students were identified from the income records submitted either to the research team or to the head teachers before the exams took place.

To understand the spillover effects of income-treated cohorts on non-poor students within the cohort, we only analyzed the subsample of colleges where administrative records were available with the colleges before the study's intervention.

Table 1.11.1 shows the sample balance using only college administrative records. Approximately 50% of the students on average in each cohort are from the lower castes, and around 10% of the students on average in each cohort have an annual family income of USD 2,340 dollars or less; thus, they are eligible for economic targeting. Fig. 1.10.3 shows the distribution of baseline GPAs for upper and lower castes, controlling for income. In our sample, the poor lower castes are .5 standard deviation below poor upper castes, which is in contrast to the educational attainment gap between black and white students, where black students score roughly 1 standard deviation below white students on standardized tests.

1.3.2 Baseline Survey Data

A survey was administered in eight colleges before the intervention. The survey collected information on whether the students received any scholarships and the exact amounts they received in scholarships. The students were also asked to provide the full names of three of their friends with whom they hang out and study after lectures. For each named friend, the students self-reported the frequency of their interactions. The question regarding the frequency of interaction was posed as a scale ranging from seven days a week, two-to-six days a week, and one day a week. Using a Likert scale, the survey also asked questions about how often they agreed or disagreed with caste- or income-based quotas or no quota policies. The survey was conducted towards the end of their lectures, with the help of the lecturer. The survey was conducted for those present in class, among which the response rate was 100%.

It could be that the students who responded to the survey were different from the entire sample. Therefore, to use the survey data to examine the possible mechanisms through which incentives work, we first needed to examine how different the respondents were from the rest of the population and whether the respondents differed across the various treatment arms. Data from the survey was merged with the student university exam mark sheet using the full names of the students, their department, the year of their degree program, and the class roll list. We were able to identify 5,905 students (45% of the sample) out of approximately 9,000 survey responses using administrative records. Students that have an identical full name or study at the same college, in the same department, and in the same year could not be uniquely traced back to their exam mark sheet. Thus, these students were dropped from the sample (there were only two cases). Most of the time, the handwriting could not be read properly to be able to trace the names back to the student's mark sheets. Table A3 compares the statistics of students who took part in the survey and whose names could be traced back to their mark sheet with students who either did not do the survey or could not be linked to their exam mark sheet. Table A4 and Table A5 show the balance along the observables of the students across the treatment arms of the survey respondents.

1.3.3 Baseline Network Data

Students first reported the friend with whom they interact most frequently after lectures. The third-reported friend has the weakest ties to the survey respondent. No survey respondent reported meeting their third friend every day after lectures. The type of network (e.g. a mostly upper-caste friendship network or a network with all high-achieving friends) signals both the resources available to the student as well as the type of student they are. Our current analysis focuses on the distributional impact of incentives on the different types of networks defined along the lines of caste and ability, where ability is measured using baseline university test scores. A network can also be defined in terms of strength, measured by the probability that a relationship is reciprocal. We analyze the impact of reciprocal networks. Reciprocity in friendship predicted positive effects on test scores.

Out of 5,905 surveys (about 45% of the full sample), data on the caste of 4,873 students (about 33% of the full sample) was successfully identified using college administrative data. Students may organize themselves into friendship networks by caste and ability. Some relevant facts about the baseline scores of the student study networks in our data are as follows: lower-caste students who have all upper-caste friends have a much higher-than-median GPA (6.82), compared to upper-caste students who have all upper-caste friends; they have, on average, a GPA score of 6.5. This shows that lower caste students who organize themselves into upper-caste friendship networks are, on average, of higher academic ability. Fig. 1.10.4 shows that students with an above average baseline GPA have friends with higher GPAs than friends of students with lower than average baseline GPAs. There is strong evidence of assortative matching by test scores. Since lower-caste poor students have, on average, very low test scores, they are likely to have low-ability close friends.

1.3.4 Post-Intervention Survey Data

A post-intervention survey was administered at all 11 colleges, which asked the students to rank the targeted and untargeted incentives given out by the researcher. Additionally, the students were also asked the following question: if they could give out prize incentives in colleges at that time, how would they allocate USD 11 between three prize winners: 1) the winner who scores the highest in the cohort, 2) the winner who scores the highest among the lower caste within the cohort, or 3) the winner who scores the highest among the poor students whose family income is USD 2,340 or less. Students were also asked, hypothetically, to mention their perception of the chances of winning the prize money under untargeted, income- or caste-targeted schemes. The students were encouraged to respond and express their opinions about the prize incentives given out. The response rate in this survey was nearly 100% among those who were present in class. This survey was self-administered and conducted in class towards the end of lectures. Thus, the responses of those students that were not present in class when the survey was conducted are not available. In neither of these cases were the students told beforehand about the survey that would be conducted during class.

After the university exams had been completed, a second survey was conducted in only one college (due to logistical reasons) to record post-intervention friendship networks. The respondents were asked to give the name of the student with whom they had studied the most for the incentivized exam. They were also asked how long they had known their study mate. This survey was conducted online, and the students were not told that it was linked to research associated with the prize money for exams. Drawing on Gächter et al. (2015), questions were asked to measure the degree to which students had bonded with their study mates. These questions aimed to test whether the targeted incentives caused competition among friends and friction between peers who had not been incentivized.

1.4 Model

Next, we present a simple extension of the basic tournament model of Lazear and Rosen (1981) in order to guide our interpretation of the empirical analysis. Given our setting of a team-based environment, where students rely on each other to perform well in exams, and where teacher absenteeism is high and the quality of teaching is poor, we abstract away from institutional channels that may impact on student achievement in order to focus on the two margins through which tournament incentives impact on performance: (i) by motivating agents to exert effort, and (ii) by changing peer group interaction.

Tournament incentives are shown to be very powerful at incentivizing performance. Experimental research shows that tournament incentives are even more powerful than predicted by the rational decision-making model (Dechenaux et al. (2015)). Research shows that people value winning in itself and enjoy the recognition afforded by relative rankings. A simple laboratory experiment shows that more than 40% of individuals are willing to exert costly effort in order to win a tournament in which the winning prize was 0 USD (Cason et al. (2010)). Several field experiments, such as Kosfeld et al. (2014), have shown that, by simply honoring the best performance publicly with a symbolic award, a manager may increase the average performance of individuals. Thus, targeted tournament incentives that may motivate contestants could also potentially demotivate colleagues who are excluded from the competition.

Although tournaments can create powerful competitive incentives, there are some disincentive effects of tournaments. Several experimental studies have shown that tournament incentives induce non-cooperative behavior among contestants (Drago and Garvey (1998)). In our experiment, we additionally explore the possibility of non-cooperation between contestants and non-contestants.

1.4.1 The Basic Building Blocks

For simplicity, we consider the two-player tournament model, where the student with the highest test score receives a fixed prize, i.e. V_j for student j. Since the exams considered here are high-stakes university exams, which are a crucial factor in the job market outcomes of students, we introduce a parameter M_j that captures the utility from the scores obtained in university exams. Additionally, we introduce the possibility of an encouragement or discouragement effect for being excluded from participating in tournaments, which may vary according to the demographic groups that are able to participate in the tournament. This we denote as T_j^t for student j. T_j^t for student j can be interpreted as how a student feels about obtaining high test scores in university exams when he will not be recognized for his achievements but his colleagues will be given the recognition.

The test score of a student is a random variable whose distribution is controlled by the student's effort to prepare for exams. However, a given realization of a test score also depends on random factors that are beyond the student's control.

Therefore, a student j has test score q_j, which follows Prendergast (1999)

$$q_j = \mu_j + a_j + \epsilon_j \tag{1.1}$$

Where, μ_j is the level of effort chosen by student j prior to the realization of the random luck component ϵ_j that has density g and cdf G. Each agent differs in ability, denoted by a_j for student j. Higher ability increases the chances of success for any given level of effort. Average skill is produced at a cost denoted by $C(.)$ that is increasing in effort μ.

We assume that $G(.)$ has a symmetric distribution; that is,

$$G(\epsilon) = 1 - G(-\epsilon)$$

for all ϵ. It follows that the density function, $g(.)$, then satisfies $g(\epsilon) = g(-\epsilon)$. We also assume that g is unimodal[5] and satisfies $\epsilon g'(\epsilon) \leq 0$. The student's education production function follows (1), and the winner of the prize is determined by the largest drawing of q. Each player plays against the "field". Each player does not observe the precise number of hours that a peer engages in study for the purpose of the exam.

We assume that any student can reach out to his/her colleagues to help him or her prepare for exams. Additionally, each student can have closer ties with a subgroup of their colleagues with whom they call friends. Group study with colleagues is assumed to reduce the cost of effort for the student in preparing for exams. To illustrate this assumption, consider a student who has not understood a concept taught in class. The student can read books to understand the concept or approach a teacher, who might not have enough time available to answer all the student's questions. Alternatively, the student can approach a student who is relatively good at their studies to explain the study material. This can save the student time in looking for the appropriate reference books or waiting for the teacher's availability to help with the required studies. Formally, it is written as follows:

$$C(\mu_j, S_j(f, c)) \tag{1.2}$$

[5] The assumption of a unimodal distribution is usual in tournament models, see, for example, Dixit (1987), Drago et al. (1996), Hvide (2002), Chen (2003)

where μ_j is the effort of the student j, $S_j(f, c)$ is the availability of study resources from classmates who are friends f, and from peers who are just classmates (denoted as c) of the student j. $C(.)$ is a convex function of effort μ_j. S is an increasing function of f and c. $C_{\mu,s}$ is a decreasing function, which implies that a higher exchange of resources among peers reduces the cost of individual effort on exams. S measures how resourceful your friends are. We do not endogenize $S(.)$. For simplicity, it is assumed to be an environment that can be affected by tournaments. For example, in the case of untargeted tournaments, the high-ability students might wish to spend more time studying on their own and might not wish to spend much of their time helping low-ability students, should they seek support in their studies. In the case of targeted tournaments, colleagues might find targeting unfair and refuse to offer help to the beneficiaries, or the beneficiaries might hesitate to ask the excluded group for their time. We denote $S_j(f, c)^C$ as the student interaction available to student j in the control group, $S_j(f, c)^U$ as the student interaction available to student j under the untargeted tournament conditions, i.e. pure merit tournament, and $S_j(f, c)^T$ as student interaction available to student j in the targeted tournament.[6]

1.4.2 Payoffs

Formally, each student maximizes the following preference function:

$$[P_j V_j]D + M_j q_j + (T_j^t * B)Dq_j - C(\mu_j, S_j(f, c)) \tag{1.4}$$

where D is an indicator variable that takes the value 1 if the cohort is assigned to a tournament, otherwise it takes the value 0. B is an indicator variable that takes the value 1 if the student is assigned as a non-beneficiary, otherwise it takes the value 0. Since the university test score is a high stakes exam, the student's expected utility has a motivation factor $M_j q_j$, regardless of incentives. V_j is the value of the monetary award to student j, where P_j is the probability of winning the tournament. The probability that any student j wins the tournament while competing with student $-j$ is:

[6] To ensure the existence of pure strategy equilibria in the tournament, we assume that,

$$sup_{\Delta\mu,\Delta_a} V|g(\Delta\mu + \Delta a)| < inf_{\mu>0} c''(\mu, S) \tag{1.3}$$

This implies that the cost curve needs to sufficiently convex

$$P_j = Prob(q_j > q_{-j}) = Prob(\mu_j - \mu_{-j} + a_j - a_{-j} > \epsilon_j - \epsilon_i) \tag{1.5}$$

where, $\epsilon_j - \epsilon_i \equiv \gamma$ and $\gamma \sim$g, where E(γ)=0 and E(γ^2)= $2\sigma^2$

It is important to note that this model is the same as the standard tournament model with asymmetric costs, with only one additional modification. In our theory, we assume that tournament incentives not only change the benefit from scoring high in exams but can also potentially change the cost of effort of the student. The way that cost of effort may change because of competitive incentives depends on whether academic support from peer's changes as a result of competition. We allow for the possibility that changes in peer interaction will differ with the income and caste of the student.

1.4.3 Equilibrium

For simplicity, we assume $V_j = V, \forall j$; $M_j = M$ $\forall j$ and $T_j^t = T^t$ $\forall j$. Insight from the model will remain unchanged if we do not impose the above restrictions and assume a representative agent from each socio-economic group.

The equilibrium characterized here is from the perspective of any student j in each of the treatment and control groups.

No Incentives: Control Group

Assuming that an interior equilibrium exists, the equilibrium level of effort is characterized by the following first order condition:

$$M = C'(\mu_j, S_j(f,c)^C) \tag{1.6}$$

Equilibrium effort is

$$\mu_j^* = C'^{-1}(M; S_j(f,c)^C) \tag{1.7}$$

Since C(.) is a convex function, $C'^{-1}(.)$ is monotonic. Here, there are two effects that drive effort in the control group. The first is the motivation to perform well on exams, and the second is the social network effect that reduces the cost of exerting effort in exams. Higher M and S induce students to exert more effort.

Untargeted Tournaments

Student j assigned to an untargeted tournament exerts an effort that satisfies the following best response function:

$$M + g(\mu_j - \mu_{-j} + a_j - a_{-j})V = C'(\mu_j, S_j(f,c)^U) \tag{1.8}$$

and student $-j$ competing with student j under untargeted tournaments exerts effort as follows:

$$M + g(\mu_j - \mu_{-j} + a_j - a_{-j})V = C'(\mu_{-j}, S_{-j}(f,c)^U) \tag{1.9}$$

using symmetry assumption of G(.) , we get $g(\mu_j - \mu_{-j} + a_j - a_{-j}) = g(-\mu_j + \mu_{-j} - a_j + a_{-j})$

The tournament contestants have three effects that determine the equilibrium effort of any student j. The first term on L.H.S of the F.O.C is the motivation of any student j to do well on high stakes exams. The second term on the L.H.S captures the competition effect of the standard theory of contests, which predicts a fall in effort as differences in ability among contestants rise. The R.H.S is the effort and the network effect on cost of effort. If all students have access to everyone in class, then costs across students within a class will differ only if student's efforts differ. If students do not have access to everyone in class but just to their close ties, such as friends, then students who have access to more resourceful friends will have a lower cost of exerting effort in exams.

Income Targeting

We have two forms of targeting in our framework: economic and social. Under the targeted incentives, student j can be a beneficiary or a non-beneficiary. We characterize the equilibrium for any student j in each of these cases.

Under economic targeting, any student j, competing with student $-j$, chooses effort μ_j that satisfies the following:

$$M + g(\mu_j - \mu_{-j} + a_j - a_{-j})V = C'(\mu_j, S_j(f,c)^{income}) \tag{1.10}$$

In addition, her competitor, say student $-j$, chooses effort μ_{-j} using the following best response function:

$$M + g(\mu_j - \mu_{-j} + a_j - a_{-j})V = C'(\mu_{-j}, S_{-j}(f,c)^{income}) \tag{1.11}$$

using symmetry assumption of G(.) , we get $g(\mu_j - \mu_{-j} + a_j - a_{-j}) = g(-\mu_j + \mu_{-j} - a_j + a_{-j})$

If $S_j(f,c)^{income} = S_{-j}(f,c)^{income}$, then the equilibrium effort will be symmetric and will only be a function of the level of competition $a_j - a_{-j}$. Greater differences in ability will induce low-ability students to exert low effort. The best response of high-ability students will then be to exert low effort in equilibrium. In the case of asymmetric costs, $S_j(f,c)^{income} \neq S_{-j}(f,c)^{income}$. The student that has a cost advantage will exert higher effort. In our set up, our treatment creates variation in cost advantages among the contestants by changing co-operation among colleagues that study for the same exam. Therefore, the above best response functions will provide predictions on the heterogeneities in outcomes among contestants of the tournaments.

Under economic targeting, if student j is a non-beneficiary, the equilibrium effort exerted by the student j is given by:

$$M + T = C'(\mu_j, S_j(f,c)^{income}) \tag{1.12}$$

The equilibrium effort reduces to the following first order condition:

$$\mu_j^* = C'^{-1}(M; T^{income}; S_j(f,c)^{income}) \tag{1.13}$$

The equilibrium effort of students excluded from the tournaments will only be affected if their motivation to do well in exams is affected $T^{income} \neq 0$ or if their quality of academic support is affected $S_j(f,c)^{income} \neq S_j(f,c)^C$.

Caste Targeting

Similarly, under caste targeting, we have the following characterization:

Any student j, competing with student $-j$, chooses effort μ_j that satisfies the following:

$$M + g(\mu_j - \mu_{-j} + a_j - a_{-j})V = C'(\mu_j, S_j(f,c)^{caste}) \tag{1.14}$$

In addition, her competitor, say student $-j$, chooses effort μ_{-j} to satisfy the following best response function:

$$M + g(\mu_j - \mu_{-j} + a_j - a_{-j})V = C'(\mu_{-j}, S_{-j}(f,c)^{caste}) \tag{1.15}$$

using the symmetry assumption of G(.) , we get $g(\mu_j - \mu_{-j} + a_j - a_{-j}) = g(-\mu_j + \mu_{-j} - a_j + a_{-j})$

Under caste targeting, if student j is a non-beneficiary, the equilibrium effort exerted by the student j is given by:

$$M + T^{caste} = C'(\mu_j, S_j(f,c)^{caste}) \tag{1.16}$$

The equilibrium effort is as follows:

$$\mu_j^* = C'^{-1}(M; T^{caste}; S_j(f,c)^{caste}) \tag{1.17}$$

Here again, the equilibrium effort of the non-beneficiary is affected only if the targeted policies change his motivation to score high in exams or his academic support system.

To better understand the motivation and peer effect channel of our model, consider the case where $S_j(f,c)^p = S_j(f,c)^C = S_j(f,c)^U$, where $p = \{caste, income\}$ $\forall j$ and $T_j^t = 0$. In this case, both untargeted and targeted incentive schemes will unambiguously increase average effort relative to the control group, where no incentives were given. Changes in effort relative to the control group for both contestants and non-contestants occur if tournaments change the motivation to perform well in addition to the academic support from peers required to perform well.

There is a third very important parameter that has not been incorporated in the model, which is the size of the tournament contest. Since our experiment varies competition types, both in relation to the size of the contest and the socio-economic features of the contestants, the interaction term of the proportion of the contest size and social groups will be the relevant parameter to test. For simplicity, we keep the size of the contest constant in the model to first understand the way the motivation channel and social network channel operate to impact on student achievement. Then, we infer how the proportion of incentivized students could impact on outcomes, e.g. the proportion of incentivized students will indicate the proportion of students who are being excluded so that if the excluded students are demotivated for being excluded from the opportunity and reduce effort in exams, the average scores of the cohort will be affected only if the disincentive effects are large enough.

Proposition 1 *If for the non-beneficiary student j, $S_j(f,c)^p = S_j(f,c)^C$, where $p = \{caste, income\}$, then the equilibrium effort of the non-beneficiaries in targeted incentives will fall relative to the control group if $T_j^t < 0$.*

If non-beneficiaries find targeted incentives unfair and feel demotivated, then keeping the quality of peer interaction unchanged relative to the control group, optimal effort exerted by the non-beneficiaries, in equilibrium will be lower than that in the control group.

The above proposition highlights the importance of the motivation channel in explaining the potential negative spillover effects of targeting. Since the experiment is about targeting students from poor social groups, the non-beneficiaries of the targeted incentives are economically better off by construction. The advantaged socio-economic groups, on average, have higher pre-test exam scores. Therefore, if targeting causes a rift between low ability beneficiaries and high ability non-beneficiaries, the high ability non-beneficiaries are less likely to experience negative effects on the quality of study support. Therefore, the above proposition that keeps the quality of the study group of the non-beneficiaries unchanged is realistic.

Proposition 2 *If $S_j(f,c)^p < S_j(f,c)^C$,where $p = \{caste, income\}$, then the effort of the beneficiaries will fall relative to the control group if the cost of effort due to changing peer interaction is higher than the gain from exerting effort in the contest.*

Targeted tournaments incentivize the beneficiaries to exert effort in exams. However, targeted incentives can have a negative externality on test scores by creating a tension between the beneficiaries and the excluded peer group. Given that effort is complementary, if the excluded peer group does not co-operate with the beneficiaries, such that the quality of academic support from peers falls compared to the control group, then the cost of effort to increase test scores will rise for the beneficiaries. The targeted tournaments can increase the equilibrium effort exerted by the beneficiaries relative to the control group only if the gain from the incentives outweighs the cost of exerting effort in exams.

Proposition 3 *If targeted financial incentives disrupt cooperation among classmates but not among friends, denoted as f in S compared to the control group, then, under such schemes, contestants with higher S, i.e. contestants with resourceful friends, will have positive effects on test scores.*

Students with resourceful friendship links will have higher S. This will reduce the cost of effort for the students. Tournaments generate incentive effects for the contestants. However, if tournament incentives disrupt co-operation among colleagues who are distant and not among the subgroup of colleagues who are close friends, then the cost of effort will be lower only for those contestants who have resourceful friendship ties. This will cause effort and hence the test scores of students with productive friendship ties to rise. The test scores of disadvantaged students who do not receive support will fall if the cost of effort outweighs the incentive effects of financial rewards.

1.5 Estimation Strategy

To test whether the provision of tournament incentives affects test scores, we estimate the intent-to-treat effects of financial incentives using the following specification for student i, in department d, in faculty f, and in year t of his studies at college c.

$$y_{idftc} = u_{ftc} + \beta A_{dftc} + \gamma C_{dftc} + \alpha I_{dftc} + \epsilon_{idftc} \tag{1.18}$$

where, y_{idftc} is the outcome variable of student i, in department d, in faculty f, in study year t, and at college c. We express grades as a z score, standardizing them by the full sample mean and full sample standard deviation, so that our coefficient of interest can be interpreted as the standard deviation (SD) change in the university grade point averages associated with the incentives. u_{ftc} refers to the randomized strata dummies over faculty, study year, and college level (Bruhn and McKenzie (2009)), A_{dftc} is the dummy for the treatment of untargeted incentives i.e. incentives for all, C_{dftc} is the dummy for the treatment of caste-targeted incentives, where only lower castes within the cohort compete, and I_{dftc} is the treatment of economic-targeted incentives, where only poor students within the cohort compete. We allow for our errors to be correlated within the cohort by clustering our standard errors ϵ_{idftc} at the cohort level. Barrios et al. (2012) view clustering as a design problem and show that within group correlation is neither a necessary nor sufficient condition for clustering standard errors. They suggest that correlation between the errors and covariates that are not randomly assigned could imply inter-cluster correlation, in which case the standard errors need to take account of this correlation. As a robustness check, we also clustered standard errors at the level of college-faculty-year of study, which reduced the number of clusters from 470 to 106.

The primary outcome of the analysis is the grade point average of the university semester exams. Given that these are high-stakes exams, dropout rates are low. 92 percent of the students took the university exams, and there is no difference in the dropout rates between the treatments and the control.

We want to estimate the effects of incentives on student effort that increase the test scores of the students. Thus, to estimate the effect of student effort on educational attainment, we cannot have teacher biases in grading confounding the impact of the incentives on student achievement. We achieve this by incentivizing university exams that are graded centrally and are not marked by teachers that teach the courses in the colleges. Additionally, the identities of the students in the exams remain anonymous. Given that college syllabuses are also set centrally by the university at the beginning of each academic year, there is no scope for the teachers or college management to change the level of the courses in response to incentives. Additionally, university level grading of approx. 360 colleges that includes more than 50,000 students in total, of which we only have a very small fraction in our sample, along with college fixed effects solves our grading on a curve problem.

Unlike other studies in the literature on performance-based incentives, such as those carried out by Kremer et al. (2009); Angrist et al. (2009); Burgess et al. (2016), we have mitigated the channel via which teachers or college authorities can impact the test scores of students in response to incentives by increasing the quality of teaching. The incentives were announced in every college towards the end of term and a month before the exams. All colleges had about two weeks of lectures left before they all closed for three weeks for the autumn festival (Durga Festival) in the month of October. Exams took place when the colleges re-opened. Additionally, all colleges suffer from a high level of student absenteeism because students prefer to study either on their own or in study groups or take private tuition rather than attending lectures because of the poor quality of teaching in public colleges. The rate of absenteeism increases towards the end of the term and closer to the exams, when students prefer to take study leave. Given the high level of student absenteeism, there was very little scope left for the teachers to increase their quality of teaching and convince students to attend lectures as a response to incentives during the last two weeks of term.

The parameters of interest are β, γ, and α, where the null hypotheses are $\beta = 0$, $\gamma = 0$, and $\alpha = 0$. Our main identifying assumption is that, in the absence of tournament incentives, there would be no differences in the test scores between the students in the treated groups, relative to the control group. Thus, for incentives to change the test scores of the treated cohort without changing the test scores of the untreated cohort, there can be no spillovers or complementarities in terms of study effort between the treated and control cohorts. Given that the treatments were randomized across departments and only among major students that do not share common curriculum across departments, the experiment design has ruled out spillovers in terms of study effort between treated and untreated cohorts.

Nevertheless, the interpretation of β,γ and α as an effect on the incentives of test scores, compared to when no incentives were provided, could be contaminated by the fact that the students in each treatment group could be aware of different incentives provided to other classes. The effect could be interpreted as the response to which incentives are provided to the class the student belongs to versus those offered to students in other classes. However, since students studying major subjects only have major class lectures among themselves, and they spend a greater number of hours in lectures amongst their classmates, the spillover effect of the above type is likely to be very low. Moreover, we believe that students are less likely to be affected by opportunities in other departments, as students studying different majors already have different career tracks and job placements.

1.6 Impact of Incentives

Table 1.12.1 shows that only tournament incentives targeted towards the poor had an impact on the test scores of the cohort that were exposed to the incentives. We find statistically significant negative effects on the test scores of the treated cohort, compared to the cohort that received no incentives. The mean deterioration in grades .14 SDs ($s.e = .053$) of student performance is significant at a 1% level of significance. The significance of the coefficient of interest is robust to college-faculty-year level, which yields $s.e = .0537$.[7] We compute randomisation inference p values for individual coefficients using Young (2015) randcmd code. The randomisation (c) p values for untargeted treatment .54, for caste targeting is .51 . The randomisation (c) p values for income targeting is significant at .013. the randomisation (t) p values for untargeted treatment is .52, for caste targeting is .50 and for income targeting is significant at .02.

This is the first evidence in the literature of performance-based financial incentives, where economic targeting has reduced the average test scores of the treated cohort. In the tournament incentive program for girls in Kenyan primary schools studied by Kremer et al. (2009), average test scores of the girls increased by .19 SD. They also observed significant improvement in teacher attendance in program schools, establishing a plausible behavioral mechanism that explains the test score gains for girls and boys in program schools. In our setting, human capital externalities in response to incentives are more likely to happen via interactions among peers than by behavioral changes of the teacher. Targeted incentives that excluded colleagues from the opportunity to participate in the achievement awards competition could have disrupted academic cooperation among peers, which could potentially explain the negative outcome.

We also find no average effect on test scores when students are exposed to caste-based incentives. We find no effect on the lower castes, who are directly incentivized, and find a small negative and statistically insignificant effect on the upper caste in the treated cohorts, relative to the cohorts that received no incentives. In the experimental economic literature, Hickman (2013) and Calsamiglia et al. (2013) studied incentive effects of affirmative actions by making disadvantaged students compete among themselves. The papers find that such tournaments increased average human capital investment and exam performance for the majority of the disadvantaged students targeted by the policy. In our setting, it had no effect on the disadvantaged social groups, i.e. the lower castes. The lower castes are over-represented (approximately 50% in each cohort) in this study's sample colleges due to affirmative actions

[7] Clustering at the level of college-faculty-year has no effect on the standard errors, thereby ruling out the possible existence of correlation across cohorts. Clustering at the level of college reduces the standard errors, indicating that, with the number of clusters reduced to 11, the estimates could be erroneous.

that are nationally in place in India. Therefore, it is plausible that the competition that includes more than half of the cohort tends to replicate the competition that includes the entire cohort. We do find that the effect of social targeting is the same as that of untargeted incentives. One may argue that lower castes do not respond to incentives due to the caste stigma attached to caste incentives. We will show in a later section that the stigma effect of caste is less likely to explain the absence of any impact of social targeting.

We find no average effect of merit-based incentives. In the performance-based incentive schemes studied by Levitt et al. (2011) and Angrist et al. (2009), they found a positive impact of such schemes on students who are more likely to win the prize, i.e. on marginal students. This caused such programs to have a positive average impact on educational attainment. Tournament incentives are also known to cause "discouragement effects" on academically weaker students (Sheremeta (2015)). Therefore, the literature has found a heterogeneous effect of tournament incentives on students of different abilities. We did not find evidence of such an effect in our quantile treatment effect reported in Table 1.12.2. All quantiles of the score distribution have no effect under both merit and caste targeting and have a negative effect on test scores under income-targeted incentives. Both students in the upper and lower quantiles in the incentivized tests performed substantially worse under income targeting, compared to what the students would have achieved with no incentives. The treatment effect is economically significant. Since poor students have lower baseline test scores on average, the uniform effect of economic targeting across the entire score distribution suggests that even poor eligible students are also negatively affected. We further tested whether heterogeneity in ability that affects the chances of winning contests affects outcomes. The Table 1.12.3 shows that heterogeneity in ability does not impact outcomes under incentives.

The results suggest that the size of the proportion of individuals targeted under affirmative action policies matters. In our setting, targeting not only changed the demographic composition of the contestants but also the number of students competing for the prize. The paper finds a significant negative effect only in the income-targeted contest which, incidentally, also incentivized a small proportion of the students (approx. 10%). In our next section, we examine whether negative impacts were driven by students not eligible for achievement awards.

1.7 Heterogeneous Effect of Income Targeting

Proposition 1 predicts that if targeted tournaments upset non-beneficiaries such that they feel demotivated for being excluded from the opportunity, then they will experience a fall in test scores. Furthermore, if tension among beneficiaries and non-beneficiaries reduces cooperation in academics, the disadvantaged beneficiaries, who are, on average, low-ability students that could potentially benefit from studying with high-ability peers, will also be adversely affected. In our set up, income targeting that excluded, on average, 90% of the cohort from tournament incentives had a significant negative average effect on test scores.

We first use the following specification to estimate the spillover effects of income targeting.

$$y_{idftc} = u_{stc} + \delta I_{dftc} * poor + \tau poor + \alpha I_{dftc} + \epsilon_{idftc} \tag{1.19}$$

where, y_{idftc} is the outcome variable of student i, in department d, in faculty f, in study year t, and at college c. We express grades as a z score, standardizing them by the full sample mean and full sample standard deviation. I_{dftc} is a dummy for cohorts where tournament incentives are provided only to poor students and *poor* is a dummy where the student has a family income of and below USD 2,340. u_{stc} is our randomization stratum. The spillover effect is captured in α, and the differential effect on the poor students due to income-targeted incentives is the coefficient δ.

Using only sample colleges, involving 5,683 students, where administrative data on income was collected from the students at the point of admission and before our intervention, we estimate the spillover effects of targeting the poor. Table 1.12.4 find that income targeting has a significant negative effect on the test scores of students who are excluded from such incentives. Furthermore, we cannot rule out the hypothesis that poor students are not equally adversely affected in terms of test scores. It is plausible that, due to low power, we are unable to detect differential negative effects on the poor. It is also possible that we do not cluster our standard errors correctly. Therefore, to address the issues on inference, we further conduct Fisher's randomisation inference test using Young (2015) randcmd code. We find randomization inference (c) p values for the spillover efects is .038 and randomization inference (t) p value to be .039.

To further show the effect of being demotivated for being excluded from the opportunity to win prize incentives, we estimate the intent-to-treat effects of financial incentives on cohorts where there are no poor students as per pre-intervention administrative records. In such cohorts no student was eligible for income targeted awards. In the subsample, where pre-intervention administrative records are available, there are 204 cohorts, of which 102 cohorts have no poor students and 67 cohorts have poor students higher than 10 per cent. We use only the subsample of 102 cohorts where there are no poor students to identify the resentment effect of announcing income targeted incentives on the test scores of students.

We use the following specification for student i, in department d, in faculty f, and in year t of his studies at college c.

$$y_{idftc} = u_{ftc} + \beta A_{dftc} + \gamma C_{dftc} + \alpha I_{dftc} + \epsilon_{idftc} \tag{1.20}$$

where, y_{idftc} is the outcome variable of student i, in department d, in faculty f, in study year t, and at college c. We express grades as a z score, standardizing them by the full sample mean and full sample standard deviation, so that our coefficient of interest can be interpreted as the standard deviation (SD) change in the university grade point averages associated with the incentives. u_{ftc} refers to the randomized strata dummies over faculty, study year, and college level (Bruhn and McKenzie (2009)), A_{dftc} is the dummy for the treatment of untargeted incentives i.e. incentives for all, C_{dftc} is the dummy for the treatment of caste-targeted incentives, where only lower castes within the cohort compete, and I_{dftc} is the treatment of economic-targeted incentives, where only poor students within the cohort compete. We allow for our errors to be correlated within the cohort by clustering our standard errors ϵ_{idftc} at the cohort level. We find significant negative effect of higher magnitude on the test scores of the students as shown in Table 1.12.5

This is the first paper to provide evidence on the negative spillover effects of targeted tournament incentives. Theoretically, Milgrom and Roberts (1988) and empirically, there is evidence of sabotage and counterproductive behavior among participants that compete in tournaments, but there is no evidence on the effect of peers that are excluded from tournaments. Targeted tournament incentives are a form of affirmative action policies implemented by policymakers to provide the weaker socio-economic groups with a level playing field. Bagde et al. (2016) and Bertrand et al. (2010) examine the impact of affirmative action policies in admissions into engineering colleges in India, using RDD design on marginal students who gain admission into university, and those who fail to gain admission. They found positive effects on the educational attainment and labor market outcomes of the beneficiaries. The average effect on educational outcomes of the non-beneficiaries that get displaced have remained unexplored. This paper provides evidence that affirmative action policies could yield negative outcomes for the non-beneficiaries and possibly the beneficiaries.

Since competition were randomized in every faculty in every colleges, we also evaluate the average effect of income targeting in competitions within every faculty in every college. Fig. 1.10.5 displays the treatment effect coefficients (and its 95% confidence interval) university test scores in each of the 59 tournaments.[8] In the majority of these groups, the income targeting had a negative average effect on student achievement. Interestingly enough, 8 of the groups have experienced statistically significant positive effects. We find groups that have positive effects of income targeting are driven by colleges in poorest district (Hailakandi district). Income records of most of these colleges were not available. While groups that have negative significant effects of income targeting have lower proportion of eligible poor students (.09) with their non poor colleagues having above median baseline scores (7.94). The effect of proportion of poor students eligible for prize incentives and effect of income targeting is reported Table B2 using only the sample of colleges where pre-intervention administrative data was used to identify poor students with family income USD 2,340 and below. This suggests that in targeted incentives, proportion of students excluded from targeting has an effect on student performance. List et al. (2014), Nalebuff and Stiglitz (1983), Loury (1979), show that contest size impacts performance by affecting the chances of winning the prize of the contestants. The novel contribution of our paper is that it shows that tournaments affects peers of contestants who are excluded from the contest and both proportion and the socio-economic characteristics of the excluded peers can potentially explain the effect on the outcomes of beneficiaries and the non-beneficiaries. Our paper cannot disentangle the two channels but can provide suggestive evidence of the importance of the role played by social and economic status of the disadvantaged groups in affecting outcomes by examining the effects of incentives on reported pre-intervention close friendship networks.

1.7.1 Explaining Negative Spillovers

The literature on affirmative action policies in firms and organizations (Harrison et al. (2006); Shteynberg et al. (2011); Lynch (1989) has found a backlash among

[8]There are 59 tournaments because we evaluated the competitions by faculty and year in non-engineering colleges and only by faculty in engineering college.

non-beneficiaries, who may feel unfairly disadvantaged by these policies. We examined whether targeted incentives generated disapproval for the groups receiving prize incentives.

We conducted an attitude survey a week after announcing the tournament incentives, where we asked students to divide USD 11 between the winners of three types of tournaments: untargeted, caste-targeted and income-targeted tournaments. Using the post intervention attitude survey of the sample, which exhibits the same effect of incentives on student test scores as the full sample, as reported in Table B3, we measured whether relative support for the winners of untargeted tournaments, i.e. pure merit tournaments, rises among cohorts where only poor students are incentivized, in comparison to cohorts where everyone is incentivized. Table 1.12.6 shows that support for the winners of pure merit tournaments rises, and Table 1.12.7 shows that support for the winners of income-targeted tournaments falls in cohorts where only the poor are incentivized, relative to cohorts where untargeted incentives are given. The randomization inference-c p values for support for winners of untargeted tournaments is .10 and randomization-t p value for the coefficient of treatment effects of untargeted tournaments is .11. The treatment coefficient of income targeting has p value of .11 under randomization inference-c and .10 under randomization inference-t. One may argue that cohorts where poor students are receiving incentives might express support for merit winners, not because they find targeting unfair but because they think they can now express support for merit winners because poor students are already receiving benefits under our scheme. By this logic, cohorts who received caste incentives should also have supported other types of winners more than lower-caste winners. However, we do not observe this pattern in the data. Table 1.12.8 find that cohorts who receive caste targeting are more likely to support lower-caste winners relative to other winners, in comparison to cohorts where everyone is eligible for the prize incentives. The p values for the coefficient randomization inference c is .07 and for randomization inference-t, it is .05. We also find that lower castes are more likely to support lower-caste winners over others.

The result on attitudes shows that targeting caused cohorts receiving income targeting to seek merit tournaments and reduce support for winners of income-based tournaments. We interpret this as a possible tension between the poor beneficiaries and the non-poor non-beneficiaries, due to targeting. This can cause non-beneficiaries to become demotivated to perform well in exams, or it may cause non-beneficiaries to perform even better in exams to show that it is them who deserve the prize incentives. The negative effect on the test scores suggests exclusion from the prize incentives possibly generated the former effect and demotivated the non-beneficiaries.

The attitude survey is also consistent with a small insignificant negative effect on caste targeting. We find that caste targeting increased support for lower-caste winners, and this was driven by the lower castes. The lower castes are the majority in university cohorts. Therefore, the negative spillover of caste targeting is very small and statistically insignificant.

Another explanation for the negative spillover could be that the spillover effect of being excluded from tournaments may be driven by high-ability students, who want to sort themselves into merit-based tournaments. The literature on tournament incentives shows that more able and less risk averse agents sort themselves into tournament incentives (see, Sheremeta (2015)). Table 1.12.9 cannot rule out no differential negative spillover effects by student ability in the sample, where administrative data on income is available. In the attitude survey, we find that students with higher baseline scores are more likely to show support for merit tournaments. It is plausible that all students excluded from prize incentives were disappointed in not being given the opportunity; however, the demotivating effect could be larger for students who are able and more likely to win prize incentives.

1.7.2 Effect on beneficiaries: *A friend in need*

By Proposition 3, if tournament incentives create tension among colleagues that are distant, without affecting intimate friendship ties within the cohort, then contestants who have high-ability friends should experience a positive impact on test scores. We used the friendship data of a large part of the sample, collected at the baseline, to estimate the causal effects of incentives on the close network types that students had organized themselves into before the study's intervention.

We examined the heterogeneous impact of friendship ties on the test scores of low- and high-ability students under different incentives. We measured the ability of students using baseline GPAs in university exams. High-achieving students are students with higher than median baseline scores (i.e. 6.5). Since students in their freshman year have not yet taken any university exams, we only included data on students from higher degree years.

Table 1.12.10 shows that low-ability students, i.e. students with below average baseline scores, experience, on average, a negative impact on test scores when only poor students are incentivized. However low-ability students that have organized themselves into a high-achieving network experience a positive effect on test scores under income targeting, compared to the low-ability students with no incentives. Here, the results are robust to the number of high-achieving friends that a student reports having. High-ability students who have high-achieving friends do not experience any effects due to income targeting, but they experience a positive effect on test scores under untargeted tournaments. Since a low-ability student is more likely to be poor, the above result can be interpreted as income-targeted incentives benefiting poor students that have organized themselves into resourceful friendship networks before our intervention. High-ability students are likely to be non-poor, and thus are ineligible under income targeting. The results suggest, therefore, that non-poor students who have resourceful friends perform well under merit tournaments. The network type predicts outcomes in response to incentives more than the ability of the students.

If the social learning mechanisms of peer effects are at work (Bursztyn et al. (2014)), the results are in line with the literature on peer effects, which suggest that peer effect is generally asymmetric; for instance, Hoxby (2000) found that low-achieving students benefit from having high-achieving friends. However, it may be that there is no change in peer interaction, and the effect of incentives is reflective of student type rather than the effect of peer support. Manski (1993) defines exogenous effects of peers as the effects driven by a peer's background that could be correlated with the student's background. The student that sorted him/herself into a resourceful network could have a background that predicts his/her outcomes under incentives. In what follows, we illustrate suggestive evidence of changing interactions among colleagues that could potentially explain the adverse outcomes of targeted policies.

1.7.3 Direct Evidence of Changing Peer Interaction

By Proposition 2 and 3, the test scores of the beneficiaries of targeted incentives can be negatively affected if colleagues do not cooperate with them in exam preparation. Using the network data of a post-intervention survey conducted only in one college, which also replicates the same impact on incentives as the aggregate sample reported in Table B4, we provide direct evidence of changing peer interaction. The response rate of the survey was 8%, and therefore we use this only as suggestive evidence of changing peer interactions and peer effects at work.

In the survey, the students were asked to name a study mate that they worked with the most for the incentivized exams. The students had the option of typing "none" if they did not have any study mates. A total of 27 students reported having no study mates in the incentivized exams. The names of the study mates were tracked back to their exam roll numbers using the exam roll list. The caste, income, and baseline GPA of the study mates were recorded using college administrative data.

Table 1.12.11 shows that, when incentivized, students turn to their close friends for support. However, the level of intimacy among study mates is higher and statistically significant in cohorts where poor students are targeted with financial incentives. The level of intimacy among study mates falls from income targeting to pure merit-based incentives. Due to small sample size, clustered robust standard errors is likely to produce erroneous results. We use randomisation inference test to estimate treatment effect coefficients (Young (2015)). We find that the p values for cohorts treated with income based prize incentives are .053 with randomization-c inference test and .044 with randomisation-t test. With incentives for all there is no effect on test scores. The p values for treatment coefficient of untargeted incentives using randomization inference-c method is .461 and randomization inference-t method is .44. The p- values for caste treatment is .74 under both randomisation-c and randomisation-t method. Our results suggest that students, on average, receive support from very close friends to prepare for exams when treated with economic targeting. This effect is reduced as more students are included in the prize tournaments. In untargeted tournaments, when all in the cohort are eligible for incentives, study support is not restricted to close friends.

Table 1.12.12 shows that students under income targeting found, on average, lower-achieving study mates when compared to the control group with no incentives. The p values for the treatment effect coefficient under randomization-c method is .07 for income targeting and no significant effect under caste or untargeted incentive treatments. This suggests that income targeting disrupts the communication between high- and low-achieving students within the class. Consequently, students found cooperation amongst relatively low-achieving students, compared to the control group where no incentives were provided.

Using the pre-intervention network data from the large sample of students, the figure depicts that there is positive assortative sorting in friendship ties. Students with above average baseline scores have, on average, high-ability friends, and students who are low ability, i.e. those with below average baseline scores, have, on average, friends who are low ability.

Combining the evidence on the post-intervention network with the evidence on the large data set of the pre-intervention network, we speculate that negative effects through income targeting come from the segregation caused by targeting and the discouragement effect. Income targeting also produces large negative effects because income targeting in universities excludes almost the entire cohort from the opportunity to win the prize incentives.

1.8 Robustness checks

Barrios et al. (2012) show that spatial correlation between randomly assigned clusters may invalidate the use of variance estimators that incorporate only cluster level outcome correlations. Our within college cluster randomisation may have spatial correlation structure which may impact precision of treatment effect estimates. Therefore, we redo our analysis of the effects of incentives where we cluster the standard errors at college-faculty-year level, which is higher than the level of our cluster, since correlations will be higher within a faculty among the same degree year students studying the same college. The precision of our estimates are robust to clustering standard errors at levels college-faculty-year level as reported in Table 1.12.13. The standard errors change from .0530 to .0537 when standard errors are clustered at college-faculty-year-of-study level which reduces clusters from 470 to 106. This shows that within college correlation between different departments are low and thus clustering at levels higher than the cohort level does not impact statistical significance of the outcomes.[9] Next, we drop the engineering dept from the sample and re do the analysis by clustering our specification at various levels. We show in column (1) of Table 1.12.14 that our results from specification that clusters standard errors at cohort level remain unchanged after dropping engineering college. We, then, also estimate the equation by clustering the standard errors at college-faculty-year level. We find that our clusters reduce from 430 to 63 and our standard errors remain unchanged at .06.

It is possible that our entire results are driven by colleges that are competitive to begin with. In India, the most competitive colleges are the public engineering colleges. Since returns to engineering degree is highest for undergraduates in India, there is a stiff competition to gain admissions into public engineering colleges which are limited in supply and highly reputed. Entry into engineering colleges not only

[9]We also did the analysis with standard errors clustered at the level of college, we find that the standard errors reduced to .048, which indicate the cluster does not work for smaller number of clusters

depend on the scores of the school leaving exams but also the scores obtained in entrance exams administered by the engineering colleges. Most of these exams are administered at national or state level. Column (1) of Table 1.12.14 shows that the removal of engineering colleges the removal of engineering colleges did not even change the magnitude of the effects of income targeting. Another concern could be that the experiment was conducted in two different states: Assam and Meghalaya, where Meghalaya is a minority dominated state which is not representative of the whole of India. It could be that the results are driven by institutions of minority state, which could have implications on the generalizability of the results in Indian context. Column (2) of Table 1.12.14 drops the colleges from Meghalaya and finds that the results are unchanged. Since the freshman year received only one prize incentives per cohort and higher years received two prize incentives per cohort, results from higher years could be different. We do the analysis including only higher years in the sample and find no change in the impact of incentives on test scores as reported in Table 1.12.14.

1.9 Conclusion

This study conducted a field experiment to provide evidence on the effectiveness of targeted and untargeted financial incentives on university test scores in poor economies. We find that economically-targeted incentives, which incentivized the smallest proportion of students, reduced the average test scores of all the cohort, while social targeting that incentivized more than half the cohort and untargeted incentives that incentivized the whole cohort did not affect the test scores. We find that the negative effect of economic targeting was driven by peers who were not given the opportunity to participate in the prize competition while their colleagues were.

As is often the case in field experiments, the interpretation of the findings and their wider applicability depend on the key features of the specific setting. In our case, two features are of note. The first is that we assigned students to each incentive scheme. In general, we expect incentives, such as merit-based or caste-based scholarships, to affect the selection of students, since different schemes will attract different types of agents. Coupled with this, the presence of peer effects might induce disadvantaged students not to apply for income-based stipends. Second, a key feature of our setting is the reliance on friends for study support due to the poor educational infrastructure in most of our sample institutions. The role of peer effects might be smaller or even absent in other settings where there is a strong culture of institutional support for weaker and low-income students.

In weak institutions, our main result has implications for affirmative action policies that exclude large demographic groups who are also resourceful. Evidence of discontent among the advantaged groups and the consequent fall in test scores of both the advantaged and disadvantaged social groups in a team-based environment indicate that policymakers should take the degree of complementarities in input into account while designing policies that cater to weak social groups.

However, the most interesting finding is that the backlash of the advantaged groups did not have any effect on disadvantaged students who had close ties with high-achieving students. We find that low ability students, in general, performed worse under income targeting but that low ability students who sorted themselves into friendship networks with high-achieving students before the intervention benefited from income-targeted incentives. We provide suggestive evidence that this effect is more than just the effect of a student's background helping the student to sort him/herself into a high-achieving friendship network. We provide evidence to indicate that income targeting which disappointed a large proportion of the cohort caused the poor students to find support from their closest friends, who, on average, happened to be of low ability like themselves and could not provide the adequate support required to increase their test scores. However, low ability students who had close friendship ties with high-achieving friends had a positive impact on test scores under income incentives.

The results above speak to the role that institutional support can play to neutralize the effect of the non-cooperation of resourceful social groups on disadvantaged beneficiaries. This result brings forth the salience of the literature on the role of teachers to improve the educational attainment of students from weak socio-economic groups (Dobbie and Fryer Jr (2009); Fryer Jr (2014); Chetty et al. (2011)).

Our experiment bundled together two important parameters that could impact on the student performance of both the beneficiaries and non-beneficiaries. Each tournament incentive in the experiment varied both the size and the demographic groups of the contest. Our results on the heterogeneous impact on network types suggest that social groups that are incentivized have an impact on test scores because a student from a minority background will have a different social network than that of a non-minority poor student. Also, if exclusion from the opportunity to achieve awards demotivates students, then the proportion of students excluded could also potentially impact on outcomes. Further research is required to disentangle the two effects in order to better design policies that cater to the need of minorities.

1.10 Figures

COLLEGE-A

ENGLISH 1^{ST} YEAR T1	ENGLISH 2^{ND} YEAR control	ENGLISH 3^{RD} YEAR control
HISTORY 1^{ST} YEAR T2	HISTORY 2^{ND} YEAR T1	HISTORY 3^{RD}YEAR T1
HINDI 1^{ST} YEAR T3	HINDI 2^{ND}YEAR T2	HINDI 3^{RD} YEAR T3
BENGALI 1^{ST} YEAR: control	BENGALI 2^{ND}YEAR T3	BENGALI 3^{RD} YEAR T2
SOCIOLOGY 1^{ST} YEAR control	SOCIOLOGY 2^{ND} YEAR control	SOCIOLOGY 3^{RD} YEAR control
PHYSICS 1^{ST} YEAR T2	PHYSICS 2^{ND} YEAR T1	PHYSICS 3^{RD} YEAR T3
CHEMISTRY 1^{ST} YEAR T1	CHEMISTRY 2^{ND} YEAR T3	CHEMISTRY 3^{RD} YEAR T1
MATHS 1^{ST} YEAR control	MATHS 2^{ND} YEAR control	MATHS 3^{RD} YEAR control
BIOLOGY 1^{ST} YEAR control	BIOLOGY 2^{ND}YEAR T3	BIOLOGY 3^{RD} YEAR control

Figure 1.10.1: A Example of Stratification by Arts and Science Faculties and by year of Study in College A. T1 represents income treatment, T2 represents caste treatment and T3 represents merit treatments.

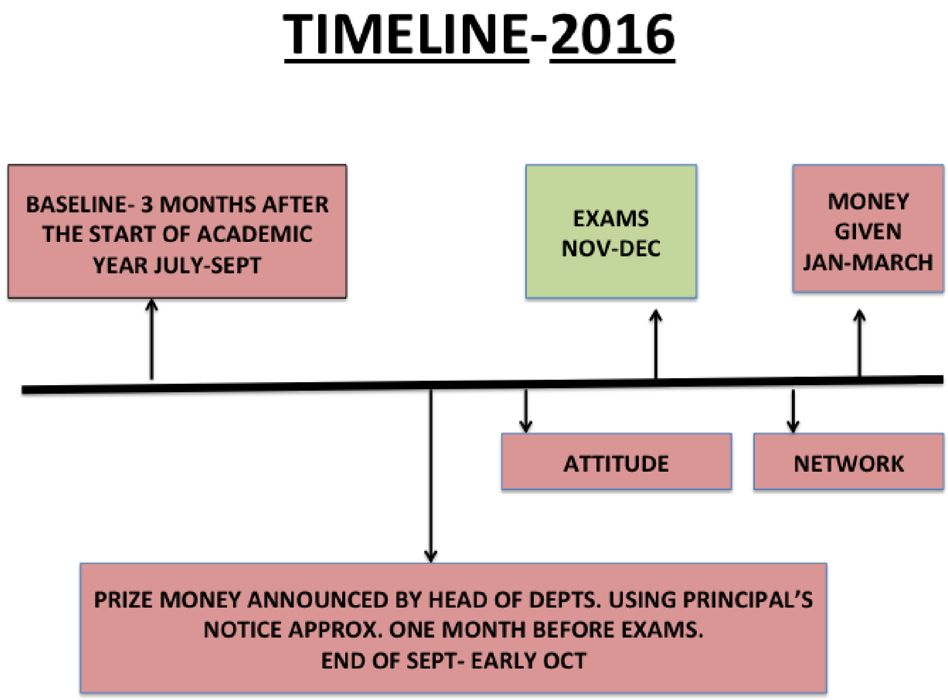


Figure 1.10.2: Time Line

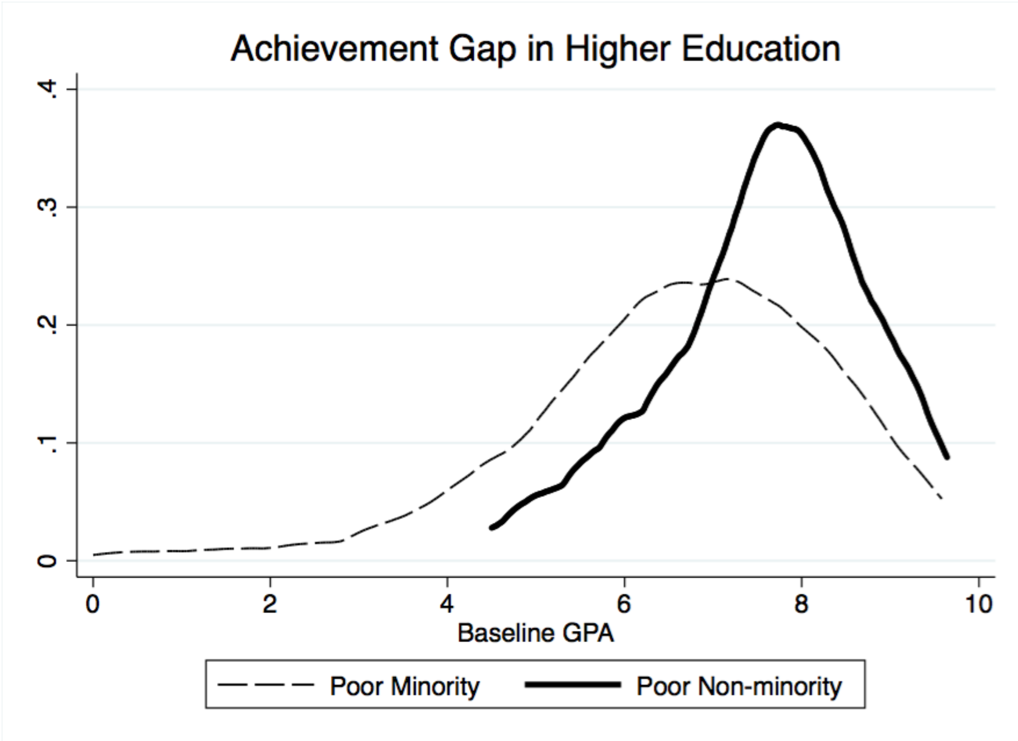


Figure 1.10.3: Distribution of Baselines Scores of Students whose parents earn USD 2,340 a year or below. Non minority refers to upper castes and Minority refers to lower castes. Grade Point Average (GPA) below 3 implies failing the university exam.

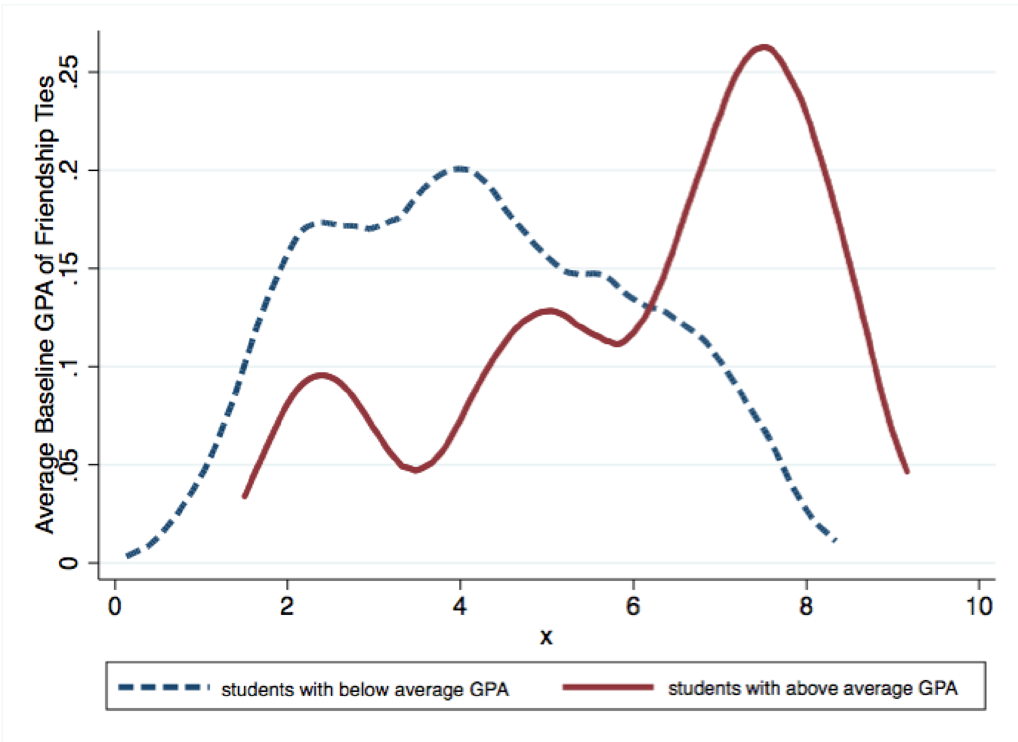


Figure 1.10.4: Assortative Matching by Test Scores

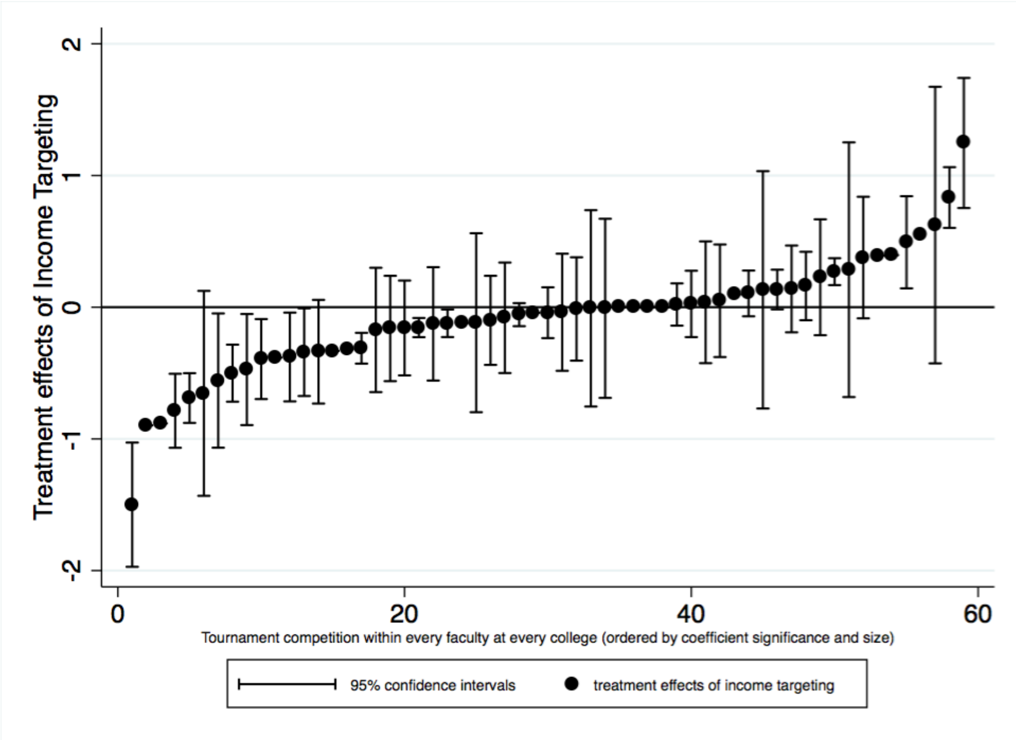


Figure 1.10.5: Treatment Effect of Income Targeting.

Notes: The sample includes all undergraduate students in my sample. The figures plot the treatment effects of income targeted incentives by faculty and and year in non-engineering colleges and by faculty in engineering colleges and their 95% confidence intervals. Tournament groups in both figures are ordered by significance and coefficient size.

SAMPLE DESCRIPTION OF MAIN TABLES

Full Sample (Intention to Treat) : 14,190 students from 11 Colleges.

Main Results of the Effect of Treatment estimated in Table 1.12.1 and Table 1.12.2 consists of students who took the exam after intervention, i.e. 12,998 students from 11 Colleges with 470 cohorts. The attrition rates are about 8 percent which are balanced across treatment arms.

Spillover effects are estimated in Table 1.12.4 using cohorts for which administrative data on income before intervention was available. There are 204 cohorts.

Pre- intervention network data were collected from 8 colleges out of 11 colleges. Post-intervention network data was collected in 1 college, which had only 8 % response rate.

Figure 1.10.6: Sample Descriptions

1.11 Summary Statistics

Here the table contains key summary statistics and randomisation tables using administrative records.

Table 1.11.1: Summary Statistics and Randomisation of Full Sample

Means Std. dev. in brackets	Control	T1	T2	T3	F-test	Obs
Baseline Scores Available	.46	.44	.45	.44	.97	14,190
	[.49]	[.50]	[.49]	[.49]		
Baseline Scores	6.05	6.07	6.28	6.05	.81	6,643
	[2.06]	[1.93]	[2.01]	[1.88]		
Drop out rates	.09	.08	.09	.08	.93	14,190
	[.28]	[.26]	[.28]	[.27]		
Lower Caste drop outs	.6	.6	.6	.6	.97	1,1,02
	[.26]	[.26]	[.28]	[.27]		
Missing data on caste	.25	.17	.21	.17	.47	14,190
	[.43]	[.38]	[.40]	[.37]		
Proportion of lower castes per cohort	.51	.49	.50	.51	.85	11,226
	[.15]	[.12]	[.13]	[.13]		
Proportion of eligible poor per cohort	.09	.09	.09	.11	.73	6,542
	[.11]	[.12]	[.11]	[.12]		
Class size	48	48	52	55	.93	14,190
	[37.25]	[42.53]	[38.45]	[51.90]		

Notes: T1 is treatment group where all compete, i.e. untargeted; T2 is treatment group where only lower castes compete. T3 is the treatment group where only poor students compete . *Baseline Scores* are the university exam scores of undergraduate students before the intervention. The baseline scores of students in their freshman year are not available for they were yet to take their university exams. Baseline scores of two colleges are not available due to administrative issues that occurred post experiment. Data on *castes* are from college administrative records. Caste data are missing for two colleges due to administrative issues. Data on students having family *income* of USD 2,340 and below are obtained from college records and are used to identify poor students. Only one college had detailed income records. Other colleges have records only for students in the freshman year because they did not maintain records in previous years. In such cases, college principal collected income certificates of students to identify eligible candidates. *Class* is defined as a cohort who study the same major degree, are in the same degree year, in the same college. This is the unit of our randomisation. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

1.12 Tables

Table 1.12.1: Effect of Incentives on test scores

	Z Scores
T1:Cohort with Income Targeted Incentives	-0.147***
	(0.053)
T2:Cohort with Caste Targeted Incentives	-0.038
	(0.050)
T3:Cohort with Untargeted Incentives	0.037
	(0.047)
Observations	12,998
Cluster	470
Randomisation inference (c) based p values for T1	.013
Bonferonni adjusted P values T1	.017
R-squared	0.498

Notes: Outcome variable is grade point average (GPA) of only major (honours) students in university exams standarised by full sample mean and standard deviation. The exams are graded centrally by the university that the colleges are affiliated to. The raw GPA is standardised using the mean and the standard deviations of whole sample of the experiment. Income Targeted Cohort is the cohort where only students whose annual family income is USD 2,340 and below are incentivised. Caste Targeted Cohort is the cohort where only lower castes within the cohort are incentivized. Untargeted cohort is a cohort where all are incentivized. Regression specification has stratifying dummies as controls. Standard errors are clustered at cohort level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The standard errors were also corrected for multiple hypothesis testing using Bonferonni's correction method.

Table 1.12.2: Negative Quantile Treatment Effect

	(1)	(2)	(3)
Z Score Quantile Reg	Median	Q(.25)	Q(.75)
T1: Income Targeted Incentives	-0.156***	-0.142***	-0.156***
	(0.0162)	(0.0257)	(0.0110)
T2: Caste Targeted Incentives	-0	-0.0158	-0
	(0.0213)	(0.0238)	(0.0198)
T3: Untargeted Incentives	0.0332	0.0189	-0
	(0.0214)	(0.0366)	(0.0211)
Observations	12,998	12,998	12,998

Notes: Outcome variable is grade point average (GPA) of only major (honours) students in university exams standarised by full sample mean and standard deviation. The exams are graded centrally by the university that the colleges are affiliated to. The raw GPA is standardised using the mean and the standard deviations of whole sample of the experiment. Income Targeted Cohort is the cohort where only students whose annual family income is USD 2,340 and below are incentivised. Caste Targeted Cohort is the cohort where only lower castes within the cohort are incentivized. Untargeted cohort is a cohort where all are incentivized. The table consists of full sample of students who sat for the university exams. Bootstrap Standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.12.3: Heterogeneous Effects of Tournament Incentives

	Z Score
untargeted	-0.112
	(0.0743)
caste targeted	0.0206
	(0.0801)
income targeted	-0.148*
	(0.0783)
low variance in scores (yes=1)	0.194*
	(0.109)
untargeted*low var	0.0495
	(0.121)
caste targeted*low var	-0.137
	(0.116)
income targeted*low var	0.00155
	(0.116)
Observations	6,401
Cluster	226
R-squared	0.445

Notes: Outcome variable is standardised grade point average (GPA) of only major (honours) students in university exams. The exams are graded centrally by the university that the colleges are affiliated to. The raw GPA is standardised using the mean and the standard deviations of whole sample of the experiment. Low heterogeneity in baseline scores is a dummy variable that takes the value 1 if the standard deviation of baseline scores of the cohort (i.e. the unit of randomisation) is below 1.3. The table consists of full sample of students whose baseline scores are available (N=6,401). Standard errors clustered at department-college-year level. The regression contains stratifying dummies. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.12.4: Heterogeneous Effect of Income targeted Incentives

	Z Score
T1:Cohort with Income Targeted Incentives	-0.136***
	(0.049)
Cohort with Income Targeted Incentives*poor student	0.038
	(0.126)
Poor student (yes=1)	0.016
	(0.104)
Clusters	109
Observations	2,884
Bonferroni corrected P value for T1	.02
Randomization Inference (c) p value for T1	.03
R-squared	0.368

Notes: Outcome variable is standardised grade point average (GPA) of only major (honours) students in university exams. The exams are graded centrally by the university that the colleges are affiliated to. The raw GPA is standardised using the mean and the standard deviations of whole sample of the experiment. Poor is a dummy if a student has family income below USD 2,340 per year. Income Targeted Cohort is the cohort where only students whose annual family income is USD 2,340 and below are incentivised. Standard errors are clustered at cohort level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The standard errors were also corrected for multiple hypothesis testing using Bonferonni's correction method.

Table 1.12.5: Demotivation Effect

	Z Score
Cohort with Income Targeted Incentives	-0.176*
	(0.0915)
Cohort with Caste Targeted Incentives	-0.0562
	(0.0747)
Cohort with Incentives for All	-0.0865
	(0.0716)
Cluster	102
Observations	1,958
R-squared	0.443

Notes: The sample consists of cohorts where proportion of poor students identified by pre-intervention administrative income data is 0. Outcome variable is standardised grade point average (GPA) of only major (honours) students in university exams. The exams are graded centrally by the university that the colleges are affiliated to. The raw GPA is standardised using the mean and the standard deviations of whole sample of the experiment. Poor is a dummy if a student has family income below USD 2,340 per year. Income Targeted Cohort is the cohort where only students whose annual family income is USD 2,340 and below are incentivised. Standard errors are clustered at cohort level. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.12.6: Effect of Targeting on Attitudes

	Merit Winners vs Other Winners
Cohort with Income Targeted Incentives	0.034*
	(0.018)
Cohort with Caste Targeted Incentives	0.006
	(0.013)
Controls	yes
Clusters	144
Observations	1,241
R-squared	0.833

Notes: The outcome variable is proportion of USD 11 dollars the survey respondent would like to allot to the student who received highest GPA in the whole cohort. The controls in this specifications are randomisation strata dummies, baseline test scores. Here cohorts with targeted incentives are compared with cohorts that receive untargeted i.e. merit based incentives.

Table 1.12.7: Effect of Targeting on Attitudes

	Poor Winners vs Others
Cohort with Income Targeted Incentives	-0.0350**
	(0.0175)
Cohort with Caste Treated Incentives	-0.0321**
	(0.0129)
Controls	Yes
Observations	1,149
R-squared	0.885

Notes: The outcome variable is proportion of USD 11 dollars the survey respondent would like to allot to the student who received highest GPA among the poor students whose family income is USD 2,340 and below in the whole cohort. The controls in this specifications are randomisation strata dummies, baseline test scores. Here cohorts with targeted incentives are compared with cohorts that receive untargeted i.e. merit based incentives.

Table 1.12.8: Effect of Targeting on Attitudes

	Lower Caste Winner vs Others
Cohort with Income Targeted Incentives	0.00369
	(0.012)
Cohort with Caste Targeted Incentives	0.0295**
	(0.011)
Lower Castes	0.0757***
	(0.009)
Baseline Scores	-0.00556**
	(0.002)
Controls	yes
Observations	1,149
R-squared	0.821

Notes: The outcome variable is proportion of USD 11 dollars the survey respondent would like to allot to the student who received highest GPA among the poor students whose family income is USD 2,340 and below in the whole cohort. The controls in this specifications are randomisation strata dummies, baseline test scores. Here cohorts with targeted incentives are compared with cohorts that receive untargeted, i.e. merit based incentives.

Table 1.12.9: Effect of Targeting on Excluded Students

	Z Score of Non Poor Students
Cohort with Income Targeted Incentives	-0.150**
	(0.0711)
Low ability student (=1 if yes)	-0.766***
	(0.106)
Cohort with Income Targeted Incentives* Low ability student	0.0328
	(0.141)
Observations	2,592
R-squared	0.497

Notes: Outcome variable is standardised grade point average (GPA) of only major (honours) students in university exams. The exams are graded centrally by the university that the colleges are affiliated to. The raw GPA is standardised using the mean and the standard deviations of whole sample of the experiment. Low ability student is dummy variable of a student whose baseline score is below average i.e. below 6.15. This regression consists of only sample of non poor students.

Table 1.12.10: Effect of Targeting on Network Types

Dep Vars: Z Score	Low-Ability	High-Ability
T1: Cohort with Income Targeted Incentives	-0.173**	-0.114
	(0.087)	(0.069)
T2: Cohort with Caste Targeted Incentives	-0.0595	-0.013
	(0.071)	(0.067)
T3: Cohort with Untargeted Incentives	-0.058	-0.084
	(0.080)	(0.067)
H: Friends with Above Median GPA (yes=1)	0.036	0.056
	(0.090)	(0.052)
T1xH	0.414**	0.094
	(0.16)	(0.082)
T2xH	-0.045	-0.002
	(0.20)	(0.075)
T3xH	0.279	0.122*
	(0.20)	(0.072)
T1+T1XH	.03	.88
T3+T3XH	.32	.21
R-squared	0.420	0.329
Observations	779	1,646

Notes: Outcome variable is standardised grade point average (GPA) of only major (honours) students in university exams. The exams are graded centrally by the university that the colleges are affiliated to. The raw GPA is standardised using the mean and the standard deviations of whole sample of the experiment. Low ability student is dummy variable of a student whose baseline score is below average, i.e. below 6.15.

Table 1.12.11: Effect of Incentives on Studying with Close Ties

Strength of Intimacy with Study Partners	IOS Scale Measure (1-7)
T1: Cohort with Income Targeted Incentives	0.663*
	(0.366)
T2: Cohort with Caste Targeted Incentives	0.432
	(0.315)
T3: Cohort with Untargeted Incentives	0.137
	(0.256)
Cluster	38
Observations	130
Randomization inference (c) p value T1	.053
R-squared	0.218

Notes: Outcome variable is measured using Simon Gaechter et al. (2015) IOS Scale that is a measure commonly used in psychology to measure how connected one feels to another. Standard errors are clustered at department-year-college level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.12.12: Effect of Incentives on Type of Studymates

	Friends Baseline scores
T1: Cohort with Income Targeted Incentives	-0.938*
	(0.484)
T2: Cohort with Caste Targeted Incentives	-0.00715
	(0.393)
T3: Cohort with Untargeted Incentives	-0.814
	(0.496)
Controls	Yes
Randomization inference (c) p value T1	.07
Observations	97
R-squared	0.269

Notes: Outcome variable is the baseline GPA of the study mate of the respondents.Standard errors are clustered at department-year-college level. Controls are the stratifying dummies. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.12.13: Robustness Checks

	Z Score
T1: Cohort with Income Targeted Incentives	-0.147***
	(0.0537)
T2: Cohort with Caste Targeted Incentives	-0.038
	(0.038)
T3: Cohort with Untargeted Incentives	0.037
	(0.048)
R-squared	0.498
Cluster	106
Bonferonni adjusted P values of T1	.02
Observations	12,998

Notes: Outcome variable is grade point average (GPA) of only major (honours) students in university exams standarised by full sample mean and standard deviation. The exams are graded centrally by the university that the colleges are affiliated to. The raw GPA is standardised using the mean and the standard deviations of whole sample of the experiment. Income Targeted Cohort is the cohort where only students whose annual family income is USD 2,340 and below are incentivised. Caste Targeted Cohort is the cohort where only lower castes within the cohort are incentivized. Untargeted cohort is a cohort where all are incentivized. Standard errors are clustered at college-faculty-degree year level. *** p<0.01, ** p<0.05, * p<0.1. The standard errors were also corrected for multiple hypothesis testing using Bonferonni's correction method.

Table 1.12.14: Robustness Checks

	(1) Drop Eng. Dept.	(2) Drop Minority Colleges	(3) Only Higher Degree Year
Income Targeted	-0.144**	-0.118**	-0.159**
	(0.0620)	(0.0510)	(0.0722)
Caste Targeted	-0.0533	-0.0468	-0.0763
	(0.0569)	(0.0450)	0.0693)
Untargeted	0.0452	0.0143	-0.0554
	(0.0538)	(0.0420)	(0.0645)
R-squared	0.540	0.367	0.517
Observations	10,743	9,209	6,344

Notes: Outcome variable is standardised grade point average (GPA) of only major (honours) students in university exams. The exams are graded centrally by the university that the colleges are affiliated to. The raw GPA is standardised using the mean and the standard deviations of whole sample of the experiment. Column(1) consists of sample excluding engineering departments. Column(2) consists of sample excluding all colleges that are from the state of Meghalaya (i.e. tribal reserved region). Column(3) consists of all college students ,excluding engineering students, who are in their penultimate or final year of their undergraduate degree Standard errors clustered at department-college-year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

1.13 Appendix A: Summary Statistics

Here the section contains detailed balance tables and summary statistics

Table A1: Balance Test of Subjects

	(1) Untargted Incentives	(2) Income Targeted Incentives	(3) Caste Targeted Incentives
P values of F test	.92	.88	.88
Observations	270	263	266
R-squared	0.160	0.161	0.169

Notes: To understand whether subjects predict treatment, we did the following cohort level regressions: Each regression is a comparison between treatment versus control groups. $T = \alpha + \sum_{allsubjects} D_{subjects} + C_{collegefixedeffect} + \epsilon$ where, T is a dummy for whether the cohort has been treated or is in control. D is a dummy for subjects, i.e. whether the cohort is english major or history major etc. There are total 47 subjects. The focus is on the joint significance i.e. P values of the F-test, to see whether subjects jointly predict treatment. And we find that they do not joint predict any treatment. There are three regressions for three treatments.

Table A2: Pairwise Tests of Full Sample

Equality of Means Test	T1=T2	T1=T3	T2=T3
P values: Proportion of Availability of Baseline Scores	.72	.66	.92
P values: Baseline Scores	.61	.52	.90
P values: Drop out rates	.56	.66	.87
P values: Proportion of lower castes among dropouts	.96	.67	.68
P values: Missing data on caste	.62	.92	.56
P values: Proportion of lower castes	.83	.33	.50
P values: Proportion of eligible poor	.89	.68	.78
P values: Class size	.74	.69	.86

Notes: T1 is treatment group where all compete, i.e. untargeted; T2 is treatment group where only lower castes compete. T3 is the treatment group where only poor students compete. *Baseline Scores* are the university exam scores of undergraduate students before the intervention. The baseline scores of students in their freshman year are not available for they were yet to take their university exams. Baseline scores of two colleges are not available due to administrative issues that occurred post experiment. Data on *castes* are from college administrative records. Caste data are missing for two colleges due to administrative issues. Data on students having family *income* of USD 2,340 and below is obtained from college records and is used to identify poor students. Only one college had detailed income records. Other colleges have records only for students in the freshman year because they did not maintain records in previous years. In such cases,college principal collected income certificates of students to identify eligible candidates. *Class* is defined as a cohort who study the same major degree, are in the same degree year at the same college. This is the unit of our randomisation.

Table A3: Selection of Respondents

Equality of Means Test	Survey Respondents	Non Respondents	P value of Diff
Baseline Scores	6.25	6.10	.28
	[1.84]	[2.06]	
Lower Castes	.54	.53	.27
	[.49]	[.49]	
Students Eligible for Income Targeting	.02	.06	.00
	[.15]	[.24]	

Notes: Standard deviation are in brackets. This sample consists of students in 8 colleges who were present in class at the date when the baseline survey was conducted. Baseline surveys were conducted in 8 out of 11 colleges. Baseline surveys in the remaining 3 colleges could not be completed on time due to logistical reasons. Survey response of those present in class was 100 per cent. But only a subset of the respondents could be tracked back to the college administrative records because some students did not type their details fully. All students whose survey details could not be analysed due to incomplete response are treated as non respondents along with the students who were not present to do the surveys on the survey date. The table shows the balance along two observables between the groups who details can be analysed and the groups whose details cannot be analysed.

Table A4: Balance Test of Estimated Sample on Networks

Means	Control	T1	T2	T3	F-test	Obs
Baseline Scores	6.17	6.38	6.42	6.45	.86	2,231
	[2.14]	[2.02]	[1.98]	[2.14]		
Lower Caste	.55	.54	.56	.57	.74	4,572
	[.49]	[.49]	[.49]	[.49]		

Notes: T1 is treatment group where all compete, i.e. untargeted; T2 is treatment group where only lower castes compete. T3 is the treatment group where only poor students compete. Standard deviations are in brackets. This sample consists of students in 8 colleges who were present in class at the date when the baseline survey was conducted. Baseline surveys were conducted in 8 out of 11 colleges. Baseline surveys in the remaining 3 colleges could not be completed on time due to logistical reasons.Survey response of those present in class was 100 per cent. But only a subset of the respondents could be tracked back to the college administrative records because some students did not type their details fully. The sample only consists of students whose survey could be traced back to their administrative records and hence analysed. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A5: Pairwise Tests of Estimated Network Survey Sample

Equality of Means Test	T1=T2	T1=T3	T2=T3
P values: Proportion of Availability of Baseline Scores	.91	.94	.86
P values: Lower Castes	.51	.71	.29

Notes: T1 is treatment group where all compete, i.e. untargeted; T2 is treatment group where only lower castes compete. T3 is the treatment group where only poor students compete. This sample consists of students in 8 colleges who were present in class at the date when the baseline survey was conducted. Baseline surveys were conducted in 8 out of 11 colleges. Baseline surveys in the remaining 3 colleges could not be completed on time due to logistical reasons. Survey response of those present in class was 100 per cent. But only a subset of the respondents could be tracked back to the college administrative records because some students did not type their details fully. The sample only consists of students whose survey could be traced back to their administrative records and hence analysed.

Table A6: Summary Statistics and Randomisation on Estimated Sample

Means Standard deviation in brackets	Control	T1	T2	T3	F-test	Obs
Proportion of Availability of Baseline Scores	.46	.44	.45	.44	.97	12,998
	[.50]	[.49]	[.49]	[.49]		
Baseline Scores	6.15	6.14	6.36	6.33	.79	6,409
	[2.00]	[1.89]	[1.95]	[1.82]		
Missing data on caste	.27	.18	.22	.18	.43	12,998
	[.49]	[.38]	[.41]	[.38]		
Proportion of lower castes	.53	.53	.53	.52	.79	10,124
	[.49]	[.49]	[.49]	[.49]		
Proportion of eligible poor per cohort	.09	.09	.09	.11	.63	5,683
	[.10]	[.11]	[.10]	[.12]		
Class size	48	51.27	54.89	58.29	.89	12,998
	[37.25]	[42.53]	[38.45]	[51.90]		

Notes: T1 is treatment group where all compete, i.e. untargeted; T2 is treatment group where only lower castes compete. T3 is the treatment group where only poor students compete.There are 8% drop outs. The sample that we estimate are students who have not dropped out of the exams(N=12,998). *Baseline Scores* are the university exam scores of undergraduate students before the intervention. The baseline scores of students in their freshman year are not available for they were yet to take their university exams. Baseline scores of two colleges are not available due to administrative issues that occurred post experiment. Data on *castes* are from college administrative records. Caste data are missing for two colleges due to administrative issues. Data on students having family *income* of USD 2,340 and below is obtained from college records and is used to identify poor students. Only one college had detailed income records. Other colleges have records only for students in the freshman year because they did not maintain records in previous years. In such cases,college principal collected income certificates of students to identify eligible candidates. *Class* is defined as a cohort who study the same major degree, are in the same degree year, in the same college. This is the unit of our randomisation. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A7: Pairwise Tests of Baseline Variables in Estimated Sample

Equality of Means Test	T1=T2	T1=T3	T2=T3
P values: Baseline Scores	.95	.41	.51
P values: Missing data on caste	.58	.95	.54
P values: Proportion of lower castes	.82	.32	.50
P values: Proportion of eligible poor	.86	.33	.33
P values: Class size	.76	.67	.82

Notes: T1 is treatment group where all compete, i.e. untargeted; T2 is treatment group where only lower castes compete. T3 is the treatment group where only poor students compete. The estimated sample is group of students who sat for the university exams. 8% of the students dropped out from each treatment arms.*Baseline Scores* are the university exam scores of undergraduate students before the intervention. The baseline scores of students in their freshman year are not available for they were yet to take their university exams. Baseline scores of two colleges are not available due to administrative issues that occurred post experiment. Data on *castes* are from college administrative records. Caste data are missing for two colleges due to administrative issues. Data on students having family *income* of USD 2,340 and below is obtained from college records and is used to identify poor students. Only one college had detailed income records. Other colleges have records only for students in the freshman year because they did not maintain records in previous years. In such cases,college principal collected income certificates of students to identify eligible candidates. *Class* is defined as a cohort who study the same major degree, are in the same degree year, in the same college. This is the unit of our randomisation.

Table A8: Summary Statistics and Randomisation of Reported Networks

Means	Control	T1	T2	T3	F-test	Obs
All four upper caste network	.05	.03	.05	.04	.16	4,870
One lower caste and three upper caste network	.04	.04	.04	.04	.81	4,870
One upper Caste and three lower caste network	.03	.02	.03	.03	.15	4,870
All lower castes network	.10	.11	.12	.12	.71	4,870
Three high achieving reported friend network	.18	.18	.18	.18	.99	2,523
Having reciprocal friends	.67	.66	.68	.70	.89	5,708

Notes: T1 is treatment group where all compete, i.e. untargeted; T2 is treatment group where only lower castes compete. T3 is the treatment group where only poor students compete. Data on castes and baseline scores of the reported friends have been used to construct the group characteristics of the reported friends. Students were asked to report up to three friends. Standard errors clustered at department-college-year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A9: Pairwise Tests of Estimated Sample of Networks

Equality of Means Test	T1=T2	T1=T3	T2=T3
P values: All Upper Castes	.08	.03	.77
P values: All Lower Castes	.68	.50	.81
P values: One Upper Caste and rest Lower castes	.06	.35	.39
P values: One Lower Caste and rest Upper Castes	.68	.71	.45
P values: All reported High achieving friends	.97	.89	.90
P values: Reciprocal relationships	.48	.82	.62

Notes: T1 is treatment group where all compete, i.e. untargeted; T2 is treatment group where only lower castes compete. T3 is the treatment group where only poor students compete. Data on castes and baseline scores of the reported friends have been used to construct the group characteristics of the reported friends. Students were asked to report up to three friends. Standard errors clustered at department-college-year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A10: Balance Check for Students who reported having a Study Mate

Means	Control	T1	T2	T3	F-test
Pre-treatment Scores	7.22	7.21	7.18	7.02	.98
Upper Castes	.45	.47	.44	.65	.015
Response Rate	.07	.07	.07	.07	.52

Notes: Post intervention survey was done only in one college. 250 out of 2,255 students responded. Out of which 27 students did not report to have any study mate. The table provides the balance along observables of the respondents who reported to have study mates across treatment and control.

1.14 Appendix B: Tables

Table B1: Placebo Test: Comparison of Baseline Scores and Post-treatment scores on same sample of students

	(1) Before-Treatment GPA	(2) Post-Treatment GPA
Income Targeting	-0.00224	-0.346***
	(0.136)	(0.122)
Caste Targeting	0.141	-0.0913
	(0.144)	(0.114)
Untargeted Incentives	0.157	-0.124
	(0.132)	(0.120)
R-squared	0.385	0.367
Observations	6,409	6,409

Notes: Outcome variable is standardised grade point average (GPA) of only major (honours) students in university exams. The exams are graded centrally by the university that the colleges are affiliated to. The raw GPA is standardised using the mean and the standard deviations of whole sample of the experiment. The table consists of full sample of students whose baseline scores and post-treatment GPA scores are available (N=6,409). Standard errors clustered at department-college-year level. *** p<0.01, ** p<0.05, * p<0.1.

Table B2: Effect of Income Targeting and Proportion of Poor

	Coefficient of Income Targeting
Proportion of Poor	0.0651***
	(0.0225)
Observations	6,542
R-squared	0.007

Notes: Using only the sample where administrative records of colleges were used to identify students we find positive correlation of treatment effect estimates of income targeting and proportion of poor studenst who are eligible for prize incentives.

Table B3: Impact of Incentives on Students of Attitude

	Z Score
Poor Treated Cohort	-0.142*
	(0.084)
Lower Caste Treated Cohort	-0.085
	(0.086)
All Treated Cohort	-0.116
	(0.087)
Observations	1,215
Controls	yes
R-squared	0.621

Notes: Outcome variable is standardised grade point average (GPA) of only major (honours) students in university exams. The exams are graded centrally by the university. The raw GPA is standardised using the mean and the standard deviations of whole sample of students in the engineering college in the experiment. Standard errors clustered at cohort level. This sample consists of the same students who reported their scholarship preferences in the attitude survey *** p<0.01, ** p<0.05, * p<0.1.

Table B4: Impact of Incentives on Survey College

	Z Score
T1:Cohort with Income targeted Incentives	-0.222**
	(0.0857)
T2:Cohort with Caste targeted Incentives	0.0126
	(0.114)
T3:Cohort with Untargeted Incentives	0.00408
	(0.102)
Stratifying Dummies	Yes
R-squared	0.042
Observations	2,255

Notes: Outcome variable is standardised grade point average (GPA) of only major (honours) students in university exams. The exams are graded centrally by the university. The raw GPA is standardised using the mean and the standard deviations of whole sample of students in the engineering college in the experiment. Standard errors clustered at department level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B5: Effect of Incentives on Test Scores

	(1)	(2)	(3)	(4)
Z Scores	Lower caste Poor	Upper caste Poor	Lower caste Rich	Upper caste Rich
T1: Income targeted	-0.203	-0.0211	-0.173**	-0.0989*
	(0.161)	(0.155)	(0.0675)	(0.0572)
T2: Caste targeted	-0.143	-0.00550	0.0161	0.00622
	(0.200)	(0.154)	(0.0674)	(0.0494)
T3: Untargeted	-0.216	-0.0537	-0.0118	0.00555
	(0.192)	(0.137)	(0.0669)	(0.0443)
Cluster	82	66	190	194
R-squared	0.232	0.707	0.231	0.655
Observations	351	179	2,311	2,650

Notes: Lower (Upper) Caste Poor implies students with annual family income of USD 2,340 and Lower(Upper) Caste rich implies annual family income above USD 2,340. Outcome variable is standardised grade point average (GPA) of only major (honours) students in university exams. The exams are graded centrally by the university that the colleges are affiliated to. The raw GPA is standardised using the mean and the standard deviations of whole sample of the experiment. The table consists of sample of lower caste students in colleges where income records are maintained. Column (1) and Column(2) consists of the sample of lower and upper caste who have been identified as having family income of USD 2,340 and below using administrative records. Column (3) and Column(4) consists of students whose income is above USD 2,340 as per college records. Standard errors are clustered at department-year-college level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B6: Group Size and Tournament Incentives

	Z Score
T1: Cohort with Income Targeted Incentives	0.0138
	(0.0587)
T2:Cohort with Caste Targeted Incentives	0.0771
	(0.0669)
T3:Cohort with Untargeted Incentives	0.00662
	(0.0607)
S: Small Class Size (=1or 0)	0.108
	(0.0737)
T1xS	-0.00304***
	(0.000916)
T2xS	-0.00212
	(0.00146)
T3xS	0.000146
	(0.00115)
R-squared	0.505
Observations	12,998

Notes: Outcome variable is standardised grade point average (GPA) of only major (honours) students in university exams. The exams are graded centrally by the university that the colleges are affiliated to. The raw GPA is standardised using the mean and the standard deviations of whole sample of the experiment. Small Class (i.e. the unit of randomisation) is dummy variable if the class size is below 48. The table consists of full sample of students in our experiment who sat the university exams (N=12,998). Standard errors clustered at department-college-year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B7: Ability and Tournament Incentives

	Z Score
T1: Cohort with Income Targeted Incentives	-0.0756
	(0.0935)
T2: Cohort with Caste Targeted Incentives	-0.0136
	(0.0979)
T3: Cohort with Untargeted Incentives	-0.0459
	(0.124)
High Scoring Class	0.370***
	(0.0853)
T1* High Scoring Class	-0.0775
	(0.111)
T2* High Scoring Class	-0.0191
	(0.117)
T3* High Scoring Class	-0.0287
	(0.136)
Observations	6,409
R-squared	0.451

Notes: Outcome variable is standardised grade point average (GPA) of only major (honours) students in university exams. The exams are graded centrally by the university that the colleges are affiliated to. The raw GPA is standardised using the mean and the standard deviations of whole sample of the experiment. High Scoring Class (i.e. the unit of randomisation) is dummy variable if the average GPA of the class is above 6.05. The table consists of sample of students for whom data on basleine scores are available. Standard errors clustered at department-college-year level. *** p<0.01, ** p<0.05, * p<0.1.

Chapter 2

Impact of British Colonial Reforms on Gender Differentiated Human Capital Investments[1]

We analyze the long-term impact of colonial legal reforms of matrimonial law introduced by the British in India in British provinces. Particularly, we examine the Child Marriage Abolition Act of 1931 that raised the minimum legal age of marriage for females to 14, and was the culmination of almost a century of matrimonial reforms in British provinces. The announcement of the abolition act in 1929 before its implementation had significant behavioural impact of varying intensity across British provinces, indicating differential beliefs about the enforceability of the law. Using this varying intensity and exploiting quasi-random variations of districts that were former British Provinces, we show, with large-scale microdata including administrative data and representative household surveys, that the reforms also had a positive sustained impact on female education and underage marriages in post-Independent India.

2.1 Introduction

There is an interest amongst economists about the persistent effect of British colonialism in India. A number of papers have shown significant persistent negative economic impacts (Banerjee and Iyer (2005), Iyer (2010)). This paper is the first to examine the persistent impact of British Colonialism on gender outcomes: female education and underage marriage of girls. We show positive effects of British colonialism on this margin.

[1]This chapter was written jointly with Dr. H.F. Tam.

Gender inequality in education is part of traditional cultures in many developing countries, where women get married earlier, are less educated and have poorer health outcomes than men Wong (2012). Besides being a serious concern in terms of equality of opportunity, it may also slow down economic growth (Klasen, 2002).[2] Our paper investigates whether the regions of India that have historically had legal institutions that fostered women's rights have better gender outcomes in the modern day. The analysis provides the basis for direct policy interventions on gender-biased social norms and practices in society. This is the first paper, to our knowledge, that investigates historical legal reforms to understand their long-run impact on gender outcomes.

The paper contributes to the literature on the impact of colonial institutions on modern day outcomes. While the literature on colonial institutions mainly focuses on changes in modern institutions (Acemoglu et al. (2001); Acemoglu and Johnson (2005)), our paper is closer to within country analysis, as in Michalopoulos and Papaioannou (2013) and Dell et al. (2015), showing that modern economic outcomes can be explained by the persistence of informal institutions. We contribute to the literature by studying the long-run impact of colonial institutions on household decisions on education and marriage, holding modern institutions constant. First, we map gender inequality in terms of education and marriage in modern India to historical political institutions. Next, we closely examine the short run and the long run impact of legal reforms introduced in two different polities.

After England took over India in 1858, India was divided into two different administrative institutions: the Princely States and British Provinces. This division ceased to exist post-independence; the State Re-Organisation Act 1956 re-divided India on the basis of linguistic identity. This led to a quasi-random distribution of Princely States and British Provinces within each modern state that made up independent India. Herein, we compare, within each modern state, the gender differential human capital investment between the regions that were under direct British rule and those that were Princely States in pre-independent India. Most of the variations in formal institutions are at the state level in India after independence. This implies that a comparison between regions within a state would allow us to control for almost all differences in formal institutions. Furthermore, the re-division of India along the lines of linguistic ethnicity allows us to further compare the impact of colonial social reforms on groups that share very similar ethnic identities.

[2](Klasen, 2002) finds that gender inequality in education is correlated with slower economic growth, both directly by lowering average human capital, and indirectly through its impact on investment and population growth.

We compare gender-related outcomes between regions that were once under the direct rule of the British, termed as British Provinces, and regions that were ruled by hereditary Indian rulers, known as native states or Princely States. We find that in former British Provinces, 5% fewer females marry under the current legal age of 18 years, and females have 1.6% higher chance of attending school between the ages of 10-16 years than those in the Princely States. This shows that regions that have different historical experiences behave differently, even after coming under the same common law.

Our hypothesis is that the legal reforms introduced by the British rulers forcibly changed the behaviour of the natives in the British provinces resulting in a positive long term effect on gender equality in India today. Before examining the long term effect of British laws, we first test the short run impact of the law using historical data to determine whether the introduction of legal reforms in British provinces in the past changed the behaviour of the natives in that region in the past. To examine this we use historical census data on marriage and literacy from 1911-1931 to estimate the impact of Child Marriage Restraint Act (1929) using the difference-in-differences strategy. The Child Marriage Restraint Act 1929, passed on 28 September 1929 in the British India Legislature of India, fixed the age of marriage for girls at 14 years and boys at 18 years. It is popularly known as the Sarda Act, after its sponsor Harbilas Sarda. It came into effect six months later on April 1, 1930 and it applied to all of British India. This created significant incentives for families to marry their children before April 1930. We use the Census data of 1911 to 1931 to capture the effect of the announcement of the law, with the Princely States as our control group. We find that announcement of the law increased the likelihood of girls married at age 5-10 by 2.8 percentage point more among the natives in British provinces, compared to the natives in the Princely States. Next, we look at the long run impact of the Sarda Act. The OLS estimates of the long run impact of the Sarda Act show that regions that were more aware of the law in 1929-1930 were less likely to marry their girls at young age for the cohorts born in 1958-1984.

Our findings highlight the importance of understanding social background when we think about how society responds to the development of the labour market. Social norms in society can persist for many years and can affect the decision to participate in education and the labour market for certain demographic groups. Even with the same formal institutions and economic environment, a society riddled with prejudices may not take full advantage of its economic transformation and development. Our paper also explains a significant part of the large regional variations in the degree of gender bias in India. The regions that were formerly British Provinces have better female education outcomes and fewer females marrying under the legal age compared to former Princely State regions. This allows for both academic and policy discussions about gender in India to go beyond geographical differences by states or by social class. We provide some explanations regarding why such differences continue to persist, but are unable to clearly determine the impact of each historical law. This would require a more elaborate analysis of historical data, which we hope to accomplish in future work.

The paper is divided as follows. Historical background is provided in Section 2, followed by a conceptual framework in Section 3. The data and empirical strategy are described in Sections 4 and 5 respectively. The Results and Discussions are provided in Sections 6 and 7, followed by a conclusion.

2.2 Historical Overview

The British first arrived in India through a trading company called the East India Company. They signed their first commercial treaty in the year 1612, granted by the Mughal Emperor Jahangir. It was not until 1757 that the British had their first military conquest. The East India Company had experimented with a number of political arrangements to maximise their commercial profits and minimise their administrative liabilities. Some states were brought directly under their control and some states entered into political and commercial treaties with the British. This experiment came to an end with the Great Revolution of 1857, when the British Government took control. The British divided areas under British rule into two territories: British India and Native (or Princely) States. British India represented all territories under the Majesty's dominion that were ruled by the Queen through the Governor-Generals. The Native States represented independent kingdoms of all the Indian kings who accepted British suzerainty. They came under the governance of the Viceroy or the Governor-General, who was the head of the administration in India and a representative of the Monarch in India. A clear distinction between "dominion" and "suzerainty" was supplied by the jurisdiction of the courts of law: the laws of British India rested upon the laws passed by the British Parliament and the legislative powers of those laws vested in the various governments of British India, both central and local; in contrast, the courts of the Princely States existed under the authority of the respective rulers of those states (The Interpretation Act 1889, British Parliament). Although the East India Company enforced indirect control over the Princely States, the rulers of those regions were not passive figures. The indigenous rulers had their own customs and laws which they insisted on pursuing (Ramusack, 2004).

India became independent in 1947, at which time it was still administratively divided into regions of British India, regions ruled by other European colonisers like the French or Danish and the Princely States. This division rendered it difficult for the administrators to rule the country. The State Re-Organisation Act was passed in 1956 that re-divided India on the basis of linguistic ethnicity. This is discussed further in the identification section.

2.2.1 Social reforms

Before the British came to administer the Indian territories, matters of marriage, maintenance, succession and legitimacy were solved using different religious laws, such as Dayabhaga and Mitakshara law for Hindus, literary traditions of Ithna Ashari and Hanafi for Muslims, and several customary laws for tribal communities. When the British took control of India, they promised not to interfere with personal laws such as marriage, succession etc. (Carroll, 1983). However, they reserved the right to intervene using statutory laws, which would override all religious laws in personal matters. Social reforms that were introduced by the British depended upon the discretion of the Governor-Generals in charge and the native social reformers (see Chitnis and Wright (2007)). All the British reforms that were introduced by the Governor-Generals were in direct conflict with the existing laws of Indian society (Carroll (1983)). Most of the social reforms were not in the interest of the British, as they created tension between the natives and their British rulers. However, the laws were passed after much deliberation by the reformist Governor-Generals. The first of the most important social reforms introduced in colonial India was the abolition of Sati in 1829. Sati was only practiced by upper caste Hindus in Bengal, Rajputana and Central India. It was a practice that involved a widow immolating herself on her husband's funeral pyre. The reform was pushed forward by a native social reformer, Raja Ram Mohan Roy. Lord William Bentinck introduced this law, arguing that the general masses of India were uncivilised and would continue this custom if the British did not bring forward a legal reform making it a punishable offence. In a speech in 1829, he pointed out that Britain could afford to abolish Sati without fearing rebellion from the natives because the majority of Indian soldiers in the British army belonged to the tribes that did not practice Sati (Fisch, 2000). Since Sati was only practiced by few ethnic groups in India, it was possible to extend the law outside British jurisdictions. The British negotiated with the Princely States to abolish Sati - Rajputana was the last native state to abolish it in 1861 (Ramusack, 2004).

Since then, most of the social reforms were implemented within British Provinces but were not enforced in the Princely States. With the initiative of the educationalist Pandit Iswar Chandra Vidyasagar, the British passed the Hindu Widow Remarriage Act of 1856. Until then, widow remarriage among upper caste Hindus had been prohibited, and Hindu widows were expected to live a life of austerity (Peers, 2013). It was introduced with the rationale of reducing female infanticide (Law Commission, 1837) and was very unpopular among the natives. The law, however, deprived childless widows of inheritance (Law Commission Report, 1856).

Although Sati was abolished in all of India, as a practice, it was not as widespread as female infanticide and child marriage (Grey, 2013), which existed across all of India and in all religions. Unlike Sati, the practice of female infanticide was not restricted to upper caste Hindus. The abolition of female infanticide (1870) and child marriage were harder to implement as they went directly against the widespread age-old customs of the natives across castes and tribes (Grey, 2011). The laws related to these practices were again confined to the British Provinces. In 1891, the Age of Consent Law was passed that raised the age of consent to 12 years. This bill created a lot of tension among the native population (Chitnis and Wright (2007); Ramusack (2004)). The reforms were slow. It took the British almost forty years to pass the Child Marriage Abolition Act (also called the Sarda Act) in 1929, which raised the age of consent to 14 years.

In our paper, we will closely examine the impact of the Sarda Act on both historical and modern marriage outcomes.The Child Marriage Restraint Act 1929, passed on 28 September 1929 in the British India Legislature of India, fixed the age of marriage for girls at 14 years and boys at 18 years. It is popularly known as the Sarda Act, after its sponsor Harbilas Sarda.[3] It came into effect six months later on April 1, 1930 and it applied to all of British India.[4] With protests from the Muslim organisation in undivided India, a personal law called as Shariat Act was passed in 1937 that allowed child marriages among Muslims with the consent of the child's

[3]Before the Sarda Act (1931), a cult group called Brahmo Samaj pioneered by Raja Ram Mohan Roy abolished the marriage of girls below 14 years of age in 1872 under an act called as the Native Marriage Act. But it only applied to the members of that group

[4]Hatekar et al. (2007) found that after the Sarda Act the probability of girls marrying below the age of 14 years dropped dramatically among the upper caste using micro data from family genealogies.

guardian. Family matters were in general governed by personal religious laws such as the Shastric law for Hindus, and Shariat law for Muslims etc. The British social reforms mostly interfered with Hindu Shastric law, using statutory laws to override customary religious laws (Carroll, 1983). Hence our analysis focuses on the Hindu population of both the British provinces and the Princely States.

We compare the impact of the Sarda Act on marriages in the British Provinces and the Princely states. However, due to paucity of census data in 1941 and 1951, we can only analyse the effect of the announcement of the law.

In contrast to the reforms in the British Raj, there were very few gender reforms in the Princely States. The only Princely States that implemented gender-related reforms were the Mysore and Kathiawar Agency of Baroda. Dewan Sheadari Iyer of Mysore in 1894 abolished the marriage of girls below the age of 8, and marriage between girls under 16 years old to men over 50. This law was less stringent than the British Sarda Act. In the face of widespread discontent among the masses, the Mysore Princely State mostly implemented this reform by occasionally prosecuting the powerless lower castes (Ramusack, 2004). The political agent, Alexander Walker, of Kathiawar agency tried to abolish female infanticide among the Jhareja and Jetwa tribes, with little success (Walker and Willoughby, 1856).

2.3 Conceptual framework

The social reforms implemented under the British rule in India may explain the differences in educational outcomes between the former Princely State regions and those that were under direct British rule, through more than one channel. We first discuss how specific reforms in British India directly affected the decision-making of the household, before discussing three potential mechanisms that could generate the long-term persistence of gender inequality, years after Indian independence from the British rule: the persistence of social norms, an information friction channel and the impact on the re-allocation of household resources.

Early marriages bring monetary savings and reduction of effort cost to the families of daughters, as after marriage, they will no longer need to be taken care of at home. If parents are happy to see their daughters married at an early age, they may only choose to educate their daughters when the net return on education is very high. Raising the legal age of marriage increases the total amount of time the daughters stay at home, and thus the cost of raising them. If there exists economic opportunities for skilled labour, the households will have an incentive to educate their daughters to participate in the labour market to reduce the cost of looking after them for a prolonged period.

With the passing of the State Reorganisation Act, the Princely States and the British Provinces came under the same jurisdiction and laws. Observed differences in gender bias in education after this reunification could either be explained by differences in the perceived return of education *specific for female* or the historical persistence of cultural bias/dis-utility generated by female participation in activities outside the household.

Our empirical exercise attempts to highlight that social reforms have a slow but persistent effect. One possible explanation for the persistent gap may be due to information friction. Correctly inferring the returns of education could be costly, and households may only make inferences based on limited experiences of other members of the same village. A larger initial stock of human capital among females could help the community identify market opportunities that are suitable for females. Thus, differences in initial human capital stock generated by historical reform in British India could translate into differences in the perceived returns of education, particularly when the return of human capital rose rapidly after trade liberalisation in India.

An alternative explanation could be linked to current debates on the subject of women empowerment (Duflo (2012) provided an in-depth discussion). Higher female education may have a direct impact on resource allocation and decision-making within the household (e.g. Quisumbing (1994) found that better educated mothers invested more in girls; Breierova and Duflo (2004) found evidence of female education on reductions in fertility and child mortality). If the mechanism holds, an exogenous shock that increases female education would have intergenerational persistence simply because better educated mothers allocate more resources to educating their daughters. This intergenerational transmission mechanism will have a larger effect if the other two mechanisms are also in operation.

With the mechanisms discussed, we test the hypothesis that regions that were historically under British rule have better gender outcomes in the short run and in the long run, compared to regions that were Princely States. In the following sections, we discuss the data and empirical strategy used.

2.4 Data

Our main source of information on the administrative division between the Princely States and the British Raj is Baden-Powell (1892), which included a detailed map on the division between the Princely States and areas under direct British rule together with the year of acquisition for each district. As the landscape of the Princely States and direct British rule was mostly settled by 1857, we define a district to be under British direct rule according to Baden-Powell (1892), otherwise it is defined as a Princely State. The geographical distribution is presented in Fig. 2.9.2.

Our measure of human capital investment comes from two independent sources: District Information System for Education and the National Sample Survey. The District Information System for Education (DISE) provides administrative records for enrolment at the school level in India. The data is designed to cover all regions of India in terms of the administrative information for each school in each academic year, including the number of students of each gender enrolled and the number of classrooms in each school. As the distinction between the Princely States and regions of direct British rule is mostly at the district level, we aggregate the information at the district level.[5] For the analysis, we aggregate all schools in each district in terms of the number of students enrolled in each class by gender for each year between 2005-2013; this gives us estimates of the ratio of male to female students enrolled in each class in each academic year for 433 districts. Summary statistics are reported in Table 2.9.1. On average, the schools in India have 9 % more boys enrolled in Class 6 compared to girls.[6]

The National Sample Survey (NSS) 64-66th round (2007-2008; 2009-2010) is another important data source that allows us to measure school attendance at the individual level. We focus on school attendance for children aged between 10-16 years old to study human capital investment decisions beyond basic literacy. It gives us 155,989 individual records (of 10-16 years old) regarding their principal activities in the past 365 days, including school attendance, participation in domestic work, and casual waged work. We report the summary statistics in Table 2.9.2 for the sample we used. The average school attendance rate is 0.85, with, however, very high variance.

[5] We excluded Karnataka in the analysis in this sample due to the lack of data availability at the time of writing.

[6] This is the ratio of raw enrolment, i.e. it does not take into account the gender ratio of the population; taking the NSS estimates of the proportion of females from 10-16 reported in Table 2.9.2 as 0.464 (which is by itself a number that shows very high gender bias), there are 15 % more boys than girls in this age range.

We further look into the percentage of marriages under the legal age in the year 2006-2007 at the district level from the District Level Household and Facility Survey (DLHS Round 3) by the Ministry of Health and Family Welfare of India. The data reported marriages under the legal age of 18 for women and recorded all marriage ceremonies held during the three years preceding the survey, covering 570 districts.[7] We use the micro data of the survey for the year 2002-2004 (DLHS Round 2) to conduct analysis with respect to Hindu female who are beyond the age of 18 at the time of interview, there are 86,214 individuals that we could merge where they are now with the historical census data we have. In addition, we obtained the district level GDP per capita from the Planning Commission of the Government of India. The geographical controls, such as latitude and distance to the coast[8] for each district, are defined at the centroid of the districts.

To study the persistence of the marriage pattern and the impact of the Sarda Act in 1929, we digitized the data from the Census of India regarding the population and marriage status of male and females at the district level for 1901-1941, covering the major British provinces and Princely States.[9] The census data are available at ten years interval for 1901, 1911, 1921, 1931 and 1941. We mainly limit our analysis to the data from 1911-1931, as the data were available for all major provinces and have consistent definition of variables across years, while in 1941 data do not exist for some regions and the definitions are inconsistent with those reported in previous years. There are changes in district names since the independence. To map the historical data to modern data, we geocode the historical district name, and compute which modern district it falls into. If more than one historical district falls into the same district in administrative division post-independence, we associate the average of records from the historical districts to the modern district. This maps into 126 modern districts.

It has been shown that princely states have higher levels of access to heath centres, schools and roads compared to British provinces (Iyer (2010)). Table 2.9.3 shows the balance of economic variables between the British provinces and princely states. We see that in the Princely States there are more households where the head of the households are literate or completed higher education. Thus one would expect that the Princely States will have fewer child marriages. Furthermore, Princely states also appear to have a higher number of females and have higher income per

[7]The data is released through DevInfo 6.0 by UNICEF.

[8]Distance from the coast is the physical distance instead of travel distance.

[9]This includes Madras, Bombay, Bengal, Rajputana, Central Provinces, Central India Agencies, Mysore, Travancore, Hyderabad, Ajmer and Punjab

capita compared to British Provinces. Therefore, princely states can be expected to have more girls going to schools. However, in our analysis that follows, we show that British Provinces do better in terms of gender outcomes compared to the Princely States.

Using within state variations in districts that were formerly princely states, we show using three independent large-scale administrative data sets, that in districts that were formerly the Princely States there is a higher rate of child marriages and fewer girls go to school. Then, we control for several economic variables that could explain the gender differences in schooling outcomes between the two provinces. We hypothesize that the differences in human capital investments by gender between the two provinces are caused by legal reforms that were introduced by British colonizers only in the British provinces. We provide suggestive evidence of long term effect of legal reforms by examining the Child Marriage Abolition Act (1929) on child marriages and schooling outcomes across the two provinces.

2.5 Empirical strategy

To study the long run impact of social reforms on human capital investment, a common challenge is to control for modern institutions and ethnicity. Different ethnic groups may be starting with different social norms. Moreover, each ethnic group may have different laws and social institutions that endogenously emerge according to the customs and culture of the group. We will describe how the State Re-Organisation Act of 1956 could help us control for both ethnicity and modern institutions.

After independence in 1947, the State of India was divided into three main regions: regions that were formerly British Provinces, regions that were under the rule of hereditary Indian rulers, and regions that were formerly under other European rulers. This division proved difficult for administrative purposes. Thus, the government of India decided to divide India on the basis of linguistic ethnicity. This proposal was very popular among the masses. The Telegu-speaking people formed the state of Andhra Pradesh, Marathi-speaking people formed the state of Maharashtra, and Kannada-speaking people formed the state of Karnataka, etc.

Linguistic ethnicity is an important determinant of identity in India. Modern India is adversely affected by conflict and riots triggered on the basis of differences in language. Since the 1920s, there has been conflict between Assamese and Bengali-speaking people. In recent times, Bihari-speaking people have been targeted in

Assam. In Maharashtra, Marathi-speaking people target migrants from Bihar and South India. In recent times, there has been a movement towards the compulsory use of the Marathi language in Mumbai, including in the Municipal Corporation. (see Baruah (2003); Weiner (2015); Murthy (2006); Mitra (1995))

Therefore, each modern state in India has people speaking the same language but with different historical experiences in terms of direct and indirect British rule. Residents of each state were subjected to same state law after 1956; this is our key source of identification. We argue that the distribution of Princely States and British Provinces are quasi-random within each state. We assume that the British did not select groups of people with a particular type of gender preferences within ethnicities to subject them to direct British rule. People with the same ethnicity tend to share norms. It is hard to imagine that people that were subjected to direct British rule had systematically different gender preferences to those of the same ethnicity living in native states at the beginning of the British India era.

In this section, we investigate the effect of Princely States (as opposed to being directly ruled by the British) on modern gender differential human capital investment. The key differences between the two forms of control were the gender-related social reforms that were highlighted in the historical section of this paper. However, there are potential confounding factors, such as differences in income and geographical characteristics, which we try to control for.

We use the following specification to test the impact of the rule of Princely States on the male/female enrolment ratio in the DISE data.

$$MFR_{sdct} = \alpha * I[princelystates]_{sd} + X'_{sd}\xi + \delta_s + \gamma_t + \mu_{sdt} \tag{2.1}$$

MFR_{sdct} measures the ratio of male/female students enroled in class c in district d within state s in year t. α, the coefficient of interest, captures whether in Princely States there are systematically more male children enrolled in school. δ_s is the state fixed effect that captures the systematic differences between states, such as the gender ratio, unobserved gender bias in social norms, and the provision of schools. X'_{sd} is the district level controls that include the proportion of rural schools in district d, the average number of classrooms in schools in district d, log GDP per capita (in 2000), and latitude and distance to the coast. γ_t is the year fixed effect that controls for yearly variations in gender differences in school enrolment.[10]

Moreover, we use the following specification to test the impact of the rule of

[10] All standard errors are clustered at district level

Princely States on school attendance and participation in waged work of women aged 10-16 years in 2006-2010 using the NSS data.

$$y_{sdi} = \beta * I[princelystates]_{sd} * female_i + \gamma_s + \phi_s * female_{sdi} + X'_{sdi}\eta + D'_{sd}\sigma + \epsilon_{sdi} \quad (2.2)$$

Where y_{sdi} is an indicator of school attendance/participation in waged work [11] as the principal activity in the 365 days before individual i in state s of district d was interviewed. $I[princelystates]_{sd}$ is a district level indicator of whether district d in state s belonged to a Princely State. The coefficient of interest is β, which is the coefficient for interaction term between $I[princelystates]_{sd}$ and $female_i$, which is a female dummy variable for person i. This captures whether females do worse in Princely States compared to direct British-ruled regions. γ_s is the state fixed effect for school attendance that captures state level differences in school attendance, such as different levels of provisions of schools. ϕ_s is a state-specific female fixed effect. This state-specific female fixed effect would mostly capture the different degrees of gender bias that exist in different states, which could be attributed to differences in gender norms between different ethnicities or differences in the labour market return of females. X_{sdi} is a set of socioeconomic controls that include the age of the child, the age of the head of the household, and the square of the age of the household head, an indicator for Muslim, Christian and other religions, an indicator of rural areas, and an indicator of the landownership of the households. D_{sd} is the geographic controls for district d in state s, which includes latitude and distance to the coast.

Moreover, we use the district level aggregate of the District Level Household and Facility survey to test the impact of Princely State's rule on the number of girls that marry under the legal age. We estimate the following equation:

$$M_{sd} = \sigma * I[princelystates]_{sd} + X'_{sd}\Phi + \kappa_s + \tau_{sd} \quad (2.3)$$

M_{sd} is a continuous measure of the percentage of marriages under the legal age in 2006-2007 for district d in state s, σ is the coefficient of interest as it tells us whether in former Princely State regions, more marriage are carried out under the legal age. X_{sd} is the district level controls that include latitude, distance to the coast and log GDP per capita (in 2000). κ_s is the state-fixed effect which controls for systematic differences across the states.

[11] This includes casual wage labor and not regular salaried work, and should be more relevant for the age range in our sample

2.5.1 Response to Sarda Act

In this paper we argue that British legal reforms affected the behaviour of the natives in British provinces by abolishing their traditional customs. To show the impact of the British legal reforms on the behavior of the natives, we begin with the study of the effects of the Sarda Act, the child marriage abolition law in 1929-1930. Fig. 2.9.3 plot the percentage of male/female married in the age group of 5-10 and 10-15 from 1901-1931 for the whole of India. The marriage pattern were stable from 1901-1921, while in 1931 the proportion of females married increase dramatically for all young age groups, particularly among the female. This is most likely due to the anticipation effect in the six months between its announcement and implementation (Census of India 1931). Fig. 2.9.4 shows the geographical distribution for the proportion of female married at age 5-10 in 1921, as well as the change from 1921-1931. It is not clear that places that experienced the highest increase in child marriage in 1931 were those that traditionally practiced child marriage in most numbers. Using historical census data, we estimate the following equation to test whether historical institutions explains the change in marriage pattern from 1921 to 1931.

$$M_{pdt} = \alpha * I[Britishdirectrule]_{pd} * I[t = 1931] + \gamma_p * t + \phi_t + \sigma_d + \tau_{pdt} \qquad (2.4)$$

M_{pdt} is the percentage of female who already got married at the age 5-10 in district d of pre-independent province p (i.e. political division before State-Reorganization Act) in year t between 1911-1931.[12] $I[Britishdirectrule]$ is an indicator which equals 1 if the district were under British direct rule.[13] α captures the differential changes in marriage pattern from 1921 to 1931 between former British direct rule regions and Princely States. Assuming there is no other factors that affect marriage pattern of the two regions differently between 1921-1931, α identify the effect of anticipation of actual implementation of the law. We control for the province specific trend ($\gamma_p * t$), district fixed effect (σ_d) and year fixed effect (ϕ_t).

2.5.2 Long run impact of Sarda Act

In this section we test the hypothesis that the awareness of the Sarda Act has long run impact on female marriage and education outcome, as measured in the DLHS

[12] It is defined as the number of married female at age 5-10 divided by the total number of female at age 5-10, reported by the Census

[13] so $I[Britishdirectrule] = 1 - I[princelystates]$, the variables we used in the specification described earlier

2002 for Hindu female aged above 18.

The equation of interest would be

$$y_{sdi} = \beta L_{sd} + \beta_2 M_{sd,1921} + X'_{sdi}\sigma + \epsilon_{sdi} \tag{2.5}$$

y_{sdi} are outcome variables measured in DLHS in 2002, for individual female i in state s, district d. It includes outcomes for marriage and education: indicator of marrying below the age of 14, marrying in the age range of 14-17, indicator for any level of schooling. L_{sd} is a measure of awareness of the Sarda Act for district d in state s since 1929. If the awareness of the Sarda Act reduces the probability of early marriage for female and increases the educational outcomes for females in the long run, we expect β to be positive. β_2 captures the accumulative effect of traditions and historical reforms before 1921 that could explains outcome in 2002.

Without a direct measure of L_{sd}, we use the proportional increase in the marriage ratio among girls aged 5-10, $m_{sd} = \frac{M_{sd,1931} - M_{sd,1921}}{M_{sd,1921}}$ as proxy, and estimate the following equation

$$y_{sdi} = \eta * m_{sd} + \gamma M_{sd,1921} + X'_{sdi}\sigma + \mu_{sdi} \tag{2.6}$$

$M_{sd,1921}$ and $M_{sd,1931}$ are the percentage of female who already got married at the age 5-10 in district d of state s in year 1921 and 1931 respectively, constructed by mapping new districts with their historical counterparts.

Given the historical context, m_{sd} should be positively related to L_{sd}, in the sense that in the districts with higher awareness of the law, more females in age range of 5-10 would be married as measured in 1931. This implies that η in equation 2.6 would have the same sign of the effect of β in equation 2.5.

2.6 Results

2.6.1 DISE

The OLS results for 2.1 of school enrolment are presented in Table 2.9.4. Column 1 reports the estimates for equation 1 for the ratio of gross enrolment of boys to girls in Class 6; the coefficient suggests that on average there are 2 percentage point more boys enrolled in schools than girls in former Princely States versus British-ruled regions. The availability of larger schools measured by the number of classrooms predicts lower boy to girl enrolment ratio. We also included log GDP per capita to control for the provision of schools and household budget constraints across districts within the same state, however, it is only marginally significant. Column 2 reports the same measure for Class 5, where the coefficient is very small and insignificant; this suggests that the results in Column 1 are mainly due to the dropping out of girls in higher grades rather than being driven by the differences in the gender ratio. Girls reach the age of 12 when they reach Class 8. Older girls may be more useful for household domestic work for which parents might take them out of school. If a community thinks that return of education of girls are low, then it is less likely for the community to invest in secondary schooling of girls.

In Table 2.9.5, we report Eq. (2.1) estimated by each class from 1 to 8. Comparing across columns, it is clear that the gender enrolment ratio only starts differing at Classes 6, 7 and 8, at which time the decision to attend school is more closely related to a human capital investment decision beyond basic literacy. The magnitudes of the coefficients across Columns 6, 7, and 8 are quite consistent at around 2-3 percentage point, suggesting that Class 6 is a critical time when, if girls drop out of school, they may not return, whereas those that stay in education are likely to proceed with similar probability to boys.

2.6.2 NSS

Table 2.9.6 reports the estimates of equation 2 on the main activities of children aged 10-16 years from the NSS data. Columns (1)-(2) report the estimates for school attendance, Columns (3)-(4) report the estimates for participation in waged work, and Columns (5)-(6) report participation in domestic work. The estimated interaction term Princely States*female is significant for school attendance, which shows that girls in Princely States are 1.6 % less likely to attend school compared to girls in British-ruled regions within the same modern state. We do not see similar significances in other outcome variables once we include the state female fixed effect to control for gender bias at the state level. The main effect of the Princely State for paid work participation (Columns (3) and (4)) is only significant when we exclude the interaction term with females, and the magnitude is small (0.5 % difference in the probability of market participation between Princely States and British-ruled regions). This could potentially be explained either by the lower age of marriage in Princely States or a small difference in the availability of market work.

However, our estimates on school attendance cannot be solely driven by the availability of market work. We further report estimates of Eq. (2.2) by Hindus and Muslims, as the historical overview section has shown that there were stark differences in how Hindus and Muslims responded to the social reforms in British-ruled regions. Column (1) in Table 2.9.7 reports the estimate for Hindus only. The coefficient is highly significant with a magnitude higher than that in the sample including all religions - females among the Hindu population are 2.1 % less likely to attend school in former Princely States, greater than the equivalent estimate of 1.6 % for the whole population. Moreover, the main coefficient of the Princely States is *positive* and is marginally significant for Hindus, which means that males are *more* likely to attend school in Princely States - this supports the hypothesis that the fundamental cause of the observed difference is driven by the persistence of cultural practices rather than the availability of education. On the other hand, the same estimate for Muslims in Column (2), despite its smaller sample size, is not only statistically insignificant but the sign of Princely States*female turns positive with a very large standard error. Instead of explaining the difference by time invariant inherent cultural differences between Hindus and Muslims, we tend to associate this difference in our estimates by how cultures interact with the implementation of the law in British-ruled regions before Independence.

In Fig. A2 we plot the percentage of married female at age 10-15 for districts that now belong to Madhya Pradesh - there were historically huge differences in how Hindus and Muslims responded to the Sarda Act of 1929 and Age of Consent Law of 1891.

2.6.3 Marriage under legal age

Table 2.9.8 reports estimates of Eq. (2.3) using the district level aggregate of the percentage of marriages under the legal age for females in 2006-2007. The coefficients estimated are positive, highly significant and robust upon inclusion of log GDP per capita (Column (1) and (2)). Our estimates suggest that Princely States have approximately 5 percentage points more marriages under the legal age for females. In Column (3) and (4), we report the results using the mean age of marriage in 2002-2004 as an outcome variable; it can be seen that districts formerly belonging to Princely States have a lower mean age of marriage by 0.4 years. An average of 22.66 % of all marriages in India take place under the legal age for female[14]; our estimated 5 percentage points difference between Princely States and direct British-ruled regions explains a significant number of underage marriages in India.

[14]From our district level aggregate not weighted by population share in each districts

2.6.4 Sarda Act using Census Data 1911-1931

Above we provide a mapping of gender inequality to different political institutions. In this section we focus on the impact of the legal reforms under two different political rule that affected behaviour in the past.

The estimates for equation 2.4 are presented in Table 2.9.9. The coefficient of *Princely states*1931* is statistically significant in column (2) where we control for province specific trend. The coefficient estimate is 2.8, which shows that among girls aged between 5-10, there are on average 2.8 percentage points more girls among natives in British provinces who got married in 1931. The natives in British provinces feared the implementation of the act in the coming months and millions of girls under the age of 14 were married off. This result is also well documented in the census reports of the British. The Sarda Act applied to only British India, however in Princely States such as Mysore and Baroda also tried to enact laws abolishing child marriages. [15] We observe a slight bunching in female child marriages in the Princely States, but in British provinces it is on average more severe.

Compare to column (1), which we did not include province specific trend, the coefficient in column (2) is more significant and with a larger magnitude, this suggesting that provinces may have differential trends before 1931.

Table 2.9.10 reports the estimates for equation 2.6, in OLS . Column (1) and (2) report the estimates for the outcome of marrying below 14 and marrying between 14-17 respectively, where the coefficients for $M_{sd,1921}$ is statistically significant, suggesting that the historical marriage pattern does explain the probability of marrying below 14 for the cohort of 18-44 years old in 2002. The estimate for $M_{sd,1921}$ in column (3) for the OLS is 0.176, which could be interpreted as 10 percentage point increase in the proportion of females married at the age of 5-10 in 1921 predicts a 1.76 percentage higher probability of a girl marrying below the legal age in 2002. The magnitude of the coefficients suggests that a significant part of child marriage in India has a very strong historical roots, going far beyond 1921.

Moreover, the estimates for m_{sd} is negative and statistically significant in column (1)-(2). It shows that one percentage point increase in the proportion of female married in 1931 (between 5-10) predict a smaller probability of a girl getting married below legal age post-independence. With the assumption that an increase in child

[15] Mysore in 1894 abolished child marriage below the age of 8. Many reformers from Mysore who pushed for the legislation of the Sarda Act could not raise the age of marriage for girls in Mysore. Therefore, in Mysore there was a weak form of child marriage restraint reform.

marriage in 1931 proxy for a high awareness of the Sarda Act, the estimate provide evidence that the Sarda Act reduced child marriages in the long run. Column (3) report estimate for the outcome of school attendance for girls. We find that 1 % increase in proportion of girls married at the age 5-10 from 1921-1931 predicts a 0.77% higher chance of females getting some education in 2002.

In column (4)-(6) we further control for state fixed effect. The coefficient for $msd, 1921$ and $M_{sd,1921}$ becomes smaller in general, but remains significant and the same signs in predicting marriage pattern in post-independence period. Controlling for state fixed-effects, we find that 1 % more girls married at the age 5-10 in 1921 predicts a 10% lower chance of females getting at least some education measured in 2002, while the coefficient of $msd, 1921$ remains positive but not significant.

We find regions that experienced bunching of marriages in 1931 have fewer girls marrying below legal age and are more likely to have experienced schooling post-independence, controlling for cultural variation across regions up to 1921. One interpretation of this long run effect could be that in British provinces the natives anticipated the implementation of the law in 1931 and did not wish to get affected by it. Therefore, the generation most affected by the Sarda Act and it's the later generations are more likely to conform their behaviour to any new law. This may explain why regions that got affected by legal reforms imposed by a foreign administrative body behave differently than regions that are culturally similar but did not get affected by the reform.

2.7 Discussion and robustness check

2.7.1 Robustness check - Princely States that potentially underwent reform

As discussed in the previous section, it was documented that in two of the Princely States (Mysore and the Kathiawar Agency of Baroda), there were reforms related to child marriages independent of similar reforms in the British Provinces. In the previous section, where we estimated Eq. (2.2) and Eq. (2.3) (school attendance in NSS and marriage under the legal age), we did not exclude Mysore and the Kathiawar Agency of Baroda because their implementation is weak from the historical description. We present the estimates for Eq. (2.2) and Eq. (2.3), excluding these two Princely States as a robustness check, in Table A1 and Table A2.

In Table A1, we report the estimates for school attendance in the NSS data excluding Mysore and Kathiawar Agency of Baroda in Column (2). The coefficient of the Princely States indicator increases slightly from 1.6 to 1.8, which implies a bigger difference in female school attendance among 10-16 years old between the Princely States and the British Provinces. Similarly, in Table A2 we report the estimates of equation Eq. (2.3) for the percentage of marriages under the legal age and the mean age of marriage, excluding Mysore and Baroda. The coefficients again increased slightly upon the exclusion of the two districts (in Columns (2) and (4)), which is what one would expect if Mysore and Baroda had weak social reforms that were similar in nature to those in the British Provinces.

2.8 Conclusion

In this paper we show that two regions that have had different legal reforms in the past behave differently when placed under the same modern institution. In particular we find that girls are more likely to go to school in regions that have had gender reforms in the past. If two regions are given the same opportunities in terms of provision of schools, we argue that the region that has had gender related legal reforms will have more females exploiting the opportunities. Our findings support policy intervention that eliminates prejudice behaviour by showing its positive long term impact. Providing infrastructure by the social planner might not be enough for economic growth, we also need to change the bottlenecks on the demand side.

2.9 Figures and Tables

Figure 2.9.1: Timeline of key historical events

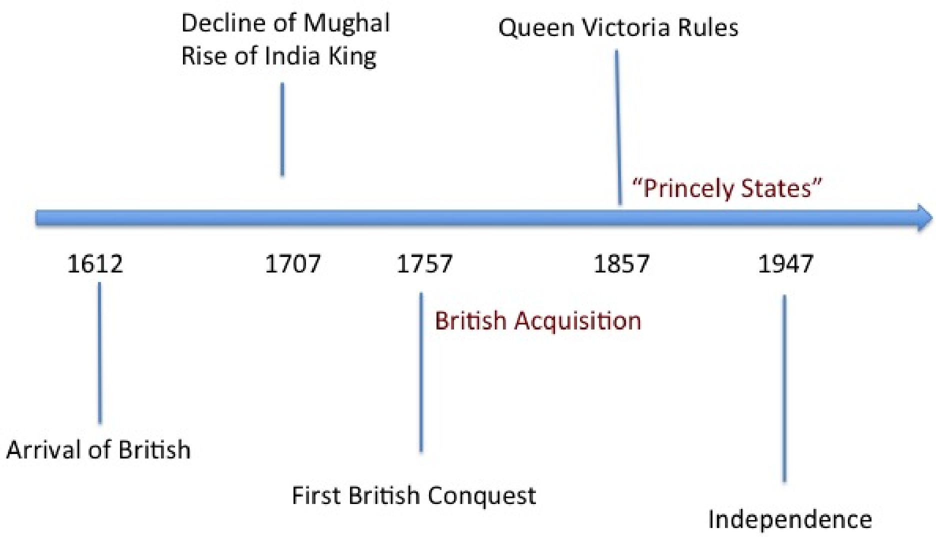
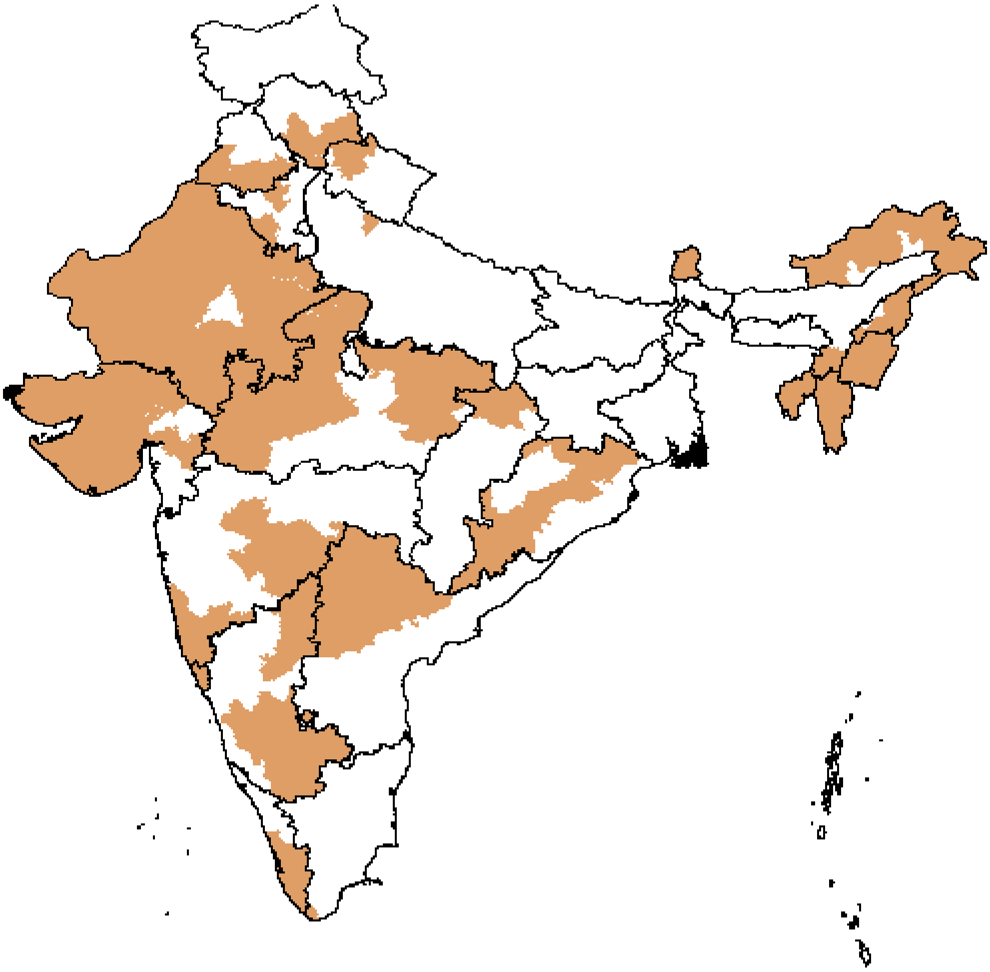
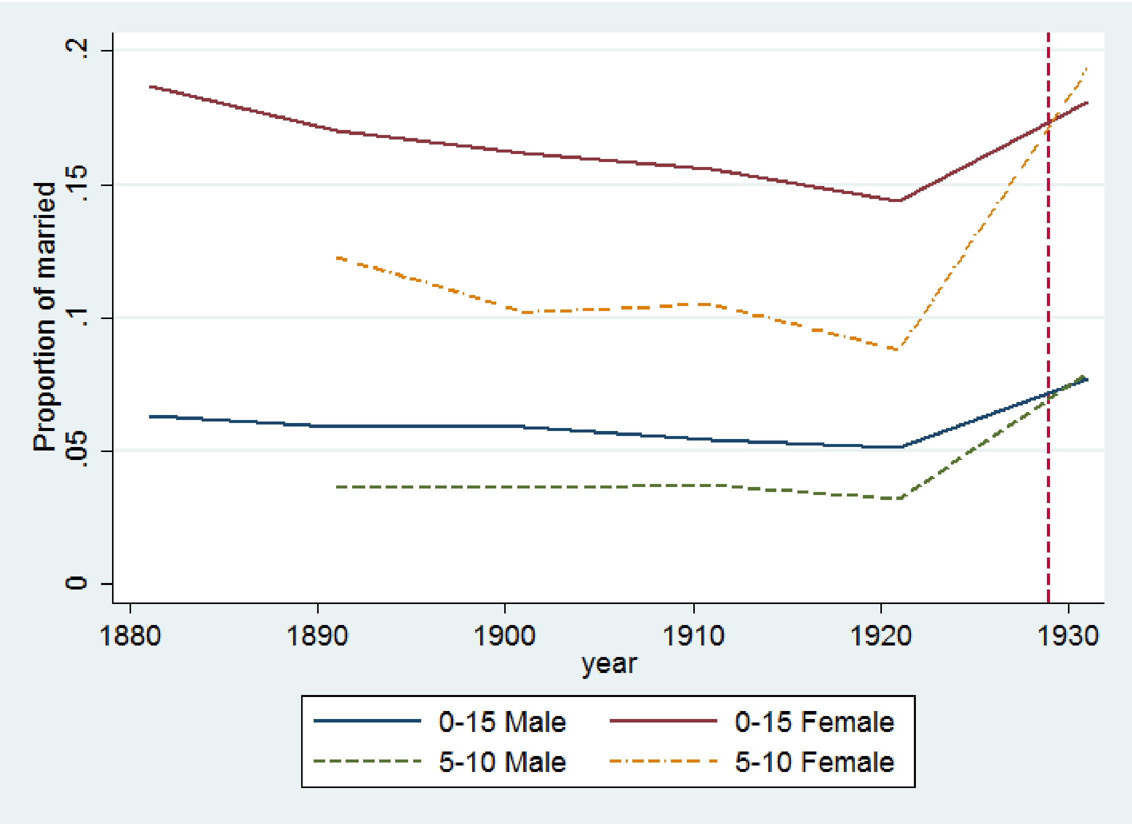


Figure 2.9.2: Distribution of Princely States and British direct rule regions



Note: The shaded parts were districts that belonged to Princely States and the white parts are districts that were under British direct rule.

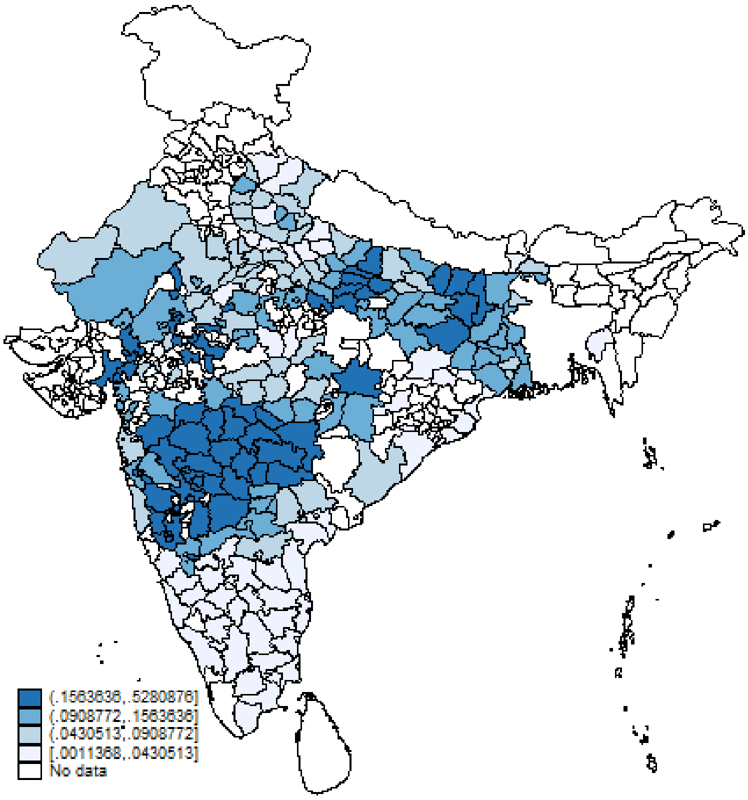
Figure 2.9.3: Marriage pattern in 1929-1930: time series



The graph plots the proportion of children married in each Census year, by gender and age group, for India as a whole.

Figure 2.9.4: Girls married 5-10 (%), all religion: Original 1931 administrative division

(a) Girls married 5- 10 (%): 1921



(b) Δ Girls married 5- 10 (%): 1931- 1921

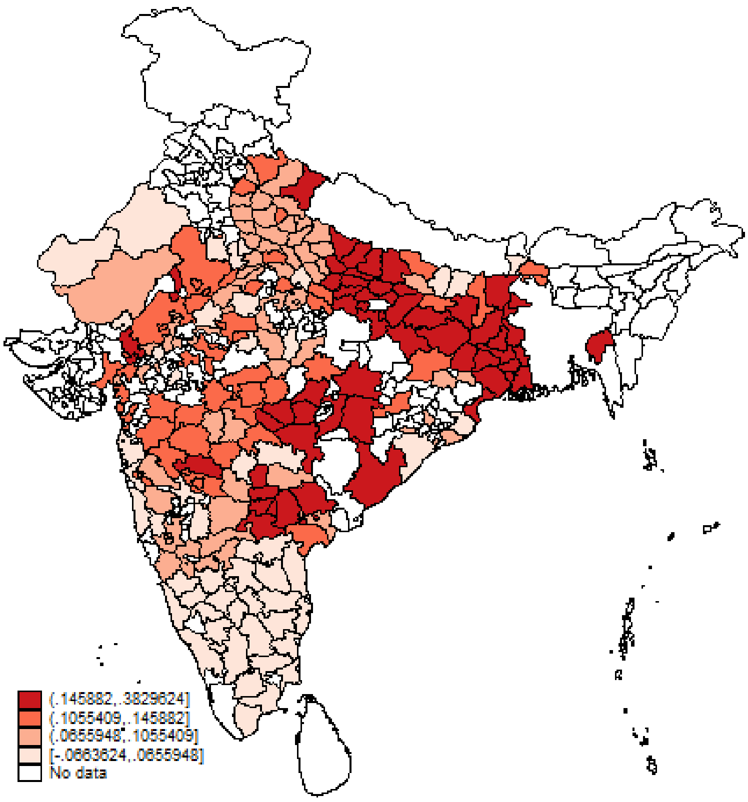


Table 2.9.1: Summary statistics of the DISE data

	mean	sd
Total boy / total girl enrollment in class 1	1.079	0.093
Total boy / total girl enrollment in class 2	1.070	0.099
Total boy / total girl enrollment in class 3	1.068	0.107
Total boy / total girl enrollment in class 4	1.069	0.115
Total boy / total girl enrollment in class 5	1.076	0.130
Total boy / total girl enrollment in class 6	1.088	0.157
Total boy / total girl enrollment in class 7	1.094	0.176
Total boy / total girl enrollment in class 8	1.103	0.195
Distance to coast	475.330	332.864
Proportion of rural schools	0.885	0.129
Princely states	0.261	0.439
Number of classrooms	4.418	1.778
Observations	2749	
Number of districts	433	

Note: Data aggregated at district level from DISE school records, forming a district level (unbalanced) panel for 2005-2013; Kerala not in the sample; Princely states is a {0,1} indicator. Distance to coast measured in kilometers from the centroid of each district.

Table 2.9.2: Summary statistics - NSS 64th and 66th round

	mean
School attendance	0.850
Female	0.464
Age	12.972
Scheduled caste	0.149
Scheduled Tribe	0.170
Head literate	0.724
Head complete primary	0.304
Head complete secondary	0.133
Head complete higher than secondary	0.185
Land ownership	0.895
Observations	155989

Note: Sample includes children in the NSS 64th and 66th round who aged 10-16 at the time of interview. School attendance is an {0,1} indicator for whether the principal activities in the past year of the child were attending school. Head literate, Head complete primary, Head complete secondary, Head complete higher than secondary were indicators for the education level of the household head, with base group illiterate. Land ownership is an {0,1} indicator for whether the household owns any land.

Table 2.9.3: Summary statistics - Princely States versus British Provinces

	British Provinces mean	Princely states mean
Princely states	0	1
Minimum distance to french port	634.3	784.3
Minimum distance to portuguese port	1134.1	1082.0
Distance to coast	477.9	357.7
Ln GDPPC (2000)	2.583	2.872
Manufacturing share of GDP (2000)	0.110	0.0821
Observations	291	98
Going to School	.843	.85
Female	.462	.467
Age	12.95	12.99
SC	.103	.23
ST	.182	.148
Owns Land	.89	.90
Observations	96,994	53,090
Head of Household literate(yes=1)	.708	.747
Head of Household primary edu(yes=1)	.301	.304
Head of Household secondary edu(yes=1)	.129	.136
Head of Household higher edu(yes=1)	.174	.203
Observations	43,215	24,080
Total boy / total girl enrollment in class 6	1.089	1.109
Total boy / total girl enrollment in class 5	1.079	1.090
Proportion of rural schools	0.887	0.847
Number of classrooms	4.402	5.203
Observations	2225	960

Note: Sample includes children in the NSS 64th and 66th round who aged 10-16 at the time of interview. School attendance is an {0,1} indicator for whether the principal activities in the past year of the child were attending school. Head literate, Head complete primary, Head complete secondary, Head complete higher than secondary were indicators for the education level of the household head, with base group illiterate. Land ownership is an {0,1} indicator for whether the household owns any land.Data aggregated at district level from DISE school records, forming a district level (unbalanced) panel for 2005-2013; Kerala not in the sample; Princely states is a {0,1} indicator. Distance to coast measured in kilometers from the centroid of each district.

Table 2.9.4: OLS Regression of Boy / Girl enrolment ratio at class 5/6: 2005-2013

Outcome: Ratio of boy/girl enrolment		
	Class 6	Class 5
	(1)	(2)
Princely states	0.0201**	0.00839
	(0.00993)	(0.00722)
Number of classrooms	-0.0201***	-0.0131***
	(0.00565)	(0.00476)
Ln GDPPC (2000)	-0.0303*	-0.0176
	(0.0166)	(0.0132)
Observations	2749	2749
State FE	Y	Y
Year FE	Y	Y

Note: Standard errors clustered at district level; The outcome variable is ratio of number of boys enrolled to the number of girls enrolled in each district, year and class, from 2005-2013; Other controls include latitude and distance to coast; Ln GDP per capita measured are district level GDP measured at 2000.

Table 2.9.5: OLS Regression of Boy/Girl enrolment ratio at all class: 2005-2013

	Outcome: Ratio of boy/girl enrolment							
	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8
Princely states	0.00280	-0.00142	-0.00374	0.0000208	0.00841	0.0201**	0.0300***	0.0256**
	(0.00589)	(0.00593)	(0.00602)	(0.00656)	(0.00723)	(0.00993)	(0.0114)	(0.0123)
Proportion of rural schools	0.0552**	0.0228	0.00334	-0.0205	-0.0478	-0.0814**	-0.0865**	-0.113**
	(0.0259)	(0.0269)	(0.0290)	(0.0318)	(0.0369)	(0.0408)	(0.0432)	(0.0472)
Number of classrooms	0.00596**	0.000161	-0.00387	-0.00744**	-0.0132***	-0.0201***	-0.0234***	-0.0266***
	(0.00269)	(0.00287)	(0.00319)	(0.00367)	(0.00473)	(0.00558)	(0.00597)	(0.00586)
Ln GDPPC (2000)	-0.00785	-0.00905	-0.00938	-0.0141	-0.0173	-0.0303*	-0.0382**	-0.0511***
	(0.00747)	(0.00840)	(0.00986)	(0.0114)	(0.0132)	(0.0166)	(0.0176)	(0.0185)
Observations	2749	2749	2749	2749	2749	2749	2749	2658

Note: Standard errors clustered at district level; The outcome variable is ratio of number of boys enroled to the number of girls enroled in each district, year and class, from 2005-2013; Other controls include latitude and distance to coast; Ln GDP per capita measured are district level GDP measured at 2000.

Table 2.9.6: Activity of 10-16 years old: NSS 64 - 66th round

	School	School	Waged work	Waged work	DW	DW
	(1)	(2)	(3)	(4)	(5)	(6)
Princely states=1	-0.00152	0.00601	-0.00565**	-0.00432	0.000973	-0.000434
	(0.00597)	(0.00580)	(0.00264)	(0.00286)	(0.00262)	(0.00162)
Princely states=1 × Female=1		-0.0160**		-0.00284		0.00299
		(0.00630)		(0.00298)		(0.00532)
Observations	150084	150084	150084	150084	150084	150084
state*female FE	Y	Y	Y	Y	Y	Y

Note: Sample includes children in the NSS 64th and 66th round who aged 10-16 at the time of interview. The outcome variables in column (1)-(2),(3)-(4) and (5)-(6) are indicators of school attendance, casual wage labor and domestic work as the principal activity in the year before survey respectively. Controls include: age at time of interview fixed effects; indicator of urban/rural; indicators for Muslim, Christian, Sikh and other religions; age of household head; indicators for education level of household head; district distance to coast and Latitude; Indicators for scheduled caste, schedule tribe and other backward caste. All regressions control for state fixed effects; Standard error clustered at district level.

Table 2.9.7: School attendance - NSS - by religion

	School		Waged work		DW	
	Hindu	Muslim	Hindu	Muslim	Hindu	Muslim
	(1)	(2)	(3)	(4)	(5)	(6)
Princely states=1	0.0115*	-0.0123	-0.00559	-0.00958	-0.0000947	-0.000354
	(0.00638)	(0.0146)	(0.00340)	(0.00641)	(0.00137)	(0.00325)
Princely states=1 × Female=1	-0.0217***	0.0147	-0.00250	0.00268	0.00405	0.00986
	(0.00746)	(0.0236)	(0.00371)	(0.00873)	(0.00534)	(0.0159)
Observations	109689	23137	109689	23137	109689	23137

Note: Sample includes children in the NSS 64th and 66th round who aged 10-16 at the time of interview. The outcome variables in column (1)-(2),(3)-(4) and (5)-(6) are indicators of school attendance, casual wage labor and domestic work as the principal activity in the year before survey respectively. Controls include: age at time of interview fixed effects; indicator of urban/rural; indicators for Muslim, Christian, Sikh and other religions; age of household head; indicators for education level of household head; district distance to coast and Latitude; Indicators for scheduled caste, schedule tribe and other backward caste. All regressions control for state fixed effects; Standard error clustered at district level.

Table 2.9.8: Marriage under legal age and mean age of marriage

Outcome:	% under legal age		Mean age of marriage	
	(1)	(2)	(3)	(4)
Princely states	5.186***	5.018***	-0.399***	-0.329**
	(1.361)	(1.369)	(0.147)	(0.149)
Distance to coast	-0.0141***	-0.0136***	0.000601	0.000172
	(0.00518)	(0.00512)	(0.000586)	(0.000529)
Latitude	1.119**	0.967*	-0.132**	-0.0910*
	(0.523)	(0.514)	(0.0572)	(0.0514)
Ln GDPpc 2000		-11.91***		1.192***
		(1.644)		(0.164)
Observations	568	508	570	508
mean	22.66	23.74	19.44	19.21
State FE	Y	Y	Y	Y

Note: Sample include percentage of marriage under legal age for female*100 reported of each districts, recorded from District Level Household and Facility Survey in 2006-2007 (from DevInfo 3.0); Mean age of marriage are district level mean age of marriage from DLHS 2002-2004 (from DevInfo 3.0); Robust standard errors reported in parenthesis.

Table 2.9.9: The impact of Sarda Act; Census 1911-1931; Difference-in-difference

	Outcome variable: Married female (%) 5-10					
	(1)	(2)	(3)	(4)	(5)	(6)
Sarda act/1931=1 × British Provinces=1	1.631	2.804**				
	(1.433)	(1.311)				
Sarda act/1931=1 × Year of direct control 69-100=1			2.930*	1.686		
			(1.505)	(1.733)		
Sarda act/1931=1 × Year of direct control 101-130=1			-2.156*	-2.069		
			(1.096)	(1.373)		
Sarda act/1931=1 × Year of direct control 131-160=1			2.290	1.241		
			(2.973)	(2.577)		
Sarda act/1931=1 × Year of direct control $\geq$ 160=1			6.558***	5.836***		
			(1.745)	(1.768)		
Sarda act/1931=1 × Ln dist. to reformers					-0.442	-1.489***
					(0.559)	(0.502)
Observations	694	694	694	694	525	525
District FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Province specific trend	N	Y	N	Y	N	Y

Note: Sample includes a panel that consists of 278 districts according to pre-independence administrative division, for the three Census year from 1911, 1921 and 1931. Sarda act/1931 is a dummy for the year 1931, capture the effect of Sarda act announced and implemented in 1929-1930. Base group in column (1)-(2) are districts that were princely states; Base group in column (3)-(4) are districts where year of acquisition are not reported from the map from Baden (1898) and thus were under indirect control. All columns include district and year fixed effects. Column (2), (3) and (6) control for province (pre-independence division) specific time trend. Robust standard errors reported in parenthesis.

Table 2.9.10: Using degree of bunching as proxy for treatment intensity

	Outcome					
	Married < 14	Married 14-17	Any Sch.	Married < 14	Married 14-17	Any Sch
	(1)	(2)	(3)	(4)	(5)	(6)
% change of married ratio 1931-1921	-0.00244***	-0.00822***	0.00774**	-0.00216**	-0.00494***	0.00161
	(0.000649)	(0.00215)	(0.00323)	(0.000841)	(0.00179)	(0.00234)
Married ratio 1921	0.176***	0.111*	0.0649	0.171***	0.0957*	-0.0963*
	(0.0405)	(0.0576)	(0.0679)	(0.0430)	(0.0545)	(0.0571)
British Provinces	-0.0215**	-0.0531***	0.132***	-0.0352**	-0.0227	0.0864***
	(0.00841)	(0.0101)	(0.0132)	(0.0135)	(0.0172)	(0.0175)
Observations	239482	239482	239482	239482	239482	239482
Number of dist.	188	188	188	188	188	188
Mean	0.0865	0.506	0.464	0.0865	0.506	0.464
State FE	N	N	N	Y	Y	Y

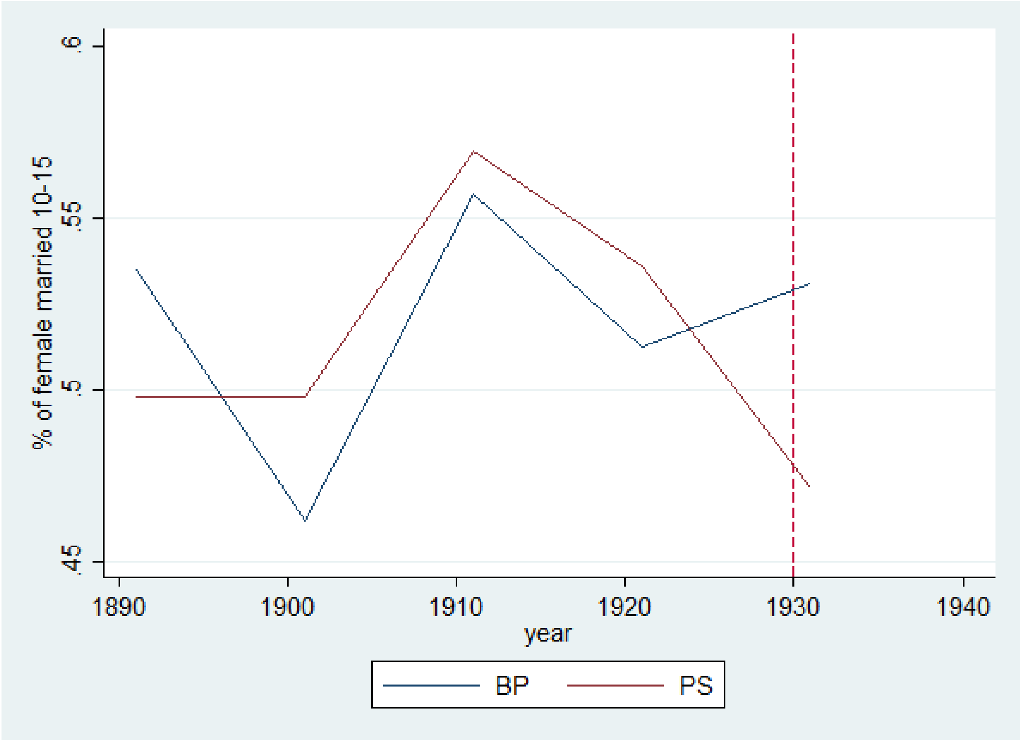
Note: Sample includes Hindu women aged 18-44 at DHLS round 2; Standard error clustered by historical districts. Married ratio defined as number of married women and widowed women at age 5-10 divided by the total number of women. % change of married ratio 1931-1921 defined as the difference of married ratio between 1931 and 1921 as proportion of married ratio 1921. Other controls includes individual controls and district level distance to coast, ln GDPpc (2000), Latitude.

2.10 Appendix

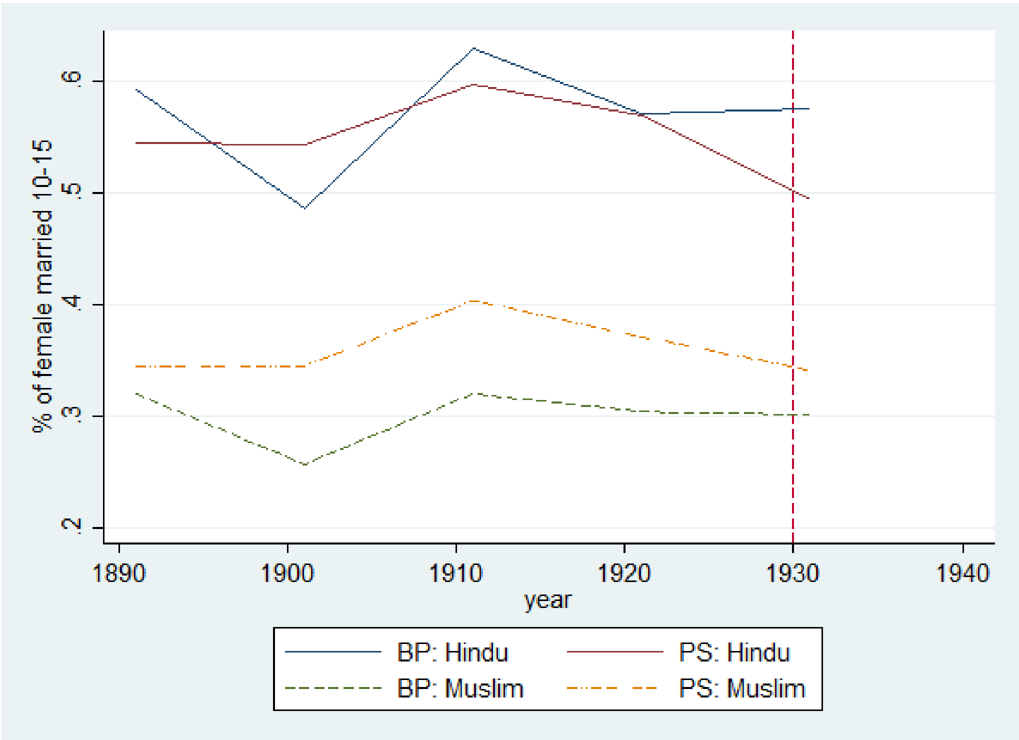
Figure A1: Geographical distribution of birth place of Pro-Sarda Act reformers

Figure A2: Percentage of married female - 10-15 years old - Madhya Pradesh

(a) All religion



(b) Hindu - Muslim



Note: Data from Census of India 1891-1931 and cover Central Provinces and Central India Agencies which belongs to Madhya Pradesh today; the red line denotes the enactment of the Sarda Act.

Table A1: Robustness check - exclusion of Mysore and Baroda: NSS Education

	School	School
	(1)	(2)
Princely states=1	0.00601	0.00556
	(0.00580)	(0.00580)
Princely states=1 × Female=1	-0.0160**	-0.0186***
	(0.00630)	(0.00650)
Observations	150084	146074
Include Mysore and Baroda	Y	N

Note: Sample includes children in the NSS 64th and 66th round who aged 10-16 at the time of interview; Specification same as in Table 2.9.6.

Table A2: Robustness check - exclusion of Mysore and Baroda: Marriage

outcome:	% under legal age		mean age	
	(1)	(2)	(3)	(4)
Princely states	5.018***	6.067***	-0.329**	-0.375**
	(1.369)	(1.367)	(0.149)	(0.151)
Distance to coast	-0.0136***	-0.0121**	0.000172	0.000217
	(0.00512)	(0.00516)	(0.000529)	(0.000532)
Latitude	0.967*	0.633	-0.0910*	-0.0896
	(0.514)	(0.529)	(0.0514)	(0.0550)
Ln GDPpc 2000	-11.91***	-11.75***	1.192***	1.190***
	(1.644)	(1.668)	(0.164)	(0.167)
Observations	508	496	508	496
Include Mysore and Baroda	Y	N	Y	N

Note: Sample include percentage of marriage under legal age for female*100 reported of each districts, recorded from District Level Household and Facility Survey in 2006-2007 (from DevInfo 3.0); Mean age of marriage are district level mean age of marriage from DLHS 2002-2004 (from DevInfo 3.0); Robust standard errors reported in parenthesis; Specification same as in Table 2.9.8

Table A3: Robustness check - exclusion of Princely States in the south: Marriage

outcome:	% under legal age			
	All states	Karnataka (exclu. Hyderabad)	Kerala	Exclu. Ke and Ka
	(1)	(2)	(3)	(4)
Princely states	5.018***	-7.346**	-7.188	6.550***
	(1.369)	(3.247)	(4.657)	(1.485)
Distance to coast	-0.0136***	0.0930***	0.0961	-0.0112**
	(0.00512)	(0.0252)	(0.0727)	(0.00522)
Latitude	0.967*	3.128	-1.270	0.168
	(0.514)	(2.497)	(1.458)	(0.553)
Ln GDPpc 2000	-11.91***	-8.846	-16.24	-11.03***
	(1.644)	(9.318)	(21.00)	(1.674)
Observations	508	23	14	467

Note: Sample include percentage of marriage under legal age for female*100 reported of each districts, recorded from District Level Household and Facility Survey in 2006-2007 (from DevInfo 3.0); Mean age of marriage are district level mean age of marriage from DLHS 2002-2004 (from DevInfo 3.0); Robust standard errors reported in parenthesis; specification same as in Table 2.9.8.

Chapter 3

Gender Bias in Education during Conflict: Evidence from Assam[1]

Using a large-scale novel panel dataset (2005-14) on schools from the Indian state of Assam, we test for the impact of violent conflict on female student's enrolment rates. We find that a doubling of average killings in a district-year leads to a 13 per cent drop in girl's enrolment ratio with school fixed effects. Results remain similar when using an alternative definition of conflict from a different dataset. Gender differential responses are more negative for lower grades, rural schools, poorer districts, and for schools run by local and private unaided bodies.

3.1 Introduction

Two billion people live in countries where development outcomes are affected by fragility, conflict, and violence (World Bank, 2016[2]). The effect of such fragility on educational attainment is generally harmful. This is not only likely to lead to lower growth in the future but may keep countries trapped in conflict. Although existing political and social institutions are unusual under stress during civil conflict, little is known about what resources can be supplied to existing schools during violent times to help wither the debilitating effects of violence.

The United Nations Millennium Development Goals have emphasized reducing schooling gender gaps since 2000 (United Nations, 2009[3]). Even as Latin America and Southeast Asia have made tremendous progress towards eliminating gender gaps

[1]This paper was coauthored with Dr. Prakarsh Singh

[2]World Bank. 2016. Helping Countries Navigate a Volatile Environment.

[3]United Nations. 2009. The Millennium Development Goals Report, 2009. New York: United Nations.

in educational attainment, the biggest room for improvement is in South Asia and West/Central Africa (Grant and Behrman (2010)). One hypothesis for the slower convergence in these regions is exposure to conflict by families residing in conflict-prone regions combined with an intrinsic bias towards investing in sons' education at the cost of daughters. The entrenched gender norms interact with a scarcity of resources to reduce school enrolment for girls. However, in Rwanda Akresh (2008) found that school-age boys and girls exposed to conflict have 0.5 and 0.3 fewer years of schooling, respectively. They argue that conflict disproportionately affects boys who had previously enjoyed an advantage in terms of education. Although the overall evidence on the impact of civil conflict on gender disparities in educational attainment is mixed, there has been little effort to analyze the supply-side factors (such as schooling inputs) that could help cushion shocks that lead to such distortions in investment decisions.

The conflict literature has transitioned from understanding cross-country correlations to a micro-level analysis of violence. By addressing reverse causality and omitted variables bias, these studies have contributed towards a deeper understanding of the consequences of civil conflict. Moreover, in recent years, the nature of armed conflict around the globe has shifted from civil wars and large-scale conflicts to more localized insurgencies. In this paper, we use school-level variation over time and regions in the north-east Indian state of Assam to uncover district-level effects of violence on girl's enrolment ratios.

Internal conflict imposes significant costs in terms of life and property. However, there may also be distortions to long-term investments due to conflict.[4] These distortions may include forgoing education or health investments if resources become more limited. The negative effects may be even larger for sub-groups that are either discriminated against or are of less economic value. Although gender inequality in education is a serious concern in and of itself, as it precludes girls from achieving equal opportunities, it could also lead to lower economic growth in the long run. For example, Klasen (2002) finds that gender inequality in education is correlated with lower economic growth, directly by reducing average human capital and indirectly through its impact on investment and population growth.[5]

[4] Blattman and Miguel (2010) find that one of the ways in which conflict depletes capital is through a massive flight of mobile forms of capital, possibly leading to low levels of new investment. In low-income countries, civil war makes poverty reduction and growth difficult to achieve (Murshed (2002), Verwimp et al. (2009) note the importance of taking the interaction between the armed actors and the households and individuals in affected communities seriously when studying violent conflict.

[5] Similarly, Baliamoune-Lutz and McGillivray (2015) demonstrate that gender inequality in

By compiling conflict data over the period 2000-14 across 22 districts in Assam, we test for the gender-differential impact of conflict on educational outcomes using a difference-in-differences approach with school, block, or district, and year fixed effects. Second, due to availability of detailed school-level data through the District Information System for Education (DISE) surveys, we test how resources for schools interact with intensity of conflict to affect enrolment for girls versus boys. Resources may be private or public and may be used to improve teacher–pupil ratios or infrastructure, through building new classrooms, equipping existing classrooms with blackboards, or stocking libraries.

Conflict's effect on education is understudied primarily due to a lack of available data from households in conflict-affected regions. Moreover, even if the data is available, it is often of very poor quality and households are not representative of the entire population. Moreover, household surveys do not allow us to consider a region-wide impact due to problems with agglomeration when using a small non-representative sample. Second, there may be spillovers on households not surveyed that may affect the biasedness of our estimates. For example, some households may decide to stop sending children to school in response to conflict and this may open up spaces for other households to send their children to school. Indeed the "treatment effect on the treated" is likely to be different from the "intention to treat" effect when considering consequences of civil conflict. Most reports on the consequences of civil wars on educational attainment for girls are case studies. However, recent empirical research in the area has found mixed evidence. Parents differentially invest in a son's secondary education as opposed to a daughter's depending on the context of conflict, and the intensity and nature of recruitment by rebel groups. Decline in women's educational attainment in war-torn societies has been observed in Chamarbagwala and Morán (2011), Shemyakina (2011) and Singh and Shemyakina (2016). However, Swee et al. (2009) and Kecmanovic (2013) find lower levels of education among the cohort of young males affected by war due to their participation in the conflict.[6]

primary and secondary education has a negative effect on income.

[6]Exposure to genocide in Rwanda resulted in a drop in educational achievement of schooling for all children but the impact was higher for boys from non-poor families (Akresh (2008)). In Nepal, educational attainment of girls who were of school age during the Maoist conflict actually increased (Valente (2013).

3.1.1 Background of conflict in Assam

The Indian state of Assam is located in the country's northeast and shares an international border with Bhutan and Bangladesh.[7] Assam has been mired in ethnic conflict since 1979, primarily between the Bodos (an ethnic tribe) and Muslim immigrants from Bangladesh. Illegal migration into Assam from Bangladesh during its independence movement in 1971 led to competition for resources and jobs in the region. The lack of economic opportunities for young males instigated the formation of militant groups. The large influx of (primarily) Muslim immigrants was a threat to the Bodos, who have sustained their community through agriculture for decades (Bhattacharjee and Phukan 2012)[8]. ULFA (United Liberation Front of Assam) was formed in 1979 under the leadership of Paresh Barua, with the aim of bringing about Assam's political separation from India, largely supported by the indigenous Bodo population. Its demands included detection of illegal immigrants, deletion of immigrants' names from voters' list, which effectively revoked their political power. When the government did not accede to their demands, ULFA targeted economically wealthy districts, abducted prominent businessmen, attacked politicians and civilians. Several militant outfits, such as the National Democratic Front of Bodoland, Bodo Liberation Tigers, and the Adivasi National Liberation Army sprang up in the 1980s.

Riots and violence between Bodos and non-Bodos have been sporadic but persistent. In recent years violent incidents have increased; one riot that erupted in 2012 killed 77 people (Asian Centre for Human Rights 2012[9]). Apart from fighting over resources and land, the Bodos under militant outfits have consistently expressed discontent with the state's policies (Goswami (2001)). Over time, different tribal factions of Assam have unsuccessfully demanded autonomy from the Indian government.[10] A recent successful attempt to evict illegal settlers by the government from protected forests provided a boost to the militant Bodo movement. For example, in the district of Kokrajhar, the Bodo heartland, Muslim migrants are regularly attacked by Bodo separatist rebels (Bhaumik 2012[11]).

[7] Assam has a population of 31 million with an area of 30,285 square miles (Census of India, 2011).

[8] "Assam Violence: A History of Conflict Rooted in Land." NDTV

[9] Asian Centre for Human Rights. 2012. National Commission for Minorities: Communalising Assam Riots.)

[10] The worst violence prompted by such tensions erupted during a controversial election in February 1983-nearly 3,000 people were left dead in that episode. After the 1983 elections, the state government tried to placate the rebels by signing an accord with the All Assam Students Union (AASU) in 1985, which was leading the campaign against the migrants. However, even though this was accepted by the moderate wing of the Bodos, the extremists opposed the accord.

[11] Bhaumik, S. 2012. What Lies behind Assam Violence? BBC News, 26 July.

The conflict data is a district-level panel for 14 years from 2000 to 2014, collated from the South Asia Terrorism Portal's (SATP's) list of all conflict events in South Asia. We employ two main indicators of violence for our regressions: total killed in violent incidents in a district-year, and total killed or injured in a district-year. The incidents have been coded manually from their "all events", which is based on news reports and may be susceptible to measurement error. However, as long as the error is not systematically correlated with educational variables, there should be no bias.

From Table 3.8.1, we find that violent conflict increases sharply in the years 2007-9 and again from 2012-14 and an average district in Assam is exposed to half as many incidents in the intervening period (2010-11). For example, in the year 2008, for the 22 districts in total, each district on average had 26 killings due to the conflict. The non-monotonicity in conflict is important for testing a causal channel between conflict and education. Military technology used by militants includes rocket launchers, grenades, ammunition, bombs, detonators, and M16 rifles. They appear to largely target businessmen and their family members and carry out their kidnappings for extortion. Along with extortion, there appears to be an upswing in the targeting of civilians such as doctors and forest personnel from the SATP data.

In Table 3.8.2, we show the variation in conflict across districts. Some districts, such as Karbi Anglong, Kokrajhar and Tinsukia suffer more than 20 civilian casualties on average every year between 2000 and 2014 due to violence. Others, such as, Karimganj, Hailakandi and Marigaon have relatively low levels of violence. Part of the reason could be that larger districts such as Karbi Anglong (population and area wise) would automatically be more prone to such incidents (just as population is a significant predictor of conflict onset in cross-country regressions, for example in Collier and Hoeffler 1998). Similarly, some districts may be more conducive to insurgencies because of their terrain. Forested areas could provide hiding space for militants. This is also a pattern seen in cross-country correlations (Fearon and Laitin 2003). Fig. 3.7.1 illustrates the high-conflict districts in the reddish spectrum whereas greener districts have lower levels of conflict from 2000 to 2014. Every district was affected by conflict over the time period. To control for larger districts having more violent incidents, or districts having more conflict because of their geography and proximity to an international border, we will include district fixed effects in our set of controls.

Fig. 3.7.2 shows the dynamics of conflict in each district. We observe that districts follow one of three patterns: first, there are several districts that have low conflict throughout the period 2000-14 (for example, Hailakandi, Jorhat, Lakhimpur, Marigaon, Nagaon). Second, there are districts that show high levels of violence in the early years but declining conflict in later years (Karbi Anglong, Tinsukia, Nalbari). Finally, there are a few districts that show an increase in the incidence of violence over time (Kokrajhar, Goalpara, Bongaigaon).

3.1.2 Gender inequality in education in Assam

Among all states in India, the social status of women was found to be the poorest in Andhra Pradesh, Assam, and Bihar (Planning Commission 2007[12]). In India's northeastern region, Assam ranks below the national average in terms of gender development index.[13] The gender gap in literacy is lower in the northeastern region than the rest of India. Over time, the literacy gender gap has narrowed down for all states except for Assam. In Assam the literacy gap has widened and the schooling enrolment gap by gender still persists in high schools.[14] Mahanta and Nayak (2013) find a greater gender gap in the enrolment ratio of grades 1 to 5 as compared to grades 6–7 over the period 1999 to 2010. This gap is highest for Assam (18.89 per cent) and lowest for Sikkim (0.84 per cent) among northeastern states. The figures are still above the national average.

In Fig. 3.7.3, we see girl's enrolment patterns over time in the different districts. We observe the following three patterns despite a positive overall time trend for nearly all districts: first, there are districts that have both high levels of girl's enrolment ratio as well as low levels of conflict (for example, Marigaon and Nagaon). Second, some districts have middling levels of conflict but vary in their trend of girl's enrolment (for example, stagnant and low enrolment in Sibsagar; sharply increasing in Darrang). Third, districts with high conflict show fluctuations in enrolment ratio (Dhubri, Karbi Anglong, Tinsukia). In order to find if these variations correlate with variations in violence, we need to run regressions across and within districts.

Another way of illustrating the state of and trends in gender inequality in Assamese schools could be by looking at the supply-side. Fig. 3.7.4 graphs the ratio of female to male teachers by district (2005-14). We find that, interestingly, districts with higher ratios of female to male teachers are usually the ones with low levels of girl's enrolment (Dibrugarh, Tinsukia, Golaghat, and Sibsagar). Similarly, several

[12]Planning Commission. 2007. Functioning and Performance of Swashakti and Swayamsiddha Programme in India

[13]Life expectancy at birth, of women in Assam is 58.1 years, lower than the national average at 63.3 years Mahanta and Nayak (2013)

[14]See Appendix Tables A1 and A2 for details

districts with lower ratios of female to male teachers had higher levels of girl's enrolment (Marigaon, Nagaon, Goalpara, and Barpeta). Thus, it is not straightforward to assume that gender inequality in educational enrolment across districts corresponds also to gender imbalance in teacher's employment across the same districts. Yet it may be the case that by increasing recruitment of female teachers in schools, the negative effects of conflict on gender balance are restrained. We test for this hypothesis in our policy recommendations section.

3.2 Data and empirical strategy

In the school-level dataset (2005-14), we have access to a rich set of variables. The data is collected by DISE at the school level (grades 1-8) for every village in all districts of Assam.[15] All schools falling under the Department of Education, tribal or social welfare department, local body, private aided, private unaided, and madrassas[16] are supposed to be covered under DISE. Most children who attend these grades would be in the age range of 6-14 years. Some of the variables at the school level include the type of management (as specified above), year of establishment, funds available for the school and the nature of funds, number of teaching staff and students, qualifications of teaching staff, and enrolment ratios of the students by grade, caste, tribe, and gender. We also know the number of students who repeat their grades by gender and grade, and presence of school facilities such as a library, blackboard, toilets, and computer facilities. This is an unbalanced panel at the school level with the average school being repeatedly surveyed 5.9 times.

We show some of the baseline means for schooling inputs by high- and low-conflict districts in Table 3.8.3. The districts are classified depending on whether their average annual killings are greater than or equal to the median or less than the median for the time period under consideration (2005-14). Girl's enrolment ratio at baseline is similar across low- and high-conflict districts. Several indicators, such as library books, male teachers, total children, total boys, and total girls are balanced between high- and low-conflict districts. However, there are significantly fewer computers in high-conflict district schools, which may also be related to lower funds available from grants for school development as well as learning material grants. Schools in high-conflict zones also appear to have significantly fewer female teachers.

The variation in conflict from the SATP dataset is at the district level and the

[15]DISE data have been used for many studies on schooling (DISE website).
[16]Schools that include the study of Islam, though this may not be the only subject studied.

number of schools in the sample is 86,558, each on average being repeated 5.9 times (n = 514,614) in our main regression). However, we will carry out a conservative check on our results by including block fixed effects – there are 149 blocks (smaller administrative units) under the 22 districts. Finally, we include school fixed effects. In all cases, we cluster our errors at the block level due to a small number of district clusters. When running the school fixed specification, we cluster our standard errors at the district level to obtain the *most conservative estimates* as the serial correlation in the error terms for all schools within a district is allowed to vary. The results are consistent when clustering at higher or lower levels. We run two sets of regressions.

In the first set, we find the first-order impact on the total boys and girls enrolled in schools in Assam. This is done using the following empirical specification:

$$y_{ijt} = \alpha_t + \beta_t + \gamma(Killed_{jt-1}) + \epsilon_{ijt} \tag{3.1}$$

where y_{ijt} is total children, total boys enroled, total girls enrolled and girl's enrolment ratio (i.e. total girls/total children) i in district j in year t. $\alpha_i$and $\beta_t$are school and year fixed effects. $Killed_{jt-1}$ is a measure of intensity of conflict in district j in year t. For running the above regression, we merge the school-level time varying data from DISE (available for 2005-14) with the conflict data collated from SATP described above. As the first academic year begins in 2005 and ends in 2006, we take district-level violence from 2005 as the "previous" year for the enrolment that is reported in 2006 at the end of the academic year. Each school is "exposed" to the conflict in its district in a year. The key innovation of the paper is to not only control for district fixed effects, the level at which conflict takes place, but to sequentially allow block fixed effects (149 blocks), and, finally, school fixed effects, as 86,558 schools are observed on average 5.9 times over the 9 years of data. If we did not control for block or school fixed effects, one could argue that attacks may be taking place in blocks within districts that had "worse" schools, either because of low human capital returns and the opportunity cost argument, or because of omitted variables that were correlated with schools having poorer outcomes and incidence of insurgent attacks. This would bias our estimates with the district fixed effects specification. In particular, if the omitted variables (for example, quality of public health services on the supply-side or parental education on the demand-side) positively affected girl's enrolment and were negatively correlated with the incidence of violence, then there would be a downward bias on our estimates. In other words, we would get a "bigger" negative coefficient on that would be biased and would show a much larger effect. Thus, controlling for block fixed effects should give a lower estimate of the impact (in absolute value) than controlling at a higher level (such as with district fixed effects).

We believe that this is the first paper in the literature to study variation in enrolment patterns *within* schools in response to conflict at the district level.

The empirical specification for our next set of regressions is as follows:

$$y_{ijt} = \alpha_j + \beta_t + \gamma(Killed_{jt-1}) + \delta Res_{it} + \omega(Res_{it} * Killed_{jt-1}) + \epsilon_{ijt} \qquad (3.2)$$

Here, Res_{ik} is an index of resources (number of computers, grants received, number of teachers, toilets, etc.) available at the school-level. γ can be interpreted as the impact of terrorism on girl's enrolment for schools having a resource index value equal to zero.

Overall, we would like to test the following hypotheses through regressions (1) and (2):

(a) γ is negative, implying a deleterious effect of terrorism on enrolment for girls after controlling for school and year fixed effects.

(b) The effect of conflict on girl's enrolment could be heterogeneous by grade. For example, if the opportunity cost of schooling is higher for higher classes (due to possibility of engaging in labor), the higher classes should experience a greater shortfall in enrolment in response to conflict. If, on the other hand, parents have a reduction in their expectations of the marginal returns to education, they may reduce education for their children in an earlier grade. This may happen if, for example, there is a higher risk of younger girls being targeted during a time of high conflict or if parents adopt a "wait and watch" policy for younger girls.

(c) δ and ω should be positive. ω refers to the marginal impact of terrorism on girl's enrolment for schools with higher resources and δ is the impact of resources on enrolment in peaceful district-years. We would expect estimates of γ to be negative and ω to be positive; thereby resources may work to cushion the negative gender-differential effects of the insurgency.

(d) We would also like to test which resources are the most effective at cushioning the effect of conflict on enrolment for girls.

The results first document if conflict affects girls differently. Second, we propose to test if there are heterogeneous effects by grade and, finally, we test for a cushioning effect on the gender gap by interacting a school's resources with conflict in the district. This will help policy makers in deciding which resources are most effective in curtailing the gender gap and fostering equal access to education during times of conflict.

3.3 Results

Table 3.8.4 shows the first order effect of conflict on total children attending schools. We show negative effects of conflict intensity on school enrolment where conflict intensity is measured by the total number of individuals killed in an incident. However, Table 3.8.5 shows that controlling for school fixed effects and year fixed effects, intensity of conflict increases total children enroled in schools. The results with the district fixed effects could have an omitted variable bias, where the omitted variable is positively correlated with total enrolmentof kids in schools and is negatively correlated with intensity of conflict, thereby resulting in large negative coefficients. Table 3.8.6 and Table 3.8.7 show the results by total girls enrolled and total boys enrolled in a school. There are two take-away messages from this table. First, the coefficient on the variable "killed" is negative and significant for both boys and girls across the different controls in specifications (1) to (4) and then (7) to (10). The school fixed effects regression gives an insignificant (and positive) coefficient for both boys and girls. This appears to suggest that school fixed effects absorb most of the variation that explains total children enrolled while total killings has little additional effect in explaining the total children enrolled over and above school fixed effects. The other pattern we notice is that conflict's effects on total girl's enrolment are more deleterious than on total boy's enrolment across all specifications. Table 3.8.7 also shows that there is not necessarily an increase in boy's enrolment rate in response to the girl's decrease (that is, we find a lack of substitution effects).

In Table 3.8.8, we illustrate the main results of the impact of conflict on girl's enrolment in the schools for the classes provided in the data set (classes 1-8). From our regression, we observe γ to be significantly negative for the effect of total killings on girl's enrolment ratios (total girls enrolled in school divided by total children enroled in school) without district and year fixed effects (column 1). The coefficient is equal to -0.0000480**.

Controlling for district and year fixed effects (in column 3), the main result is still significant at the 5% level but the girl's enrolment rate drops to -0.0000351**. However, our regression is high-powered because of the high number of observations. Note that the standard errors are clustered at the block level (the level of aggregation below a district) because we only have 22 districts in the sample, making clustering unreliable if the number of clusters are less than 42 (Angrist and Pischke (2008)). Results remain robust to clustering at the village level and/or adding block fixed effects (149 dummy variables) instead of district fixed effects. With block fixed effects (column 4), the estimate is -0.0000322** hinting at a small downward bias in the regression with district fixed effects.

In the most conservative regression specification with year and school fixed effects, we observe significance at the 5 per cent level in column (5). The interpretation for -0.0000339** observed can be thought of as follows. There are, on average, 90 children per school and 3,934 schools per district per year in our sample, and we are ultimately interested in the costs of conflict (at the level of the district) on the girl's enrolment ratio at the district level. This implies that to get the estimate of the number of female students who stop going to school in a district due to an extra killing in that district, the coefficient can be multiplied by 90*3,934 (= 354,060).

For every additional killing in a district in a year (6.16 is the mean of annual killings per district during 2005-14), we should see a decline in the district's girl's enrolment by -0.0000339*354,060 which equals 12 girls who are missing in school. For 6 killings per district (in an average year), 72 girls appear to drop out of school on average in that district. Going from the 5th percentile to a 95th percentile conflict-prone district increases killings from 0 to 21 in a year. This would imply a dropping out of 252 girls. Thus, effects of additional killings in the district are large if girl's enrolment is calibrated at the district level rather than the school level. Another way of thinking about the magnitude of the impact is to understand the impact on girl's enrolment rate at the school if the killings in a district double (increase by 100 per cent). This is done by running a regression of log of girl's enrolment on log of killed. As shown in Table 3.8.9, the impact is significant and about 13 per cent for doubling of killings in a district-year (from 6 to 12 killings).

When the definition of conflict is expanded to include the number of injured civilians in a district-year as well, we find less strong negative impacts of conflict on girl's enrolment (Table 3.8.21). This may mean that killings alone have more predictive power for reducing school enrolment of girls.

There could be a concern that a fall in girl's enrolment ratio in schools is reflective of pre-existing gender imbalance in the population where more boys are born (or survive) relative to girls over time and it is not due to conflict per se. Using Census data on child sex ratio, defined as number of females between the age of 0-6 years per thousand males between the age of 0-6 years, from the Directorate of Economics and Statistics, we find that child sex ratio has been on the rise in both low and high conflict districts. Kokrajhar, a district that has had one of the highest number of reported killings per year on average in the year 2000, has seen a rising trend in child female population between 2001-2011. Highest decline in child sex ratio has been reported in low conflict districts of Dhemaji and highest rising trend in child female population has been reported in the low conflict district of Hailakandi.

The results are in line with Chamarbagwala and Morán (2011), Shemyakina (2011), and Singh and Shemyakina (2016) as they had found a greater negative impact on girl's education from household surveys, but different from Swee et al. (2009)and Kecmanovic (2013), who had uncovered a larger negative effect for boys.

In Table 3.8.10 and Table 3.8.11, we split the samples by different school managements. We find that our main results are driven by two types of school management systems—local body and private unaided body. In both these cases, these schools appear to be locally administered and relatively autonomous. The results also show that schools, the majority of which are run by the Department of Education or Social Welfare Department, or private aided schools, do not show a decrease in girl's enrolment rates in response to violence (coefficients are insignificant and positive). Similarly, the coefficient on madrassas is insignificant (although negative). This may mean that households that send girls to attend public schools are "different" from those that send girls to attend private unaided schools and are unlikely to change schooling in response to conflict. On the other hand, it could also imply that public schools are better at retaining girls during times of uncertainty.

In Table 3.8.12, we illustrate heterogeneous effects of conflict on girl's enrolment by class or grade that would (most likely) be taking place at the schools run by local and private unaided bodies. Surprisingly, we find that the effects are not driven by (older) girls in higher classes. They seem to be driven by girls enrolled in classes 3 and 4. In fact, the enrolment ratio for older girls is positive, suggesting that perhaps they have crossed the conflict trap either from the demand-side, by allowing parents to send them for additional education, or from the supply-side, by giving them more opportunities to study in higher secondary schools established by the government. Although the eighth class is widely considered to be the critical barrier during peaceful times, it may not necessarily be the margin to focus on when studying the effects of conflict on girl's enrolment.

We also find that the significant results are driven by schools located in rural areas in Table 3.8.13. This may be because the conflict was focused in rural areas, but it could also be that most of the schools surveyed under DISE were in rural areas, lowering the power for the urban area regressions. Nevertheless, the results are in line with the effects on rural girl's enrolment found in Singh and Shemyakina (2016)

A large literature recognizes that lower incomes and poorer growth may be reasons for both onset and persistence of conflict, leading to a conflict-poverty trap. We check whether the responses to conflict for girl's enrolment are greater for poorer districts. This may be because the parents may be more affected by conflict and may feel a greater need to either switch to only invest in boys (at the cost of girls) or reduce the schooling expenditure for both boys and girls. Although we lack individual-level data on family incomes, district per capita income in 2005 that is available from the Directorate of Economics and Statistics in Assam is the next best proxy available. By defining a dummy for high gross domestic product (GDP) per capita to equal 1 for districts that are above the median income at baseline, we find that poorer districts have on average a three times higher response to civilian casualties than the richer districts, where effects are muted as shown in column (5) of Table 3.8.15. The effect for richer districts can be found by adding the two coefficients in the last regression. The stand-alone dummy for "High GDP" not present in columns (3)–(5) because it gets absorbed by district fixed effects. However, we do not find differential effects of Killed on enrolment by GDP growth rates, suggesting that baseline economic indicators are more important for explaining heterogeneity than the growth rates (which might also be endogenous to conflict).[17]

In Table 3.8.16, we delve into the mechanisms for "safety" of girls to isolate the impact of economic deprivation from simply, security of girls in that district during the insurgency. Here, we use data on the incidence of rapes against women from 2005, as provided by the Ministry of Home Affairs, New Delhi. We calculate the median rapes per capita and again categorize districts as "High Rape" depending on whether or not they had greater than the median level of rape incidence. If safety was a pertinent issue, we might expect that the interaction between casualties and insecurity of girls would lead to an even greater negative response on girl's enrolment. However, we do not find the interaction effect to be significant although it has the sign we would expect (negative).

[17]Available upon request: The median GDP growth rate is 6 per cent per year for a district between 2005 and 2012 according to the data available from the Directorate of Economics and Statistics, Assam.

3.4 Robustness checks

Schools have been targeted in isolated incidents in 2013. From the SATP data, we observe that on February 11, 2013, three schools were set ablaze in Goalpara district on the eve of the local village elections (Panchayat elections). In a similar incident, three schools in the rural Kamrup district were partially burnt down in the run-up to the same set of elections. We test for the robustness of the results by running our main specification (as in Table 4.1) for different sub-samples. We test for the robustness of our main result in Table 3.8.17 by excluding:

(a) Hailakandi, Goalpara, and Kamrup as these were the districts that suffered direct attacks on schools. The estimate remains significant and similar (the estimate is -0.0000229**) with school fixed effects implying that these districts were not driving the main results.

(b) The year 2013 from our analysis when such incidents took place in the run-up to the local elections. The estimate remains significant and similar (the estimate is -0.0000283**) in the most conservative regression (with school fixed effects) suggesting that direct violence targeted at schools does not drive our main results.

(c) International border districts of Cachar, Dhubri, Kokrajhar, and Karimganj, as the violence surrounding border areas was usually carried out by different rebel groups hiding in foreign countries and thus the dynamics of conflict may be different from the rest of Assam. The estimate remains significant and even higher in absolute value, the estimate is -0.0000552*** in the most conservative regression hinting that border district violence does not drive our main results.

In Table 3.8.18, we test for the impact of killed per capita in a district (dividing killed by district population in 2005, available from the Directorate of Economics and Statistics, Assam). The results remain consistent with this specification as well, suggesting that both aggregate levels of violence and likelihood of getting affected by violence are highly correlated. Results also remain robust to including per capita income and population as additional controls.

Next, we check for lagged effects of conflict by taking number of civilians killed in that district in the calendar year before the onset of the academic year (2.5 years before the end of the academic year). Observing the coefficients presented in Table 3.8.19, we find significant and negative effects on girl's enrolment rate but the size of the coefficient is smaller than that found in Table 3.8.8.

We were concerned that the conflict data may be biased as it is from collated news reports on the SATP website. The Bureau of Investigation's Special Branch under the Ministry of Home Affairs in Assam shared with us their dataset on district-wise bombings (and civilians, extremists, and security forces killed in those bombings) by the main insurgent group ULFA in the state from 2005 to 2012. We use this data to perform further checks on our results. In Table 3.8.20, we find consistent results that bombings that caused civilian or extremist casualties were correlated with a lower girl's enrolment ratio within the most conservative specification.

Additionally, we wanted to check if results may have been driven by extensive differential migration rates across conflict-prone districts and this may be related to having a girl or a boy. However, due to a lack of individual data on migration, we are unable to rigorously test this assumption. Nevertheless, we checked the Indian Human Development Survey from 2012 for Assam and found that out of 4,598 households sampled from seven districts, only 53 households (1.1 per cent) had migrated from another district in the last five years.

In Table 3.8.21 we redo the analysis using a different measure of conflict intensity. Here, we use the total number of killings or injured as the measure of conflict intensity and our findings are consistent with other measures of conflict.

Another way of measuring conflict is using the total number of violent incidents as a measure of conflict intensity. Incidents of violence are costly to collect. Additionally, the interpretation of the effect of the number of violent incidents on schooling outcome is not clear. Militancy in Assam involves bombing in urban centers, kidnappings of businessmen and targeted murders of ethnic minorities such as Hindi speaking Bihari migrants. Violent incidents that target rich business person or tea garden managers may not affect the schooling decisions of middle income or lower income households. Due to the lack of data on income, linguistic ethnicity of the students in DISE data, we are unable to use a total number of incidents as a measure of conflict on schooling outcomes.

Furthermore, we redo the analysis of girls enrolment rate using a balanced sample in Table 3.8.22 with school and year fixed effects. The results are consistent with the main findings.

3.5 Policy recommendations

Next, we elicit policy recommendations by checking for heterogeneities of responses to violence by school resources. Columns (1) and (2) in Table 3.8.23 show that even though grants to schools, such as teaching and learning grant, and school development grant, are useful in increasing girl's enrolment rate, they do not cushion the effects of conflict. All the regressions control for district and year fixed effects but results are robust to controlling for block fixed effects.

Computers and library books are neither strong predictors of girl's enrolment in peace nor in conflict (columns 3 and 4). What appears to matter most for improving the gender balance is having more teachers per pupil who are professionally qualified and recruiting more female teachers per pupil (columns 5 and 6). One additional professionally qualified teacher increases girls enrolment ratio by .02 percentage points. This is 377 times higher effect than that of conflict on girls enrolment ratio. Similarly, an increase of one female teacher per pupil increases girl?s enrolment ratio by .07 percentage points. This is about 1250 times more than the negative effects of conflict on girl?s enrolment. This implies that policies should focus more on incorporating more skilled human resources in schools and encourage more women to become teachers. Policies that stress physical resources within schools are less effective in combating the harmful effects of violence on women's empowerment.

3.6 Conclusion

In her 2014 Nobel Peace Prize acceptance speech, Malala Yousafzai said, "I tell my story, not because it is unique, but because it is not. It is the story of many girls". This paper tells the story of many girls who are missing in schools because of localized insurgencies. We find negative effects on school enrolment for girls and these responses are greater for lower primary school girls studying in schools run by local and unaided private bodies. Gender enrolment ratios in rural schools and poorer districts seem to be particularly negatively affected by conflict. The effects are robust to including district, block, or school fixed effects, along with year fixed effects, and to a host of other robustness checks. There also does not appear to be a corresponding increase in total boy's enrolment.

The sprouting of several Assamese militant outfits representing local tribes (such as Bodos) does not augur well for gender inequality in education. Recently, it was reported that "nine organizations representing the indigenous and tribal communities of Assam joined hands to form a political alternative for the coming Assembly elections in Assam." Hopefully, political concessions will lead to more peace in the state that has suffered from loss of life and capabilities. On the other hand, the government's policies to improve girl's enrolment during violent times in Assam should consider providing incentives for younger girls, and focusing efforts to build more public schools and monitoring local body schools.

Moreover, although school grants are useful for improving gender balance during peaceful times, policies that revolve around hiring female teachers and professionally qualified teachers appear to have the greatest impact on improving girl's enrolment. Nevertheless, it is also important to understand the socio-psychological reasons that lead some children away from school and these results, paired with finding demand-side explanations, can help policy makers spend resources more efficiently for gender equality and development.

3.7 Figures

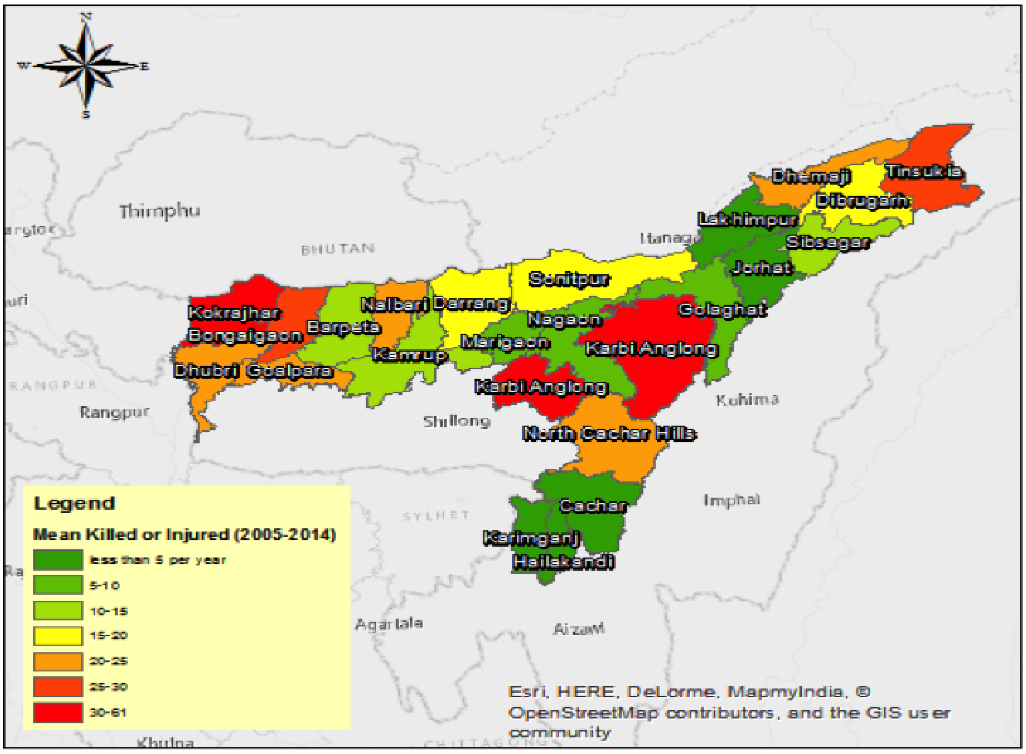


Figure 3.7.1: District heat map of Assam with mean killed or injured.

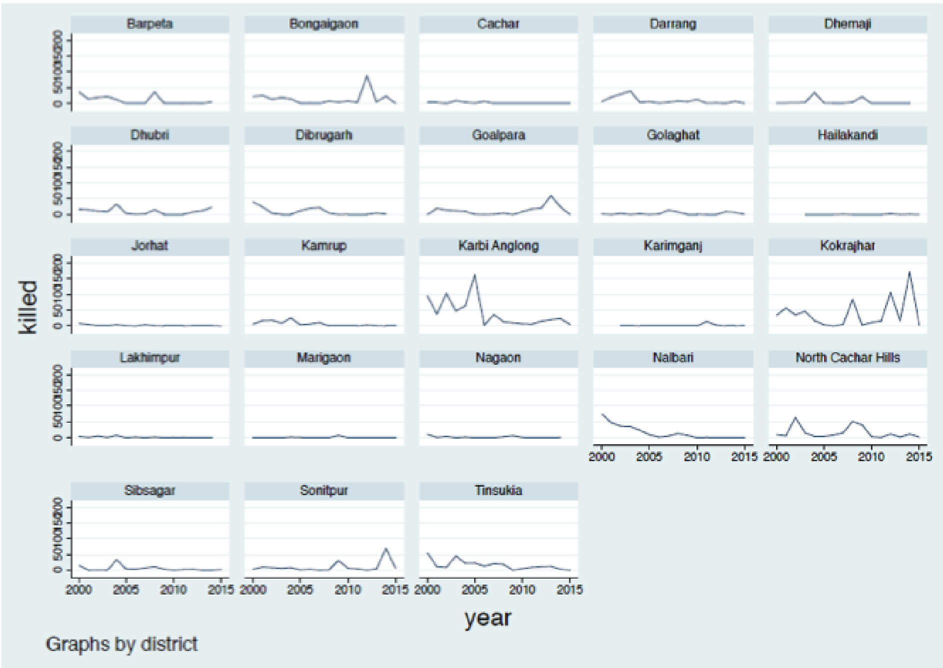


Figure 3.7.2: Total civilians killed in the Assam insurgency, by district (2000-15). Source: Author's compilation based on SATP data.

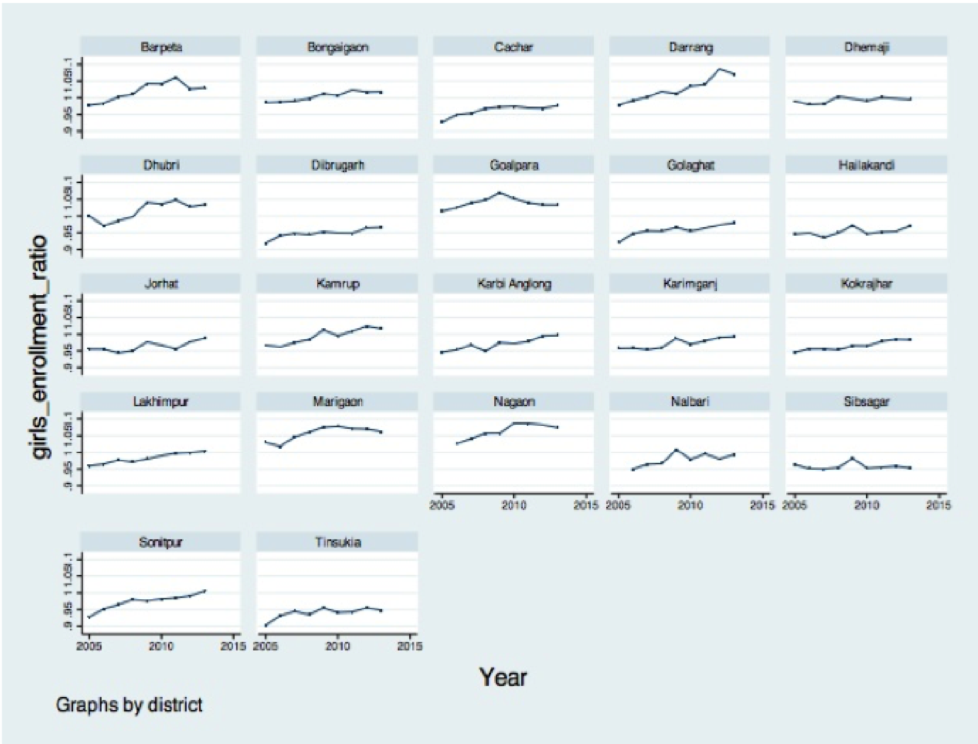


Figure 3.7.3: Girl's enrolment ratio using DISE surveys, by district (2005-14). Source: Author's compilation based on DISE data.

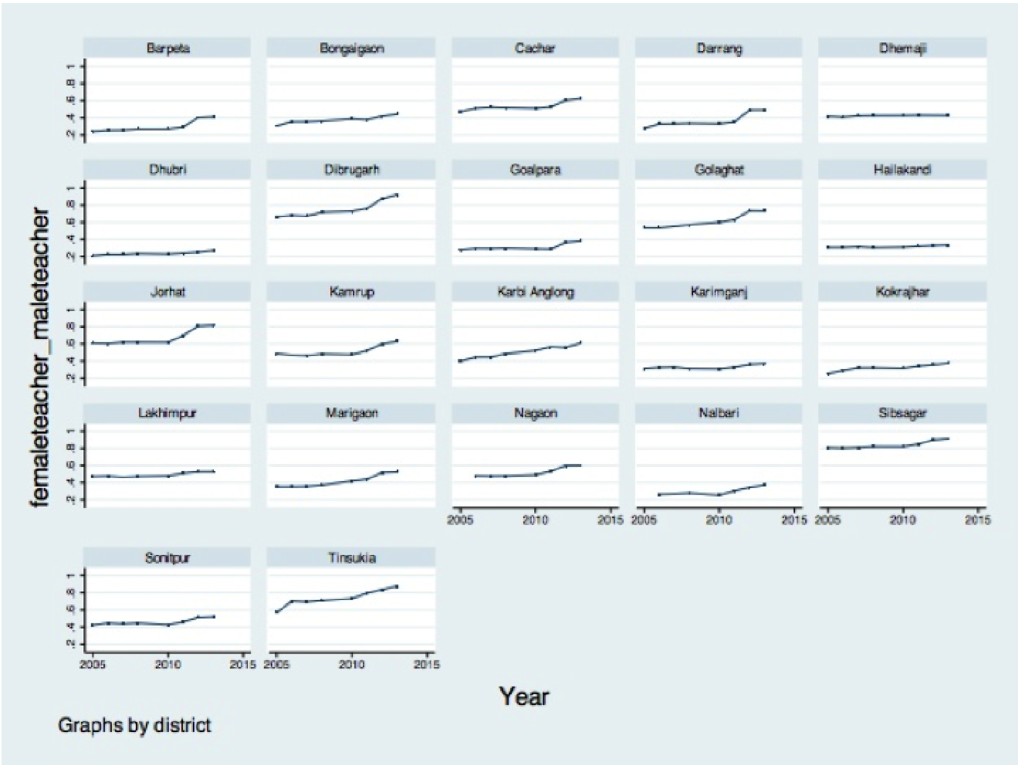


Figure 3.7.4: Ratio of female to male teachers using DISE surveys, by district (2005-14). Source: Author's compilation based on DISE data.

3.8 Tables

Table 3.8.1: Conflict in Assam 2000-2014 using SATP data

Year	Mean annual killed or injured per district
2000	26
2001	24
2002	22
2003	17
2004	34
2005	15
2006	11
2007	19
2008	26
2009	17
2010	7
2011	5
2012	18
2013	17
2014	24

Note: Means for every year are calculated per district in Assam. Killed or injured are total civilians killed or injured in insurgency-related incidents. Source: Author's compilation based on SATP data.

Table 3.8.2: Conflict in Assamese districts (2000) using SATP data

District	Mean annual killed or injured	Mean annual killed
Barpeta	15	13
Bongaigaon	27	15
Cachar	2	2
Darrang	17	9
Dhemaji	23	7
Dhubri	21	10
Dibrugarh	15	9
Goalpara	22	12
Golaghat	5	3
Hailakandi	1	0
Jorhat	4	2
Kamrup	14	6
Karbi Anglong	61	41
Karimganj	2	2
Kokrajhar	49	38
Lakhimpur	3	2
Marigaon	8	1
Nagaon	8	2
Nalbari	21	16
North Cachar Hills	24	15
Sibsagar	13	5
Sonitpur	17	10
Tinsukia	28	16

Source: Author's compilation based on SATP data

Table 3.8.3: Differences in Schooling for High vs Low Conflict districts for the baseline year 2005-2006

Variables	Mean for Low Conflict Districts	Mean for High Conflict Districts	Differences in Means
	(1)	(2)	(3)= (2)-(1)
girl's enrolment ratio	0.500	0.500	-0.000
			(0.00161)
total number of girls enroled	44.15	47.464	3.314
			(2.420)
total number of kids enroled	88.58	95.080	6.500
			(4.604)
total number of boys enroled	44.43	47.616	3.186
			(2.198)
total number of female teachers	1.381	0.971	-0.410^{***}
			(0.101)
total number of male teachers	2.727	2.678	-0.0491
			(0.0811)
number of library books available	52.53	57.928	5.398
			(5.234)
total number of computers	0.292	0.237	-0.0549^{**}
			(0.0256)
amount received as school development	4442.3	3979.1	-463.2
			(300.8)
amount spent as school development	4279.6	3786.5	-493.1^{*}
			(278.6)
amount of teaching and learning grant	1187.0	1083.8	-103.2^{**}
			(44.50)
amount received for purchasing teaching	1213.0	1096.5	-116.5^{**}
			(45.75)

Notes: School development grant is the amount received under Sarva Siksha Abhiyan, an educational program run by the central government. High conflict districts are defined as those districts that had a higher than or equal to the median annual killings on average between 2005-2014. Standard errors for the differences in the means for parentheses. Errors are clustered at the block level. *p< 0.1, ** p <0.05, *** p <0.01. Source: Author's compilation based on SATP data

Table 3.8.4: Effect of conflict on total kids enroled in school

Dep var: Total kids enroled in school	(1)	(2)	(3)	(4)
Killed	-0.0679*	-0.0365**	-0.0670	-0.0303*
	(0.0395)	(0.0177)	(0.0449)	(0.0178)
District Fixed Effects	no	yes	no	yes
Year Fixed Effects	no	no	yes	yes
Observations	518301	518301	518301	518301
adj.R-sq	0.000	0.050	0.005	0.055

Notes: Standard errors in parentheses clustered at the block level. District fixed-effects are dummies for 23 districts. Year fixed effects include dummies for years from 2005 till 2014. Killed measures total number of civilians killed in a district-year in Assam in insurgency-related events. Injured measures total number of civilians injured in a district-year during the insurgency (2005-2014). *** p¡0.01, ** p¡0.05, * p¡0.1.

Table 3.8.5: Effect of conflict on Total kids enroled in school

Dep var: Total kids enroled in school	
Killed	0.00943***
	(0.0035)
Observations	518,301
School Fixed Effects	Yes
Year Fixed Effects	Yes
R sq	.002

Notes: Standard errors in parentheses clustered at the block level. District fixed-effects are dummies for districts. Year fixed effects include dummies for years from 2005 till 2014. Killed measures total number of civilians killed in a district-year in Assam in insurgency-related events. Injured measures total number of civilians injured in a district-year during the insurgency (2005-2014). *** p¡0.01, ** p¡0.05, * p¡0.1.

Table 3.8.6: Impact of conflict on total girls enroled in school

Dep var: Total girls enroled in school	(1)	(2)	(3)	(4)	(5)
Killed	-0.0578**	-0.0557**	-0.0161*	-0.0161**	0.0151
	(0.0223)	(0.0234)	(0.00943)	(0.00732)	(0.00768)
Year Fixed Effects	no	yes	yes	yes	yes
District Fixed Effects	no	no	yes	no	no
Block Fixed Effects	no	no	no	yes	no
School Fixed Effects	no	no	no	no	yes
N	518301	518301	518301	518301	518301
adj. R-sq	0.000	0.005	0.054	0.088	0.002

Notes: District fixed effects are dummies for 22 districts. Year fixed effects include dummies for years from 2005 till 2013. Killed measures total number of civilians killed in a district-year in Assam in insurgency -related events. Standard Errors are clustered at the block level. $***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data.

Table 3.8.7: Impact of Conflict on total boys enroled in school

Dep var: Total boys enroled in school	(1)	(2)	(3)	(4)	(5)
Killed	-0.0457**	-0.0449**	-0.0140	-0.0152**	0.00531
	(0.0211)	(0.0224)	(0.00868)	(0.00747)	(0.00759)
Year Fixed Effects	no	yes	yes	yes	yes
District Fixed Effects	no	no	yes	no	no
Block Fixed Effects	no	no	no	yes	no
School Fixed Effects	no	no	no	no	yes
N	518301	518301	518301	518301	518301
adj. R-sq	0.000	0.003	0.042	0.071	0.004

Notes: District fixed effects are dummies for 22 districts. Year fixed effects include dummies for years from 2005 till 2013. Killed measures total number of civilians killed in a district-year in Assam in insurgency-related events. Standard Errors are clustered at the block level. $***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data.

Table 3.8.8: Impact of conflict on girl's school enrolment

Dep Var: Girl's enrolment in school	(1)	(2)	(3)	(4)	(5)
Killed	-0.0000480**	-0.0000463**	-0.0000351**	-0.0000322**	-0.0000339**
	(0.0000229)	(0.0000230)	(0.0000140)	(0.0000145)	(0.0000144)
Year Fixed Effects	no	yes	yes	yes	yes
District Fixed Effects	no	no	yes	no	no
Block Fixed Effects	no	no	no	yes	no
School Fixed Effects	no	no	no	no	yes
N	514,164	514,164	514,164	514,164	514,164
adj. R-sq	0.000	0.001	0.003	0.005	0.005

Notes: District fixed effects are dummies for 22 districts. Year fixed effects include dummies for years from 2005 till 2013. Killed measures total number of civilians killed in a district-year in Assam in insurgency-related events. Injured measures total number of civilians injured in a district-year during the insurgency (2005-14). Errors are clustered at the block level in (1)-(4) and at the district level in (5). $***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data.

Table 3.8.9: Impact of conflict on school enrolment using log specification

Dep var: Log Girl's enrolment in school	(1)	(2)	(3)	(4)	(5)
Log Killed	-0.00106	-0.00150*	-0.00102*	-0.00127**	-0.00131**
	(0.000820)	(0.000855)	(0.000586)	(0.000610)	(0.000615)
Year Fixed Effects	no	yes	yes	yes	yes
District Fixed Effects	no	no	yes	no	no
Block Fixed Effects	no	no	no	yes	no
School Fixed Effects	no	no	no	no	yes
N	316118	316118	316118	316118	316118
adj. R-sq	0.000	0.002	0.003	0.006	0.005

Notes: District fixed effects are dummies for 22 districts. Year fixed effects include dummies for years from 2005 till 2014. Log of Killed measures the logarithm of total number of civilians killed in a district-year in Assam in insurgency -related events. Standard Errors are clustered at the block level. Block fixed efefcts are dummies for 149 blocks. $***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data.

Table 3.8.10: Impact of conflict on girl's enrolment by school management

Dep var: Girls enrolment ratio	Dept. of education	Tribal and Social welfare dept	Local body
	(1)	(2)	(3)
Killed	0.00000792	0.000727	-0.000312**
	(0.0000169)	(0.00361)	(0.000126)
Year Fixed Effects	yes	yes	yes
District Fixed Effects	yes	yes	yes
N	374527	1476	4673
adj. R-sq	0.003	0.042	0.029

Notes: Standard errors in parentheses clustered at the block level. District fixed-effect are dummies for 22 districts. Year fixed effects include dummies for years from 2005 till 2013. Killed measures total number of civilians killed in a district-year in Assam in insurgency related events (2005-2014). $***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data.

Table 3.8.11: Impact of conflict on girls enrolment by school management

Dep var: Girls enrolment ratio			
Local body	private aided body	private unaided body	madrassa
	(1)	(2)	(3)
Killed	0.0000995	-0.000336***	-0.0000381
	(0.0000642)	(0.000107)	(0.000292)
Year Fixed Effects	yes	yes	yes
District Fixed Effects	yes	yes	yes
N	44911	26838	774
adj. R-sq	0.015	0.051	0.273

Notes: Standard errors in parentheses clustered at the block level. District fixed-effect are dummies for 22 districts. Year fixed effects include dummies for years from 2005 till 2013. Killed measures total number of civilians killed in a district-year in Assam in insurgency related events (2005-2014). $***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data.

Table 3.8.12: Heterogeneous effect of conflict on girl's enrolment by class

	class 1	class 2	class 3	class 4	class 5	class 6	class 7	class 8
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Killed	-0.0000272	-0.0000328	-0.0000624**	-0.0000730**	-0.00000139	-0.0000278	-0.0000294	0.0000561**
	(0.0000192)	(0.0000256)	(0.0000250)	(0.0000323)	(0.0000393)	(0.0000162)	(0.0000208)	(0.0000229)
Year Fixed Effects	yes	yes	yes	yes	yes	yes	yes	yes
School Fixed Effects	yes	yes	yes	yes	yes	yes	yes	yes
N	396838	388191	377112	368021	204181	118917	117909	65935
adj. R-sq	0.000	0.000	0.001	0.001	0.001	0.005	0.004	0.002

Notes: Dependent Variable is girl's enrolment rate. Errors are clustered at the district year level. Year fixed effects include dummies for years from 2005 till 2013. Killed measures total number of civilians killed in a district-year in Assam in insurgency related events. $***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data.

Table 3.8.13: Impact of conflict on girl's school enrolment in rural schools

Dep Var:Girl's enrolment ratio in school	(1)	(2)	(3)	(4)
Killed	-0.0000478**	-0.0000336**	-0.0000304**	-0.0000347**
	(0.0000233)	(0.0000137)	(0.0000144)	(0.0000142)
Year Fixed Effects	yes	yes	yes	yes
District Fixed Effects	no	yes	no	no
Block Fixed Effects	no	no	yes	no
School Fixed Effects	no	no	no	yes
N	486305	486305	486305	486305
adj. R-sq	0.001	0.003	0.006	0.005

Notes: Standard errors in parentheses clustered at the block level. District fixed effects are dummies for 22 districts. Year fixed effects include dummies for years from 2005 till 2014. Killed measures total number of civilians killed in a district-year in Assam in insurgency related events. Injured measures total number of civilians killed in a district-year in Assam in insurgency related events. $***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data.

Table 3.8.14: Impact of conflict on girl's school enrolment in urban schools

Dep Var: Girl's enrolment in school	(1)	(2)	(3)	(4)
Killed	-0.0000219	-0.0000640	-0.0000502	-0.0000313
	(0.0000608)	(0.0000510)	(0.0000526)	(0.0000272)
Year Fixed Effects	yes	yes	yes	yes
District Fixed Effects	no	yes	no	no
Block Fixed Effects	no	no	yes	no
School Fixed Effects	no	no	no	yes
N	27855	27855	27855	27855
adj. R-sq	0.000	0.003	0.009	0.005

Notes: Standard errors in parentheses clustered at the block level. District fixed effects are dummies for 22 districts. Year fixed effects include dummies for years from 2005 till 2014. Killed measures total number of civilians killed in a district-year in Assam in insurgency related events. Injured measures total number of civilians killed in a district-year in Assam in insurgency related events. $***p < 0.01$, $**p < 0.05$ $*p < 0.1$.
Source: Author's compilation based on SATP and DISE data.

Table 3.8.15: Heterogeneity of Impact of Conflict by GDP per capita

Dep var: Girls enrolment ratio in school	(1)	(2)	(3)	(4)	(5)
Killed	-0.0000845	-0.000155*	-0.000189***	-0.000149***	-0.0000961***
	(0.0000858)	(0.0000869)	(0.0000314)	(0.0000547)	(0.0000279)
High GDP	-0.00483***	-0.00532***			
	(0.00159)	(0.00160)			
Killed*High GDP	0.0000657	0.000153*	0.000184***	0.000141**	0.0000742**
	(0.0000894)	(0.0000901)	(0.0000330)	(0.0000560)	(0.0000296)
Year Fixed Effects	no	yes	yes	yes	yes
District Fixed Effects	no	no	yes	no	no
Block Fixed Effects	no	no	no	yes	no
School Fixed Effects	no	no	no	no	yes
N	514164	514164	514164	514164	514164
adj. R-sq	0.000	0.001	0.003	0.005	0.005

Notes: High GDP is a dummy variable that takes value 1 if that district in Assam had higher than median level of GDP per capita in 2005 or greater than Rs 17048. GDP data was obtained from Directorate of Economics and Statistics, Assam. District fixed effects are dummies for 22 districts. Year fixed effects include dummies for years from 2005 till 2014. Killed measures total number of civilians killed in a district-year in Assam in insurgency related events. lnjured measures total number of civilians killed in a district-year in Assam in insurgency related events. Errors are clustered at the block level (1)-(4) and at district level in (5).$***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data and Directorate of Economics and Statistics, Assam.

Table 3.8.16: Heterogeneity of Impact of Conflict by safety of girls

Dep var: Girls enrolment ratio in school	(1)	(2)	(3)	(4)	(5)
Killed	-0.0000140	-0.00000349	-0.0000260*	-0.0000267*	-0.0000234*
	(0.0000233)	(0.0000233)	(0.0000136)	(0.0000136)	(0.0000114)
High Rape	0.00616***	0.00652***			
	(0.00152)	(0.00153)			
Killed*High Rape	-0.0000742	-0.000114	-0.0000454	-0.0000265	-0.0000545
	(0.0000749)	(0.0000692)	(0.0000368)	(0.0000444)	(0.0000350)
Year Fixed Effects	no	yes	yes	yes	yes
District Fixed Effects	no	no	yes	no	no
Block Fixed Effects	no	no	no	yes	no
School Fixed Effects	no	no	no	no	yes
N	514164	514164	514164	514164	514164
adj. R-sq	0.000	0.001	0.003	0.005	0.005

Notes: High Rape is a dummy variable that takes value 1 if that district in Assam had higher than median level (>0.04)of rape per capita ('000 population) in 2005. Data on rape was obtained from Ministry of Home Affairs, New Delhi and population in 2005 from Directorate of Economics and Statistics, Assam. District fixed effects are dummies for 22 districts. Year fixed effects include dummies for years from 2005 till 2013. Killed measures total number of civilians killed in a district-year in Assam in insurgency related events. lnjured measures total number of civilians killed in a district-year in Assam in insurgency related events. Errors are clustered at the block level (1)-(4) and at district level in (5).$***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data and Directorate of Economics and Statistics, Assam, Ministry of Home Affairs, Delhi.

Table 3.8.17: Robustness Checks

	Excluding 2013		**Excluding districts where school attack took place**		**Excluding international border districts**	
Dep Vars: Girl's enrolment ratio	(1)	(2)	(3)	(4)	(5)	(6)
Killed	-0.0000283***	-0.0000285**	-0.0000219**	-0.0000229**	-0.0000577***	-0.0000552***
	(0.0000137)	(0.0000109)	(0.0000123)	(0.0000106)	(0.0000174)	(0.0000138)
Year Fixed Effects	yes	yes	yes	yes	yes	yes
Block Fixed Effects	yes	no	yes	no	yes	no
School Fixed Effects	no	yes	no	yes	no	yes
N	456349	456349	438197	438197	411269	411269
adj. R-sq	0.005	0.004	0.005	0.005.	0.005	0.004

Notes: Year Fixed Effects include dummies for years from 2005 till 2014 and Block Fixed Effects are dummies for 149 blocks. Killed measures total number of civilians killed in a district-year in Assam insurgency related events (2005-2014). $***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data.

Table 3.8.18: Using killed per capita as the main independent variable

Dep var: Girl's Enrolment in school	(1)	(2)	(3)	(4)	(5)
Killed	-0.0332*	-0.0337*	-0.0285**	-0.0260**	-0.0291**
	(0.0199)	(0.0200)	(0.0117)	(0.0122)	(0.0123)
Year Fixed Effects	no	yes	yes	yes	yes
District Fixed Effects	no	no	yes	no	no
Block Fixed Effects	no	no	no	yes	no
School Fixed Effects	no	no	no	no	yes
N	514164	514164	514164	514164	514164
adj. R-sq	0.000	0.001	0.003	0.005	0.005

Notes: District fixed effects are dummies for 22 districts. Year fixed effects include dummies for years from 2005 till 2014. Killed per capita measures total number of civilians killed in district-year in Assam in insurgency related events divided by the population (in '000) in that district in 2005 as obtained from Directorate of Economics and Statistics, Assam. Errors are clustered at the block level in (1)-(4) and at the district level in (5). $***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data.

Table 3.8.19: Using Lagged conflict as the main independent variable

Dep var: Girl's Enrolment Ratio in School	(1)	(2)	(3)	(4)	(5)
Lag Killed	-0.0000275	-0.0000377*	-0.0000260**	-0.0000281**	-0.0000262***
	(0.0000185)	(0.0000197)	(0.0000110)	(0.0000115)	(0.00000724)
Year Fixed Effects	no	yes	yes	yes	yes
District Fixed Effects	no	no	yes	no	no
Block Fixed Effects	no	no	no	yes	no
School Fixed Effects	no	no	no	no	yes
N	428300	428300	428300	428300	428300
adj. R-sq	0.000	0.001	0.003	0.006	0.004

Notes: District fixed effects are dummies for 22 districts. Year fixed effects include dummies for years from 2005 till 2013. Killed per capita measures total number of civilians killed in district-year in Assam in insurgency related events divided by the population (in '000) in that district in 2005 as obtained from Directorate of Economics and Statistics, Assam. Errors are clustered at the block level in (1)-(4) and at the district level in (5).$***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data.

Table 3.8.20: Alternative definition of Conflict

Dep var: Girl's enrolment ratio	(1)	(2)	(3)
Civilians killed	-0.0000649***		
	(0.0000180)		
Security Force killed		-0.000116	
		(0.000122)	
Extremists Killed			-0.00128**
			(0.000573)
Year Fixed Effects	yes	yes	yes
School Fixed Effects	yes	yes	yes
N	348186	348186	348186
adj. R-sq	0.002	0.002	0.002

Notes: Data on ULFA bombing was obtained from Bureau of Investigation Special Branch, Department of Home Affairs, Assam. Civilians killed measures the civilians killed in ULFA bombing incidents in Assam by district and year. Security forces killed measures security forces killed in bombing incidents in Assam by district and year. Extremists killed measures extremists killed in bombing incidents in Assam by district and year. $***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on Bureau of Investigation Special Branch, Department of Home Affairs, Assam, and DISE data.

Table 3.8.21: Impact of conflict on girl's school enrolment

Dep var: Girl's Enrolment ratio in school	(1)	(2)	(3)	(4)	(5)
Killed or Injured	-0.0000109	-0.000000895	-0.00000696	-0.00000219	-0.0000174*
	(0.0000177)	(0.0000184)	(0.0000112)	(0.0000119)	(0.00000930)
Year Fixed Effects	no	yes	yes	yes	yes
District Fixed Effects	no	no	yes	no	no
Block Fixed Effects	no	no	no	yes	no
School Fixed Effects	no	no	no	no	yes
N	5141,64	5141,64	514,164	514,164	514,164
adj. R-sq	0.000	0.001	0.003	0.005	0.005

Notes: District fixed effects are dummies for 22 districts. Year fixed effects include dummies for years from 2005 till 2013. Killed measures total number of civilians killed in district-year in Assam in insurgency-related events. Injured measures total number of civilians injured in a district-year during the insurgency (2005-2014). Errors are clustered at the block level in (1)-(4) and at the district level in (5).$***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data.

Table 3.8.22: Impact of conflict on girl's school enrolment

	Girls Enrolment
killed	-2.59e-05**
	(1.08e-05)
Observations	252,790
School Fixed Effect	Yes
Year Fixed Effect	Yes
Number of schoolcode	28,176
R-squared	0.007

Notes: The analysis is done on the balanced panel of schools over 9 years. The standard errors are clustered at school level. The sample consists of about 48 per cent of the full sample.$***p < 0.01$, $**p < 0.05$ $*p < 0.1$. Source: Author's compilation based on SATP and DISE data.

Table 3.8.23: Impact of Conflict on School Enrolment with Resource Effects

Dep var: Girl's enrolment ratio	(1)	(2)	(3)	(4)	(5)	(6)
Killed	−0.0000331**	-0.0000333**	-0.0000296**	-0.0000297**	-0.0000535***	-0.0000565**
	(0.0000151)	(0.0000146)	(0.0000141)	(0.0000149)	(0.0000157)	(0.0000225)
Teaching grant pp(TL)	0.0000178**					
	(0.00000653					
TL*Killed	7.24e-08					
	(0.000000563)					
School Dev. Grant pp (SDG)		0.00000107**				
		(0.000000471)				
SDG*killed		1.60e-08				
		(1.93e-08)				
No of computers pp			0.00272			
			(0.0210)			
No of computers pp*killed			-0.00159			
			(0.00181)			
No of Library Books pp (NB)				0.000275		
				(0.000168)		
NB*killed				-0.00000343		
				(0.000104)		
Professionally Qualified Teachers (Q)					0.0200*	
					(0.0109)	
Q*killed					0.00157***	
					(0.000525)	
No of female teachers pp						0.0740***
						(0.0200)
No of female teachers pp*killed						0.00194*
						(0.00108)
District Fixed Effects	yes	yes	yes	yes	yes	yes
Year Fixed Effects	yes	yes	yes	yes	yes	yes
N	514021	510623	514161	514162	462307	462307
adj. R-sq	0.005	0.005	0.005	0.005	0.005	0.007

Notes: Standard errors in parentheses clustered at the block level. District fixed effects are dummies for 22 districts. Year fixed effects include dummies for years from 2005 till 2013. Killed measures total number of civilians killed in a district year in Assam in insurgency related events.$***p < 0.01$, $**p < 0.05$ $*p < 0.1$. School development grant is the amount received under Sarva Siksha Aviyan. Source: Author's compilation based on SATP and DISE data.

3.9 Appendix

Table A1: Cross-state literacy gaps in 2001

	GDI in 2001	Literacy gap in 2001	Literacy gap in 2011
A. Pradesh	0.48	20.33	14.12
Assam	0.49	8.64	11.54
Manipur	0.58	18.17	13.32
Meghalaya	0.51	5.73	3.39
Mizoram	0.67	4.56	4.32
Nagaland	0.42	9.19	6.60
Sikkim	0.59	15.60	10.86
Tripura	0.56	16.10	9.03
India	0.54	21.60	16.68

Source: Planning Commision (2001); www.indiastat.com

Table A2: Cross State literacy in classes IX-XII

	Class IX-X (14-15 years)			**Class XI-XII (16-17 years)**		
	Boys	Girls	Gap	Boys	Girls	Gap
Ar Pradesh	73.3	67.9	5.4	49.1	45.7	3.4
Assam	52.0	46.9	5.1	18.2	14.6	3.6
Manipur	83.5	80.1	3.4	39.0	32.1	6.9
Meghalaya	49.0	49.9	-0.9	13.7	17.3	-3.6
Mizoram	75.4	78.3	-2.9	41.2	40.2	1.0
Nagaland	27.4	29.5	-2.1	18.3	16.7	1.6
Sikkim	44.9	50.3	-5.4	27.6	29.5	-1.9
Tripura	73.0	73.3	-0.3	31.9	25.0	6.9
India	69.0	60.8	8.2	42.2	36.1	6.1

Bibliography

Acemoglu, D. and S. Johnson (2005). Unbundling institutions. *Journal of political Economy 113*(5), 949–995.

Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review 91*(5), 1369–1401.

Akresh, R. (2008). *Armed conflict and schooling: Evidence from the 1994 Rwandan genocide*, Volume 3516. World Bank Publications.

Angrist, J., D. Lang, and P. Oreopoulos (2009). Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics 1*(1), 136–163.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Ashraf, N., O. Bandiera, and K. Jack (2012). No margin, no mission. *A Field Experiment on Incentives for Pro-Social Tasks, CEPR Discussion Papers 8834*.

Attanasio, O. P., C. Meghir, and A. Santiago (2011). Education choices in mexico: using a structural model and a randomized experiment to evaluate progresa. *The Review of Economic Studies 79*(1), 37–66.

Baden-Powell, B. H. (1892). *The Land-systems of British India: book III. The system of village or Mahál settlements*, Volume 2. Clarendon Press.

Bagde, S., D. Epple, L. Taylor, et al. (2016). Does affirmative action work? caste, gender, college quality, and academic success in india. *American Economic Review 106*(6), 1495–1521.

Baliamoune-Lutz, M. and M. McGillivray (2015). The impact of gender inequality in education on income in africa and the middle east. *Economic Modelling 47*, 1–11.

Banerjee, A. and L. Iyer (2005). History, institutions, and economic performance: The legacy of colonial land tenure systems in india. *American economic review 95*(4), 1190–1213.

Banerjee, A. V., S. Cole, E. Duflo, and L. Linden (2007). Remedying education: Evidence from two randomized experiments in india. *The Quarterly Journal of Economics 122*(3), 1235–1264.

Barrios, T., R. Diamond, G. W. Imbens, and M. Kolesár (2012). Clustering, spatial correlations, and randomization inference. *Journal of the American Statistical Association 107*(498), 578–591.

Baruah, S. (2003). Citizens and denizens: Ethnicity, homelands, and the crisis of displacement in northeast india. *Journal of Refugee Studies 16*(1), 44–66.

Bertrand, M., R. Hanna, and S. Mullainathan (2010). Affirmative action in education: Evidence from engineering college admissions in india. *Journal of Public Economics 94*(1), 16–29.

Blattman, C. and E. Miguel (2010). Civil war. *Journal of Economic literature 48*(1), 3–57.

Bracha, A., A. Cohen, and L. Conell-Price (2015). Affirmative action and stereotype threat. *SSRN Electronic Journal*.

Breierova, L. and E. Duflo (2004). The impact of education on fertility and child mortality: Do fathers really matter less than mothers? Technical report, National bureau of economic research.

Bruhn, M. and D. McKenzie (2009). In pursuit of balance: Randomization in practice in development field experiments. *American economic journal: applied economics 1*(4), 200–232.

Burgess, S. M., R. D. Metcalfe, and S. Sadoff (2016). Understanding the response to financial and non-financial incentives in education: Field experimental evidence using high-stakes assessments.

Burns, J., L. Corno, and E. La Ferrara (2015). Interaction, prejudice and performance. evidence from south africa. Technical report, Working paper.

Bursztyn, L., F. Ederer, B. Ferman, and N. Yuchtman (2014). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica 82*(4), 1273–1301.

Calsamiglia, C., J. Franke, and P. Rey-Biel (2013). The incentive effects of affirmative action in a real-effort tournament. *Journal of Public Economics 98*, 15–31.

Cameron, A. C. and D. L. Miller (2015). A practitioners guide to cluster-robust inference. *Journal of Human Resources 50*(2), 317–372.

Carneiro, P., J. J. Heckman, and E. J. Vytlacil (2011). Estimating marginal returns to education. *The American economic review 101*(6), 2754–2781.

Carrell, S. E., R. L. Fullerton, and J. E. West (2009). Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics 27*(3), 439–464.

Carroll, L. (1983). Law, custom, and statutory social reform: the hindu widows' remarriage act of 1856. *The Indian Economic & Social History Review 20*(4), 363–388.

Cason, T. N., W. A. Masters, and R. M. Sheremeta (2010). Entry into winner-take-all and proportional-prize contests: An experimental study. *Journal of Public Economics 94*(9), 604–611.

Chamarbagwala, R. and H. E. Morán (2011). The human capital consequences of civil war: Evidence from guatemala. *Journal of Development Economics 94*(1), 41–61.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2011). The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood. Technical report, National Bureau of Economic Research.

Chitnis, V. and D. Wright (2007). Legacy of colonialism: Law and women's rights in india. *Wash. & Lee L. Rev. 64*, 1315.

Dechenaux, E., D. Kovenock, and R. M. Sheremeta (2015). A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics 18*(4), 609–669.

Dell, M., N. Lane, and P. Querubin (2015). State capacity, local governance, and economic development in vietnam. *NBER Working Paper*, 1–40.

Desai, S. and V. Kulkarni (2008). Changing educational inequalities in india in the context of affirmative action. *Demography 45*(2), 245–270.

Dobbie, W. and R. G. Fryer Jr (2009). Are high quality schools enough to close the achievement gap? evidence from a social experiment in harlem. Technical report, National Bureau of Economic Research.

Drago, R. and G. T. Garvey (1998). Incentives for helping on the job: Theory and evidence. *Journal of labor Economics 16*(1), 1–25.

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature 50*(4), 1051–79.

Duflo, E., R. Glennerster, and M. Kremer (2007). Using randomization in development economics research: A toolkit. *Handbook of development economics 4*, 3895–3962.

Fisch, J. (2000). Humanitarian achievement or administrative necessity? lord william bentinck and the abolition of sati in 1829. *Journal of Asian History 34*(2), 109–134.

Fisher, R. A. (1937). *The design of experiments.* Oliver And Boyd; Edinburgh; London.

Fryer, R. G. and G. C. Loury (2005). Affirmative action in winner-take-all markets. *The Journal of Economic Inequality 3*(3), 263–280.

Fryer Jr, R. G. (2014). Injecting charter school best practices into traditional public schools: Evidence from field experiments. *The Quarterly Journal of Economics 129*(3), 1355–1407.

Fryer Jr, R. G., S. D. Levitt, and J. A. List (2015). Parental incentives and early childhood achievement: a field experiment in chicago heights. Technical report, National Bureau of Economic Research.

Gächter, S., C. Starmer, and F. Tufano (2015). Measuring the closeness of relationships: a comprehensive evaluation of the'inclusion of the other in the self'scale. *PloS one 10*(6), e0129478.

Glewwe, P. and M. Kremer (2006). Schools, teachers, and education outcomes in developing countries. *Handbook of the Economics of Education 2*, 945–1017.

Goswami, S. (2001). Ethnic conflict in assam. *The Indian Journal of Political Science*, 123–137.

Grant, M. J. and J. R. Behrman (2010). Gender gaps in educational attainment in less developed countries. *Population and Development Review 36*(1), 71–89.

Grey, D. J. (2011). Gender, religion, and infanticide in colonial india, 1870–1906. *Victorian Review 37*(2), 107–120.

Grey, D. J. (2013). Creating the "problem hindu": Sati, thuggee and female infanticide in india, 1800–60. *Gender & History 25*(3), 498–510.

Hanna, R. N. and L. L. Linden (2012). Discrimination in grading. *American Economic Journal: Economic Policy 4*(4), 146–168.

Hanushek, E. A., L. Woessmann, E. A. Jamison, and D. T. Jamison (2008). Education and economic growth. *Education Next 8*(2).

Harrison, D. A., D. A. Kravitz, D. M. Mayer, L. M. Leslie, and D. Lev-Arey (2006). Understanding attitudes toward affirmative action programs in employment: Summary and meta-analysis of 35 years of research.

Hartog, J. and H. Oosterbeek (1988). Education, allocation and earnings in the netherlands: 0verschooling? *Economics of Education Review 7*(2), 185–194.

Hatekar, N., R. Mathur, and P. Rege (2007). 'legislating'social change: Strange case of the sarda act. *Economic and Political Weekly*, 145–153.

Hickman, B. R. (2013). Pre-college human capital investment and affirmative action: a structural policy analysis of us college admissions. *University of Chicago. Unpublished*.

Hoxby, C. (2000). Peer effects in the classroom: Learning from gender and race variation. Technical report, National Bureau of Economic Research.

Iyer, L. (2010). Direct versus indirect colonial rule in india: Long-term consequences. *The Review of Economics and Statistics 92*(4), 693–713.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics 125*(2), 515–548.

Kecmanovic, M. (2013). The short-run effects of the croatian war on education, employment, and earnings. *Journal of Conflict Resolution 57*(6), 991–1010.

Klasen, S. (2002). Low schooling for girls, slower growth for all? cross-country evidence on the effect of gender inequality in education on economic development. *The World Bank Economic Review 16*(3), 345–373.

Kosfeld, M., S. Neckermann, and X. Yang (2014). Knowing that you matter, matters! the interplay of meaning, monetary incentives, and worker recognition.

Kremer, M., E. Miguel, and R. Thornton (2009). Incentives to learn. *The Review of Economics and Statistics 91*(3), 437–456.

Krishna, K. and A. Tarasov (2016). Affirmative action: One size does not fit all. *American Economic Journal: Microeconomics 8*(2), 215–252.

Krueger, A. B. (1999). Experimental estimates of education production functions. *The quarterly journal of economics 114*(2), 497–532.

Lavy, V., M. D. Paserman, and A. Schlosser (2012). Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *The Economic Journal 122*(559), 208–237.

Lavy, V. and E. Sand (2016). The effect of social networks on student?s academic and non-cognitive behavioral outcomes: Evidence from conditional random assignment of friends in school. *Unpublished manuscript, University of Warwick*.

Lazear, E. P. and S. Rosen (1981). Rank-order tournaments as optimum labor contracts. *Journal of political Economy 89*(5), 841–864.

Levitt, S. D., J. A. List, S. Neckermann, and S. Sadoff (2011). The impact of short-term incentives on student performance. *Unpublished mimeo, University of Chicago*.

List, J., D. Van Soest, J. Stoop, and H. Zhou (2014). On the role of group size in tournaments: Theory and evidence from lab and field experiments. Technical report, National Bureau of Economic Research.

Loury, G. C. (1979). Market structure and innovation. *The quarterly journal of economics*, 395–410.

Lynch, F. R. (1989). *Invisible victims: White males and the crisis of affirmative action.* Greenwood Press.

Mahanta, B. and P. Nayak (2013). Gender inequality in north east india.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies 60*(3), 531–542.

Michalopoulos, S. and E. Papaioannou (2013). Pre-colonial ethnic institutions and contemporary african development. *Econometrica 81*(1), 113–152.

Milgrom, P. and J. Roberts (1988). An economic approach to influence activities in organizations. *American Journal of sociology 94*, S154–S179.

Mitra, S. K. (1995). The rational politics of cultural nationalism: subnational movements of south asia in comparative perspective. *British Journal of Political Science 25*(1), 57–77.

Murshed, S. M. (2002). Conflict, civil war and underdevelopment: an introduction.

Murthy, K. D. (2006). Rajkumar and kannada nationalism. *Economic and Political Weekly*, 1834–1835.

Nalebuff, B. J. and J. E. Stiglitz (1983). Prizes and incentives: towards a general theory of compensation and competition. *The Bell Journal of Economics*, 21–43.

Peers, D. M. (2013). *India under colonial rule: 1700-1885*. Routledge.

Prendergast, C. (1999). The provision of incentives in firms. *Journal of economic literature 37*(1), 7–63.

Quisumbing, A. R. (1994). Intergenerational transfers in philippine rice villages: Gender differences in traditional inheritance customs. *Journal of Development Economics 43*(2), 167–195.

Ramusack, B. N. (2004). *The Indian princes and their states*. Cambridge University Press.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly journal of economics 116*(2), 681–704.

Schotter, A. and K. Weigelt (1992). Asymmetric tournaments, equal opportunity laws, and affirmative action: Some experimental results. *The Quarterly Journal of Economics 107*(2), 511–539.

Shemyakina, O. (2011). The effect of armed conflict on accumulation of schooling: Results from tajikistan. *Journal of Development Economics 95*(2), 186–200.

Sheremeta, R. M. (2015). Behavior in group contests: A review of experimental research. *Journal of Economic Surveys*.

Shteynberg, G., L. M. Leslie, A. P. Knight, and D. M. Mayer (2011). But affirmative action hurts us! race-related beliefs shape perceptions of white disadvantage and policy unfairness. *Organizational Behavior and Human Decision Processes 115*(1), 1–12.

Singh, P. and O. N. Shemyakina (2016). Gender-differential effects of terrorism on education: The case of the 1981–1993 punjab insurgency. *Economics of Education Review 54*, 185–210.

Swee, E. L. et al. (2009). On war and schooling attainment: The case of bosnia and herzegovina. Technical report, Households in Conflict Network.

Valente, C. (2013). Education and civil conflict in nepal. *The World Bank Economic Review 28*(2), 354–383.

Van der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: A regression–discontinuity approach. *International Economic Review 43*(4), 1249–1287.

Verwimp, P., P. Justino, and T. Brück (2009). The analysis of conflict: A micro-level perspective.

Walker, A. and J. Willoughby (1856). *Measures adopted for the suppression of female infanticide in the province of kattywar, &c.*

Weiner, M. (2015). *Sons of the soil: Migration and ethnic conflict in India*. Princeton University Press.

Wong, Y. N. (2012). World development report 2012: Gender equality and development. In *Forum for Development Studies*, Volume 39, pp. 435–444. Taylor & Francis.

Young, A. (2015). Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *E, 0: 0–0*.