Essays on Development and Education Economics

Md Amzad Hossain
Dhaka, Bangladesh

M.A. Economics, University of Virginia,2017
M.S.S. Economics, University of Dhaka, 2010
B.S.S. Economics, University of Dhaka, 2009

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of Doctor of Philosophy

Department of Economics

University of Virginia
April, 2022

Committee Members:

1. Sheetal Sekhri
2. Sandip Sukhtankar
3. Leora Friedberg
4. Gaurav Chiplunker (External)

Abstract

My dissertation comprises three independent but thematically related chapters on education and development economics. The first chapter, "Performance Gains from Gender Match in Higher Education: Evidence from a Setting with Entrenched Gender Stereotypes," investigates whether female college students in a male-dominated discipline benefit academically from being taught by female instructors in a developing-country setting. I use a novel and confidential administrative dataset from a renowned economics program in Bangladesh to show that when matched with female teachers, female students gain in terms of both grade performance (nearly eight percent of a standard deviation) and longer-term outcomes such as degree completion time and the likelihood of enrolling in an economics master's program. I address endogeneity concerns using student, cohort, and teacher-by-course fixed effects. The quasi-random allocation of students to mandatory courses with no scope for selection of courses or instructors and double-blind final exams enable me to address selection further. I present suggestive evidence that the gain from matching is driven by the *role model effect* rather than bias in teachers' assessments.

In my second chapter, "Unintended Consequences of a Well-Intentioned Policy: Impact of Credit on Child Labor in Bangladesh," I find a 7.1 percentage points increase in child labor in response to an agricultural credit expansion program in Bangladesh, based on a field experiment. I show that this increase in child labor is due to new opportunities for children to work in household self-employment activities. I also find that treated households with fewer working adults use more child labor and spend less on education. While I do not see any effect on schooling outcomes, the time budget survey reveals that children from treated areas spend significantly less time studying (17 percent less than the control mean of 20.7 hours per week). These findings raise concerns about the unintended inter-generational consequences of easing credit constraints to increase self-employment.

My third chapter, "Access to Colleges, Human Capital, and Empowerment of Women," (co-authored with Sheetal Sekhri and Pooja Khosla, both UVA), uses the variation in college construction in India under a college construction grant policy to explore whether increased access to colleges in home districts improves human capital and the agency of women. Our estimates indicate an 11 percent gain in years of schooling (64 percent of which comes from college enrollment) for women in treated districts in the post-policy period over the baseline control mean of 8.6 years. The benefits accrue to the rural women but do not spill over to the neighboring districts. We also find suggestive evidence of changes in spousal quality in the marriage market.

JEL classification: I23, J16, O15, G21
Keywords: higher education, gender gap, human capital, child labor

Acknowledgement

I am grateful to my advisors, Sheetal Sekhri, Sandip Sukhtankar, and Leora Friedberg, for their guidance, support, and time throughout this dissertation process. Their insightful feedback, continuous encouragement and persistent belief in my abilities have been instrumental in instilling a sense of confidence in me to excel in research and sharpen my thinking. I extend my gratitude to Gaurav Chiplunkar, Eric Chyn, Jonathan Colmer, Shan Aman-Rana, and Sarah Turner, for valuable feedback at many stages of this work. I am also thankful to the many faculty and students at the University of Virginia and seminar participants at the Eastern Economic Association Meeting for their helpful comments and suggestions. I would also like to thank my colleagues and friends, Pooja Khosla, Moogdho Mahzab, Eric Robertson, Nirman Saha, and Luan Santos. Finally, I am thankful to the UVA Graduate School of Arts and Sciences and Bankard Fund for Political Economy for their financial support.

This dissertation is dedicated to my parents and siblings

Table of Contents

Chapter 1: Performance Gains from Gender Match in Higher Education: Evidence from a Setting with Entrenched Gender Stereotypes 9

Abstract 9

1 Introduction 10

2 Background and Data 16

3 Empirical Strategy 19

3.1 Selection into Courses and Professor Choice 19

3.2 Empirical Specification in the Quasi-Experimental Setting 20

4 Main Results and Discussion 25

4.1 Grade Performance 25

4.2 Other Medium- and Longer-Term Outcomes 26

5 Possible Mechanisms 29

5.1 Is There Grading Bias? 29

5.2 Do Students' and Teachers' Observable Characteristic Matter? 31

5.3 Role Model Effect VS Teacher Effectiveness 32

6 Conclusion 35

References 38

Appendix Figures and Tables 57

Chapter 2: Unintended Consequences of a Well-Intentioned Policy: Impact of Credit on Child Labor in Bangladesh 64

Abstract 64

1 Introduction 65

2 Program Description and Experimental Design 72

2.1 Background and Description of the BCUP Program 72

2.2 Experimental Design 73

2.3 Sample Selection 74

2.4 Baseline Sample Description and Balance 75

2.5 End-line Survey and Attrition 77

3 Empirical Specifications 78

4 Results and Discussions 79

4.1 Effect on Borrowing Behavior 79

4.2 Effect on Self-Employment Activities ... 82
4.3 Effect on Child Labor ... 83
4.4 Evidence from Time Use Survey ... 84
4.5 Heterogeneity Analysis ... 86
4.6 Education Expenditure and Schooling Outcome ... 87
5 Robustness ... 88
5.1 Multiple Hypothesis Testing and External Validity ... 88
5.2 Other Robustness and Sensitivity Checks ... 90
6 Conclusion and Policy Implications ... 91
References ... 94
Supplementary Appendix ... 112
A Appendix Figures and Tables ... 112
B LATE Estimates ... 123
C Difference-in-Differences (DiD) Estimates ... 128
D Regressions Using Only Endline Data ... 132
E Regression Using Lasso Controls ... 136
F Regression Using Wild Cluster Bootstrap-t Procedure ... 140
Chapter 3: Access to Colleges, Human Capital, and Empowerment of Women ... 144
Abstract ... 144
1 Introduction ... 144
2 Background ... 152
3 Data ... 154
3.1 College Expansion and Eligibility ... 154
3.2 Educational and Other Outcomes ... 155
3.3 Other Ancillary Data ... 156
3.4 Summary Statistics ... 156
4 Empirical Strategy ... 157
4.1 Event Study ... 157
4.2 Trend-Break Estimation ... 159
4.3 Evidence from RDD ... 161
5 Main Reduced Form Results and Discussion ... 162
5.1 Educational Outcomes ... 162
5.2 Trends in Other Potential Confounders and Other Robustness Tests ... 165

5.3 Heterogeneity by Rural Status and Wealth 169
5.4 Spillovers 170
5.5 Benefits of Education: Labor and Marriage Market Outcomes 171
5.6 Other Development Outcomes 172
6 Conclusion 173
Notes 176
References 180
Supplementary Appendix 198
Appendix Figures and Tables 204

Chapter 1: Performance Gains from Gender Match in Higher Education: Evidence from a Setting with Entrenched Gender Stereotypes

Md Amzad Hossain

Abstract

Worldwide, women are severely underrepresented in science, technology, engineering, and mathematics (STEM) education, and in developing countries in particular. This study sheds light on whether female college students in a male-dominated STEM course such as economics benefit academically from being taught by female instructors in a setting in which gender stereotypes are an entrenched social phenomenon. I use a novel and confidential administrative dataset from a highly rated economics program in Bangladesh to show that when matched with female teachers, female students gain in terms of grade performance as well as longer-term outcomes such as degree completion time and the likelihood of enrolling in an economics master's program. The quasi-random allocation of students to mandatory courses, with no scope for students to select courses or instructors, alleviates the self-selection concerns. Comparing the test scores of blindly and non-blindly graded exams for the same course, I rule out the explanation whereby the gain from gender matching is driven by gender reference in teachers' assessment. I find that the benefit of matching increases with female teachers' rank, experience, and academic qualifications. I show suggestive evidence that female teachers' effectiveness in teaching female students is an important channel, aside from the frequently cited role model effect, through which same-sex teacher assignments improve female students' academic achievement.

1 Introduction

Over the past three decades, female students in most developed countries have made remarkable gains in education relative to male students (Vincent-Lancrin, 2008; Goldin et al., 2006). Women's participation in science, technology, engineering, and mathematics (STEM) fields, however, remains notably low. For instance, the probability of obtaining a college degree in science and engineering is 37 percent lower for female students in the U.S. (Summers, 2005), and women comprise only one-fourth of the STEM workforce in the U.S. (Green, 2007). The situation is similar for economics, which is often considered part of STEM.[1] Since 2000, the share of female Ph.D. students and female assistant professors in economics programs has remained stagnant at 35 percent and 30 percent, respectively (Lundberg and Stearns, 2019).

This begs the question of why so few women are in STEM fields. Aptitude and preparedness for college do not seem to explain this (Carrell et al., 2010).[2] Social and environmental factors have been proposed as alternative explanations. One such factor is socio-cultural stereotypes (Ambady et al., 2001; Nguyen and Ryan, 2008); for example, many people believe that women are not as good as men in math and science. The literature on stereotypes documents that negative stereotypes can substantially impede performance (Aronson et al., 1998; Croizet and Claire, 1998; Spencer et al., 1999).

A related second factor is the college environment (Hill et al., 2010), such as the gender of the teachers female students are exposed to in college (Carrell et al., 2010; Hoffmann and Oreopoulos, 2009). For instance, teachers may behave differently toward

[1] In its list of science, technology, engineering, and math (STEM) degree program, the Department of Homeland Security (DHS) in 2012 included Quantitative Economics.
https://www.ice.gov/sites/default/files/documents/Document/2014/stem-list.pdf

[2] The small differences that do exist in high school math and science can't explain this (Xie and Shauman, 2003), nor can the nearly nonexistent differences in college preparatory math and science courses (Goldin et al., 2006).

students of different genders because of their beliefs about a student's innate abilities or relatability. In a related vein, Carlana (2019) finds that female students tend to underperform in math, and sort themselves into less demanding high schools when they are matched with teachers who have stronger gender stereotypes. Similarly, students may respond differently to teachers of a different gender, due to teaching styles, pedagogical techniques, or effective role modeling (Hoffmann and Oreopoulos, 2009). I focus on the role of teacher gender in this study.

In this paper, I estimate the gains from the gender matching between students and teachers at the college level in the context of a developing country. More specifically, I test whether female students gain academically when taught by female teachers. I focus on developing countries in which gender prejudices and stereotypes are an entrenched social phenomenon (Jayachandran, 2015). Societal norms and cultural values often discourage girls and women from studying STEM in many such countries, believing that it is a domain for boys and males. According to a 2015 UNESCO report, only 20 percent of STEM researchers in South and West Asia are women, whereas the corresponding worldwide figure is 30 percent.[3] These wide gender gaps in developing countries are purely based on choice, since girls and women in developing countries are often held back by biases, social norms, and expectations that influence the subjects they study.

I focus on the academic performance of female college students for several reasons. First, economists have documented a positive correlation between undergraduate GPA and post-college earnings (Wise, 1975; Filer, 1983; Jones and Jackson, 1990), either through human capital or screening or both. Second, undergraduate GPAs are crucial in Ph.D. admissions decisions (Attiyeh and Attiyeh, 1997). Third, in the context of developing countries, in which information asymmetry is more prevalent (Luoto et al., 2007), GPA often

[3] http://unesdoc.unesco.org/images/0023/002351/235155E.pdf

works as a crucial factor in job screening.[4] This is truer for jobs at universities and research institutes.[5] Finally, in deciding whether to write letters of recommendation for students, teachers use students' academic performance as one of the most important criteria. Hence, college exams are high-stakes tests.

I estimate the effect of student-teacher gender matching using a novel administrative dataset from a highly selective economics program in Bangladesh. My data track every student by semester for all cohorts between 2007 and 2017. These data contain comprehensive information on students' course-level academic outcomes; medium-term academic outcomes, including number of years repeated by students, the likelihood that they will complete their degree in four years, and their cumulative grade point average (CGPA); longer-term outcomes, such as the likelihood of enrolling in an economics master's program; and detailed demographic information for the corresponding instructors.

Two important identification challenges arise when investigating the role of a professor's gender in the academic outcomes of female students. First, students could choose their courses. Second, students could choose their professors. My setting allows me to address these selection pitfalls. Students in the program take mandatory courses each semester, for which they have no choice as to course or instructor, and I limit my analysis to these mandatory courses. I rely on two sources of variation for my identification: First, I use within-cohort cross-sectional variation, whereby students of a particular cohort are taught the same mandatory courses by the same sets of teachers. Second, I exploit across-cohort variation in teacher gender over the same courses. This temporal variation is caused by attrition (due to retirement and different types of leave), or due to the alternative assignment of teachers across years, whereby several teachers aspire to teach the same course, or the same teacher aspires to teach different courses over time.

---

[4] https://www.thedailystar.net/op-ed/politics/does-cgpa-matter-getting-job-1274815

[5] https://tbsnews.net/feature/pursuit/low-cgpa-it-end-job-opportunities-78283

These data on students of different cohorts across different semesters enable me to control for a rich set of fixed effects and to estimate the differential effect between female and male students of being exposed to a female teacher in the same course. The gender match varies for students within cohorts for different courses and across cohorts within the same course in each semester. This allows me to control for student, course, cohort, teacher, and teacher-by-course fixed effects. This helps eliminate several concerns about omitted variable bias that arise in the literature on the effects of gender match on academic outcomes.

I find that female students perform relatively better in courses taught by female teachers. Female students' grades in the mandatory courses improve by approximately 9 percent of one standard deviation when the assigned course teacher is female. This effect is similar in magnitude to that of Carrell et al. (2010), and much higher in magnitude than that of Hoffmann and Oreopoulos (2009).[6] While estimates of the gain from pairing female students with female teachers are significant for math-intensive courses such as mathematics, statistics, and econometrics, the increase in test scores is twice as high for other courses than for those in mathematics, statistics, and econometrics.[7]

I also estimate the effect of the proportion of mandatory courses a student takes from female professors, both on degree completion time and number of years repeated by a student. I find that the proportion of mandatory courses taken by a female student and taught by female professors increases the likelihood of graduating with a regular cohort for female students and results in a higher GPA. In addition, more exposure to female teachers during the bachelor's degree increases the likelihood of enrolling in an economics master's program

[6] Carrell et al. (2010) find that female students' performance in math and science courses improves by around 10 percent of a standard deviation when the course is taught by a female professor, whereas Hoffmann and Oreopoulos (2009) find an increase of at most 5 percent of its standard deviation in grade performance when matched with a same-sex teacher.

[7] These other courses include, among others, introductory microeconomics and macroeconomics, intermediate microeconomics and macroeconomics, international trade, development economics, public economics, etc.

for female students. I bolster my findings by performing a number of robustness checks to rule out the possibility that the assignment of teachers is based on the gender composition of the class or performance of the female students.

I then explore the mechanisms through which female students benefit from taking classes with female teachers. A potential mechanism is that the improvement in female students' performance may be due to bias in teachers' assessment of tests (Bar and Zussman, 2012; Hanna and Linden, 2012; Burgess and Greaves, 2013). A feature of this study's setting, which allows me to address this, is that the students take a non-blindly graded midterm and a blindly graded final exam for each course. I do not find any evidence that teachers favor students of the same gender.

Once I establish that the gain from matching is not driven by gender bias in teacher's assessment, I look at alternative mechanisms. Are female teachers more effective in teaching female students? Or, do female teachers serve as "role models" for their students? Earlier studies mostly cite role-model effect as the driving mechanism without going into much detail (Carrell et al., 2010; Hoffmann and Oreopoulos, 2009; Bettinger and Long, 2005). While it is very much difficult to distinguish between the role-model effect and the teacher-effectiveness channel, I provide suggestive evidence that female teachers' effectiveness in teaching female students is an important channel through which same-sex teacher assignments improve female students' academic achievements. I show this in two ways. First, I produce several evidence that the role-model effect cannot explain; Rather the teacher-effectiveness channel can more persuasively explain those. Second, I use the insights from the theoretical models in the literature that formalizes the notion of "role-model effects" as distinct from teacher effectiveness. Using the testable implications from these models (Gershenson et al., 2018), I show evidence that teacher effectiveness is an important channel.

I make a novel contribution to the literature on student-teacher gender match in three ways. First, I provide the first causal estimates of gender matching between students and teachers at the college level in the context of a developing country. Several studies investigate the role of teacher gender at the post-secondary level in developed-country settings (Rothstein, 1995; Neumark and Gardecki, 1998; Bettinger and Long, 2005; Hoffmann and Oreopoulos, 2009; Carrell et al., 2010), but to the best of my knowledge, no such study exists for developing countries. Because stereotypes and discrimination are much more pervasive in developing countries (Jayachandran, 2015), therefore, my paper fills an important gap. Second, by focusing on college, I examine to what extent the gender-interaction effects observed in schools in developing countries (Muralidharan and Sheth, 2016) are present at later ages. This is important since academic performance at the college level has more immediate implications for STEM higher education and consequently, for STEM labor market. Finally, the quasi-random allocation of students in my setting, whereby the students cannot choose courses or instructors, allows me to estimate the effect of student-teacher interaction on different student outcomes while avoiding the self-selection problem often encountered in prior research (Carrell et al., 2010).

I also contribute to the literature on gender bias in teacher assessments. Several studies have exploited the difference in the test scores of blind and non-blind exams to document evidence of such biases in teachers' grading, a technique pioneered by Lavy (2008). Several studies find substantial bias against boys in teacher assessments (Lavy, 2008; Falch and Naper, 2013; Cornwell et al., 2013; Lindahl, 2016), while others document no such gender bias (Hinnerich et al., 2011). In this paper, I go one step further and, following Lindahl (2016), ask whether teachers favor students of their own gender. However, unlike Cornwell et al. (2013), who find that female teachers are less generous when grading the non-blind exams of female students, I do not find evidence to support the view that teachers either favor or disfavor students of the same gender. My paper also differs from earlier studies, in which blind and non-blind exams are graded by different graders. In my settings, both exams are graded by the same person.

My paper also adds value to the literature by investigating the mechanism through which female students benefit from being paired with female teachers. Previous researchers cite the role-model effect as the key mechanism without going into much detail (Carrell et al., 2010; Hoffmann and Oreopoulos, 2009). However, in this paper, I provide suggestive evidence that female teachers' effectiveness in teaching female students can also be an important channel.

Lastly, my paper makes a direct contribution to the debate concerning increasing female representation in a male-dominated field such as economics. Increasing female representation in STEM fields has frequently been advocated (Hill et al., 2010). This paper studies the impact of matching female students with female teachers, and whether such matching can influence student academic performance. Against the backdrop of ongoing debates and legislative changes with respect to affirmative action in developing countries, this topic warrant investigation.

The remainder of the paper is organized as follows. Section 2 describes the dataset and Section 3 discusses the empirical strategy employed in this paper. In Section 4, I present my main results. Section 5 investigates the mechanisms, and Section 6 concludes.

2 Background and Data

My analysis is based on the administrative data from the economics department of one of the most reputed public universities in Bangladesh. Admission into this university is highly competitive, and its economics program is equally selective. Fewer than one out of 200 applicants are accepted for admission into the department, for a total of 150 seats. The department offers a four-year degree in economics, where each student has to complete eight semesters of coursework. Each student is therefore expected to graduate in four years, although some students take more time due to repetition.[8] Each semester, a student takes a combination of mandatory and optional courses. The number of mandatory and optional courses is the same for all students of the same cohort, although it may differ across cohorts. A student failing to achieve the minimum threshold GPA at the end of her year must repeat the year with the subsequent cohort. Each student in the same cohort has the same course instructor for a particular course, and multiple sections of the same course taught by different professors are not offered. Thus, for a given course, the student cannot select the course teacher.

Each course is graded on a scale of 100 points. This is composed of two parts – sessional exams and Final exam – each worth 50 points. The sessional exams consist of one midterm examination (30 points) and in-class exams (20 points). The Final exam carries 50 points and is held at the end of the semester. The same instructor grades both the sessional and final exams. But, there is an important distinction between these two types of exams.

The sessional exams are non-blindly graded exams, where the instructor – who is also the grader, knows the identity of the students. The final exam, on the other hand, is blindly graded exam where the instructor-cum-grader is unaware of the student's name and gender.[9]

I collect detailed information on course outcomes that include scores received in both the sessional exam and the final exam for every course offered for every student enrolled at the department from January 2007 to January 2017. My sample consists of approximately 1500 students from 11 cohorts. Approximately 35 percent of the students in my sample are female

[8] This system is different from a four-year college degree in countries like the U.S., where a student must choose his/her major in the sophomore (second) or junior (third) year. In Bangladesh, students must choose their major before enrollment into colleges, and once enrolled, they cannot switch to other disciplines.

[9] For the first four cohorts, the exams were sent to a second examiner for additional blind grading. In that case, the final grade was the average of the two grades.

(Table I). My data also includes information on the academic records of students repeating a year. Further, I collect student's baseline characteristics at entry such as merit score in the entrance exam, the major in high school, and whether the student was admitted on reserved seats.[10]

Academic performance measures in the data consist of Sessional grades (50 points), Final grades (50 points), total grades (100 points), and grade point (on a scale of 4.00) for every individual student by course and semester. Students at the Department of Economics are required to take a mix of approximately 35 mandatory and optional courses, but in my analysis, I use the data for the mandatory courses as there is no selection possibility in these courses.

Grades points are assigned on an A+, A, A, B+, B, B-, C+, C, D, F scale, where an A+ is worth 4-grade points, an A is 3.75-grade points, an A is 3.50-grade points, a B+ is 3.25-grade points, and so on (see Appendix Table A.1 for a detailed description). The sample cumulative grade point average (CGPA) for female and male students is 3.14 and 2.95, respectively. I standardize both the course grades and grade points, such that each variable has a mean of zero and a variance of one, within each course offered to a particular cohort.

For each of the courses taught during different semesters in the department, individual instructor-level data were obtained from the department. I collect data on each teacher's gender, academic rank, the highest level of education (M.A. or Ph.D.), and years of tenure in the department. These were merged with the data on students' academic achievement. During the period I study, 66 different teachers taught all courses. Of these, 30 percent

[10] Five percent of the total seats are reserved for the students belonging to an ethnic minority or to freedom fighter family and to the children of the university staffs.

(20 of 66) were female and taught nearly 42 percent (101 of 243) of the mandatory courses. In this sample, 50 percent of the female faculty members had a Ph.D. degree, while the corresponding number for the male faculty members is approximately 43 percent.

3 Empirical Strategy

3.1 Selection into Courses and Professor Choice

The existing literature faces two main identification challenges in estimating the impact of female students' exposure to female teachers. First, a student could choose their courses. As explained earlier, in my setting, students must take certain mandatory classes each semester, in which they have no choice. My primary analysis focuses on these classes. Another advantage of restricting my analysis to the mandatory courses is that the class size of those courses is large—typically more than 100 students—which reduces the likelihood that my results are based on anomalous circumstances (Hoffmann and Oreopoulos, 2009).

The second selection issue is that students could choose their instructor if multiple instructors were to teach the same course. Again, this problem does not arise in my setting. This is because every student in the same cohort is exposed to the same teacher for a particular course. This instructor is assigned by the department on the basis of expertise and experience. Therefore, the identification of female student-teacher interactions comes from the rich cross-cohort variation in teachers' gender for different courses (see Figure 1 for a stylized illustration of this). The cross-cohort variation in teachers' assignments to different courses is due to several factors, which I treat as exogenous. First, there is attrition due to retirement or for personal reasons. For example, a large number of young teachers leave the department to pursue graduate studies. Second, when several teachers specialize in a particular field, they are assigned relevant courses to teach in alternate years. Third, sabbaticals of different natures also affect which professors teach mandatory courses in a particular semester.

The fact that I limit my analysis to mandatory courses, coupled with the fact that students have no role in selecting teachers, gives me a quasi-experimental design to estimate causal effects. I also provide empirical support for my claim that the allocation of students to classrooms is random. To this end, I compare the female students taught by male and female teachers, and the male students taught by male and female teachers. If the assignment of the students to male and female teachers were random (not based on observable student characteristics), we would expect the groups of male students and female students to be similar, in terms of observable characteristics, when they are taught by male and female teachers, respectively. I report the sample mean for all students when they are taught by male and female teachers in Table II. I show the characteristics of female students by gender of the teacher in Panel A, and the characteristics of male students by gender of the teacher in Panel B. We do not observe any significant difference in the mean characteristics of the two groups of students by teacher's gender. This suggests that students are not more likely to be matched with a same-sex teacher on the basis of observable characteristics.

I do several other robustness checks to show student assignments to teachers are random. For instance, a concern may be that female teachers could be assigned to specific classes based on gender composition of the class or performance of the female students. In order to rule this out, I run regressions to show that the proportion of mandatory courses taught by female teachers in a single semester is not influenced by gender composition of the class (Appendix Table A.2) or female students' performance in the preceding semester (Appendix Table A.3).

3.2 Empirical Specification in the Quasi-Experimental Setting

Every student belonging to a particular cohort is assigned only one teacher for a specific mandatory course. However, across cohorts, this same course may be taught by a female or a male professor. Thus, I rely on two sources of variation: I use the within-cohort cross-sectional variation, where students from the same cohort are taught by different teachers for different courses. In addition, I use the across-cohort variation, in which students from different cohorts are taught by different teachers for the same course. To clarify my identification strategy, consider the following: Students of the first cohort enter the program. They are obligated to take a certain number of mandatory courses in each semester, from first to eighth. The order of these courses is fixed by the department, and students cannot take those courses in a different order. Students of this cohort take mandatory courses, some of which are taught by male teachers, and some by female teachers. This is the cross-sectional variation. Now, students of the second cohort enter. They are offered coursework in a similar order. The only thing that is changed is that different teachers are assigned for their courses relative to the first cohort (due to the reasons outlined before). This gives rise to circumstances where a course that was taught by a male teacher for the first cohort is taught by a female teacher for the second cohort. This across-cohort temporal variation in the gender of instructors teaching the same courses allows me, therefore, to account for not only student, teacher, course, and cohort fixed effects, but also teacher-by-course fixed effects.

The rich set of fixed effects helps in identification in the following way. First, there is a possibility that students taking courses taught by female teachers are systematically different from those who are not, regardless of their gender. This absolute sorting, while not likely in my setting, is accounted for by including student fixed effects in my regression model. Second, it is possible that female students take courses from teachers who are systematically different from other teachers. I eliminate this concern by including individual teacher fixed effects. Third, it might be the case that male and female teachers are systematically different in teaching certain courses. To mitigate this concern, I include teacher-specific course fixed effects. Finally, students within a particular cohort-semester are taking the same courses and are exposed to similar classroom-specific shocks, such as the time of day or different other external disruptions. I include classroom (cohort-semester-course) fixed effects to control for this concern. Absorbing these fixed effects helps me compare the difference in outcomes of male and female students of the same cohort when they are matched with female teachers compared to those when taught by a male teacher.

To estimate the effect of student-teacher interaction, I start with the regression model of the following form:

$$y_{ijkst} = \alpha_0 + \alpha_1 \times FT_{jkst} + \alpha_2 \times FS_{ikst} + \alpha_3 \times FT_{jkst} \times FS_{ikst} + X_{ijkst}\beta + u_{ijkst} \quad (1)$$

where I index students by i, teachers by j, course by k, cohort by s, and semester by t. FT_j and FS_i are dummy variables taking values equal to one if teacher j and student i are female. X_{ijkst} is a set of observable variables, such as teacher's experience and highest level of education. Finally, u_{ijkst} is the set of unobservables. I am primarily interested in α_3, which shows the differential effect between female and male students after being exposed to a female teacher for a particular course. A positive value of this parameter implies that female students perform relatively better than male students, when taught by a female teacher.

Including all these fixed effects in equation (1), I obtain a regression model of the following form:

$$y_{ijkst} = \alpha_3 \times FT_{jkst} \times FS_{ikst} + \gamma_i + \lambda_j + \varphi_{kst} + u_{ijkst} \quad (2)$$

where γ_i are the student fixed effects, λ_j are the teacher fixed effects, and φ_{kst} are cohort semester-course fixed effects (henceforth, classroom fixed effects). Including these fixed effects allows me to overcome many threats to internal validity. However, when we include these fixed effects in Equation (2), student and teacher variables at the level form are dropped, due to multicollinearity with either of the fixed effects. The parameter α_3 estimates the effect of student-teacher interaction.

One remaining possibility is that the female students gain from being exposed to female teachers because female teachers favor female students in grading. To test for gender bias in teachers' assessments, I exploit a rule that dictates that while the sessional exam (mid-term) of a course is non-blind, the final exam must be blind. Therefore, it may safely be assumed that scores in the blind test would not be influenced by any bias resulting from graders' gender stereotypes. To the contrary, scores in the non-blind exam may be affected by biases resulting from graders' stereotyping and discrimination. Since both the blind and nonblind exams test the same cognitive skills, I can use the blind score as a valid counterfactual of the non-blind score, which might be influenced by stereotyped discrimination.[11]

The sessional exams and the final exams are different in some respects. First, the sessional exams do not cover the complete syllabus of the course, whereas the final exams are based on the full syllabus. Second, the sessional exams are shorter (midterms take two hours, whereas the final exams take three hours). Yet, at the same time, they bear many similarities. First, both the sessional (non-blind) and final exams (blind) measure students' academic performance. Second, both exams in a particular course are set by the same professor, and although the final questions are considerably longer, the questions' nature is quite similar. Third, both exams take place in a similar type of exam-taking environment. This reduces the concern that differences in the exam environment that might be correlated with certain stereotyped feminine characteristics (e.g., possible higher anxiety levels) pose threats to the identification. All these similarities imply that both blind and non-blind tests measure the same cognitive skills. Therefore, once I control for the exam (i.e., blind versus non-blind) fixed effects, blind scores could be construed as a valid counterfactual of the non-blind score. I explore the difference between male and female students' gaps between the scores of the blind and non-blind exams in order to capture potential gender bias. To control for systemic

[11] In this regard, my identification strategy resembles that of Lavy (2008), which was in turn inspired by Goldin and Rouse (2000) and Blank (1991).

differences between the two tests, I control for examination-type fixed effects that control for any time-invariant unobservable characteristics. More precisely, to examine whether teachers favor students of their gender, I run the following model,

$$y_{ijkste} = \alpha_3 \times FT_{jkst} \times FS_{ikst} + \alpha_4 \times A_e + \alpha_5 \times FT_{jkst} \times A_e + \alpha_6 \times FS_{ikst} \times A_e$$
$$+ \alpha_7 \times FT_{jkst} \times FS_{ikst} \times A_e + \gamma_i + \lambda_j + \varphi_{kst} + u_{iec} \tag{3}$$

where A_e is an indicator variable taking a value of 1 if the exam is blind (anonymous) and 0 otherwise (non-blind). The main parameter of interest is α_7, which shows the difference between the blind (anonymous) scores of male students and those of female students, when they are matched with a female instructor, given the respective difference in the non-blind scores. A positive value of α_7 indicates that female teachers favor female students while grading.[12]

Finally, I measure how instructors' gender influences longer-term outcomes, such as the number of years repeated, or the likelihood to graduate with regular cohort, or a student's cumulative grade point average. To this end, I run the following model, which is a variant of the regression model (2)

$$y_{ist} = \lambda_s + \alpha_1 \times FS_{is} + (\alpha_2 + \alpha_3 FS_{is}) \frac{\sum FT_{jst}}{n_{ist}} + u_{ist} \tag{4}$$

where y_{is} is the outcome of student i belonging to cohort s up to semester t. The outcome variable is an indicator variable taking values of 1 if a student graduates with their regular cohort, or is a count variable indicating the number of years repeated by the student. $\frac{\sum FT_{jst}}{n_{ist}}$ is the share of mandatory course teachers j who were female for student i belonging to

[12] Though unlikely, it is possible that teachers speculate about students' genders from their handwriting in the test. In order to address this, I use yet another feature of the testing system, which I discuss in a later section concerning robustness tests.

cohort s up to semester t. I include this variable in the model to help estimate the average impact of being exposed to more female teachers in mandatory courses. As in Equation (1), I am primarily interested in α_3, which estimates the differential effect across male and female students of being assigned to more female teachers in the mandatory courses.

4 Main Results and Discussion

4.1 Grade Performance

Estimates of the effect of the interaction between female students and female teachers on standardized test scores are reported in Table III. In Column 1, I report a basic specification without any fixed effects. In Column 2, I control for baseline characteristics, such as merit score in the entrance exam, the students' fields of study in high school, and whether the student was admitted as part of a quota. I add an array of fixed effects in the subsequent columns to identify the parameter of interest: Column 3 includes student fixed effects, and Column 4 adds teacher fixed effects. I add classroom (cohort-semester-course) fixed effects in Column 5, and course by teachers' gender fixed effects in Column 6. I cluster the standard errors at the classroom (cohort-semester-course) level.

Table III shows that the estimate of the female student-teacher interaction is statistically significant across all specifications. In specification that include student fixed effects, female students' grades in mandatory courses increase by about nine percent of a standard deviation when a female instructor teaches the course. Once student fixed effects are included, these results are robust to the inclusion of teacher, cohort-semester or course fixed effects. This increase in grade performance is much higher compared to that estimated in Hoffmann and Oreopoulos (2009), who found that a same-gender teacher improves the average grades of female students by, at most, five percent of a standard deviation.

Two other results are worth noting, although the identification is not as strong as when we include the fixed effects. First, in column 1, where we do not include student fixed effects, we find that female students perform better than their male counterparts regardless of the gender of the teachers. This result reflects the finding of Table I, which shows that female students perform much better in terms of grade performance and medium-term outcomes. Second, male students also perform better when they are matched with female teachers. Male students' grades increase by 2.8 percentage points of a standard deviation when they take classes with female teachers instead of male teachers. The estimates of the gain are much higher when we include student fixed effects (Column 2 of Table III).

I then examine whether the increase in female grade performance from being matched with a female teacher differs across types of subjects. To this end, I divide all the mandatory subjects into two groups: (a) math, statistics and economics, and (b) all other courses. The results are documented in Table IV, which demonstrates that although the estimates of the gain from pairing female students with female teachers are significant for both groups of subjects, the increase in test scores is twice as high for other courses than those in mathematics, statistics and econometrics. This result is qualitatively similar to Hoffmann and Oreopoulos (2009), who found that gains from the interaction tend to be higher in social science courses compared to math courses.

I also shed light on which groups of students benefit more from gender-matching, in terms of ability. To this end, I run a quantile regression model of the student-teacher interaction. The results are depicted in Figure 2. It is evident that female students belonging to the middle tier of the distribution of normalized scores gain most when they are exposed to female teachers. I supplement this distributional analysis by estimating the interaction effect at different quantiles of the merit score in the entrance exam and find consistent results (Appendix Table A.5).

4.2 Other Medium- and Longer-Term Outcomes

I also look at the impact of matching female students with female teachers on medium-term outcomes such as cumulative grade point average, the likelihood of being retained in the regular cohort, and the number of years repeated by the students. I run the model specified in Equation 4 for these outcomes. As seen in Table V, the estimates of the interaction of female students and the proportion of female teachers in all mandatory courses on cumulative grade point average are statistically significant. Given that around 43 percent of all the mandatory courses are taught by female teachers, and the standard deviation of CGPA for male students is 0.71, the interaction coefficient of .0054 translates to a gain of around 33 percent of a standard deviation for the female students compared to the male students.[13] Similarly, as students are exposed to more and more female teachers as they progress through college, the probability of being retained in their regular cohort increases, and the number of years repeated declines. For instance, a one percentage point increase in the share of female teachers is associated with a 0.4 percentage point increase in the likelihood of passing with the regular cohort for the female students.

I then look at longer-term outcomes such as the likelihood of enrolling into an economics master's program and cumulative grade point average in master's conditional upon enrollment. Column 4 of Table V shows that a one percentage point increase in the share of female teachers in the compulsory undergraduate classes increase the likelihood of enrolling into economics master's program by 1.14 percentage point for the female students. Conditional on enrollment, however, we do not see any significant effect of female student-teacher matching on female students' academic performance in the master's degree.

One point to note that female students' exposure to female teachers in the first year alone has no impact on academic achievement in the later years. Table XII shows that the coefficient of the interaction of female students and the proportion of female teachers in first-year compulsory courses is statistically insignificant for most of the medium and longer-term outcomes such as Cumulative grade point average (CGPA), or the number of

[13] (0.0054 x 43) =0.23, which is approximately 33 percent of 0.71.

years repeated by a student, or the probability of enrolling into an economics master's program.[14] These results are similar to those of Carrell et al. (2010), who shows that having a high proportion of female teachers in the introductory classes does not affect longer-term outcomes such as withdrawal from the program within the first two years of enrollment or likelihood of graduating with a STEM degree. On the contrary, as Table V suggests, it is the continual interaction of female teachers and female students that affects medium-term outcomes such as overall CGPA, or the likelihood of graduating with the regular cohort, and longer-term outcome such as the likelihood of enrolling into economics master's program. These results suggest that it is the continuous exposure of female students with the female teachers, rather than the one-shot exposure, that affects female students' medium and longer-term gains.

However, one should interpret the results on medium-and longer-term analysis with one caveat in mind: the identification is not as strong as was in the case of my main analysis in Section 4.1 since I could not control for student and teacher fixed effects. Regardless, this analysis suggests longer-term gains from same-sex student-teacher pairing. One result that particularly stands out is that having more female teachers leads to an increase in the likelihood of enrolling into a master's program in economics for female students. Female under-representation in graduate economics courses has been subject to much attention in recent times (Lundberg and Stearns, 2019). My results suggest that hiring more female teachers can reduce the gender gap in economics higher education without harming the boys.

[14] The only exception is the probability of passing with the regular cohort, for which more exposure to female teachers in the first-year compulsory courses actually worsens the outcome significantly.

5 Possible Mechanisms

In this section I explore the mechanisms through which female students benefit from taking classes with female teachers. First, I test whether the gains from matching are due to bias in teachers' assessment of test scores. Then, I examine whether the gains from matching can be explained by student and teacher observable characteristics. Finally, I discuss suggestive evidence about two potential mechanisms: the role-model effect and differential teacher effectiveness.

5.1 Is There Grading Bias?

Several papers document bias in teachers' assessments and grades.[15] I therefore aim to test whether the improved grade performance of female students, when matched with female teachers, is driven by assessment bias. Following the seminal paper of Lavy (2008), I use the scores obtained by students on their blind and non-blind exams to test whether such biases in teachers' assessments exist. Table VI shows the effect of the match between female students and female teachers on normalized test scores for blind and non-blind exams separately. It is evident that the interaction effect of female students taught by female teachers is quite similar for both blind and non-blind exams. Female students' performance in mandatory courses increases by approximately nine percent of a standard deviation in the non-blind exam when a female is an instructor. In contrast, the corresponding increase is nearly seven percent of a standard deviation in the blind exams. The Chow test suggests that this difference is not
significant.

To test whether estimates of the interaction between female students and female teachers are driven by grading bias, I use a different approach by estimating a triple-difference model by interacting blind exam dummy with a female student dummy and female teacher dummy

---

[15] See, for example, Bar and Zussman (2012); Hanna and Linden (2012); and Burgess and Greaves (2013).

(Equation 3). The result is reported in the first column of Table VII. The main parameter of interest is the coefficient of the triple interaction of female student, female teacher, and the blind exam dummies. It turns out that the estimate of the triple interaction is insignificant, which implies that female teachers do not favor students of their gender when grading.

However, it is important to note several caveats regarding my analysis of blind and non-blind exams. First, the non-blind exams take place in the middle of the semester, while the blind exams are given at the end. This may bias my estimates if male and female students have different learning trajectories after the non-blind exams. I attempt to control for this by including student-by-exam-type fixed effects. Second, although the blind exams are anonymous, since the same examiner grades both the blind and non-blind exams, it is possible that the grader can infer students' identities from their handwriting. However, I believe this does not pose a threat to my study for several reasons. One, with a class size of more than 100, it is difficult to remember individual handwriting. Two, if the practice of inferring identity from handwriting is teacher-specific, that would be absorbed in teacher fixed effects. Three, for the first four cohorts in my study, the blind exams were graded by two examiners—the course teacher and an external examiner. The final grade was the average of the two grades. If the course teacher favored a student by inferring identity from handwriting, the interaction effect of the match between female students and female teachers would be different in the double-blind and single-blind exams for these cohorts. Yet, as can be seen from Appendix Table A.4, the estimate of the interaction is statistically similar in both single- and double-blind exams, and the Chow test of the equality of the two coefficients cannot be rejected. These facts suggest that inferring gender identity from a student's handwriting should not be a serious threat in the context of my study.

5.2 Do Students' and Teachers' Observable Characteristic Matter?

In this section, I examine heterogeneity in the treatment effect by teachers' and students' observable characteristics, with the goal of identifying mechanisms further. Table XI indicates that the gains from matching vary significantly with teachers' experience and rank. For instance, the gains are much higher when female students are matched with female teachers with higher hierarchical ranks (Column 1-3). Notably, the interaction effect is negative when entry-level teachers teach courses and are successively higher for assistant or associate professors and full professors. Similarly, I find that the gains from gender matching are higher in courses taught by teachers with more experience (Column 4-6). Finally, we see that estimates of the interaction of female students and female teachers are approximately 21 percent higher when the teacher has a Ph.D. than when the teacher does not (Column (8), Table XI). However, the Chow test reveals that the difference is not statistically significant. In Table VIII, I show these results by including the full set of interactions, in which I interact female student dummy, female teacher dummy, and teachers' observable characteristics. It is evident that gains from matching are driven by the rank and experience of the teachers.

I also examine whether observable student characteristics drive the gains from gender matching. To this end, I interact several student-level variables with the teacher gender dummy and student gender dummy to see whether the treatment effect of gender matching varies by student characteristics such as merit score and merit position in the entry exam, high school major, and quota status. However, Table IX shows that none of the estimates of the triple interaction variable are statistically significant. This suggests that student observable characteristics are not driving the female student-teacher interaction effect.

5.3 Role Model Effect VS Teacher Effectiveness

My results suggest that the improvement in female students' academic achievements is not driven by assessment bias. Previous literature offers two other classes of explanations (Dee, 2004). The first can be generalized as a role-model effect (Hoffmann and Oreopoulos, 2009). According to the role-model effect, an instructor's gender identity, by its very presence creates a role model that encourages students to exert more effort (King, 1993). The second class of explanations is that female teachers may be more effective in teaching female students. For example, female teachers may alter their pedagogical styles to meet the needs of female students, rendering female students more receptive to the material being taught.

Though it is difficult to disentangle the role-model effect from teacher effectiveness, previous literature identifies the role model effect as the main mechanism through which gender interaction benefits female students (Carrell et al., 2010; Hoffmann and Oreopoulos, 2009; Bettinger and Long, 2005). However, in this paper, I provide suggestive evidence that teacher effectiveness is also an important channel. I show this in two ways. First, I highlight some findings the role model effect cannot explain; the teacher-effectiveness channel seems to more plausibly explain those findings. Second, I borrow insights from the theoretical models in the literature, which differentiates between the role-model effect and teacher effectiveness in the case of same-race pairing Gershenson et al. (2018). Using their testable predictions, I find suggestive evidence that teacher effectiveness is an important channel through which same-sex teacher matching improves female students' academic outcomes.

5.3.1 Findings the Role-model Effect Cannot Explain

Gains from matching are negative for entry-level female teachers

As discussed above, Table XI shows that the interaction effect is negative when entry level teachers teach courses and is successively higher for assistant or associate professors and full professors. While one can argue that senior/experienced female teachers are better role models (Breda et al., 2020), the role-model effect cannot explain why female students perform worse when they are paired with entry-level female teachers. It seems unlikely that entry-level/young female teachers create a negative role model for female students, such that students exert less effort. If anything, female students might be encouraged more by a younger female teacher in a developing country, since they see a female who is marginally older overcoming the gender norms and stereotypes to become a teacher in an elite university. A more plausible explanation might be that female teachers with more teaching experience have learned over time how to bring the best out of the female students and incorporate this in their pedagogical style (Ladson-Billings, 1995; Foster, 1990).

Gains from matching are higher for below-median students

As seen earlier in Figure 2, the gain from female teacher-student matching is higher for students in the middle tier of the normalized score distribution; the interaction effect is actually negative for students who are above the 80th percentile. Again, the role-model effect cannot explain why top-performing female students perform worse when matched with female teachers. A more plausible explanation might be that female teachers target average/low-performing female students (Egalite et al., 2015). It is possible that female teachers design their lectures with average female students in mind, and therefore those female students perform better from gender matching.

Gains from matching in math, statistics, and econometrics are significantly lower than all other courses

Table IV shows that the gains from matching in math-intensive courses such as math, statistics, and econometrics are significantly lower than in the remaining courses. Earlier literature suggests that there is a lack of female role models in math-intensive courses at the college level (Freeman, 2004). Thus, if the role-model effect is the key mechanism, gains from matching should be higher in math courses, in which the dearth of female role models is greater than in social science courses. The literature on racial interaction suggests that black teachers help black students by incorporating *culturally relevant pedagogies* (LadsonBillings, 1995) and teaching *hidden curricula* (Foster, 1990).[16] To the extent that employing culturally relevant pedagogies or teaching hidden curricula is easier in social science courses, the teacher-effectiveness channel can explain the lower gains from matching in math-intensive courses.

5.3.2 A Formal Test to Distinguish the Role-Model Effect from Teacher Effectiveness

To formally distinguish between the role-model effect and teacher-effectiveness mechanism, we use the testable implications of the model proposed by Gershenson et al. (2018), which distinguishes between these two mechanisms in the case of same-race student-teacher interaction. They posit an education production function with two inputs and two periods of investment. The two inputs—student investment (effort) and teacher quality—determine educational achievement. Guided by the pedagogical literature, they allow black teachers to raise the achievement levels of the black students more than white teachers to capture the teacher-effectiveness channel (Ladson-Billings, 1995; Foster, 1990). In order to incorporate the role-model effect, Gershenson et al. (2018) assume black students have incorrect prior beliefs about the education production function by underestimating the returns to effort. They argue that black teachers are role models in the sense of providing "signals" to students about the true return to effort that leads students to update their beliefs about the education production function and increase their effort (investment). Again, they are guided by earlier studies that model the role-model effect as shifting beliefs about a parameter in a production function (Steele, 1997; Cunha et al., 2013; Papageorge et al., 2020).

[16] See Gershenson et al. (2018) for a detailed review.

According to the model of Gershenson et al. (2018), the effectiveness channel suggests that two black teachers would have a larger impact than one. In contrast, the role-model effect suggests that a second teacher may have a smaller effect if crucial information regarding returns to human capital has been transmitted by the first black teacher. Their model formalizes this insight and generates the following testable implications using *dosage models*, i.e., the marginal effect of having additional female teachers: If the teacher-effectiveness channel is the key mechanism, we should not observe diminishing returns to additional female teachers. On the other hand, if the role-model effect is the main mechanism, we would expect to observe diminishing returns.

I use the testable implications of Gershenson et al. (2018) about the marginal effect of being assigned an additional female teacher to distinguish between these two potential mechanisms. To this end, I run a variant of equation (4) to include the square of the percentage of female teachers and its interaction with the female student dummy. The results in Table X show that the interaction of the square of the proportion of female teachers with the female student dummy is positive and statistically significant for cumulative grade point average (CGPA), and the likelihood of enrolling in an economics master's program. I do not find any significant diminishing returns to having additional female teachers on other longer-term outcomes, such as the likelihood of graduating with the regular cohort or number of years repeated. Thus, according to the testable implications described earlier, teacher effectiveness is an important channel through which female students benefit from having female teachers.

6 Conclusion

This study is the first to estimate the causal effects of female student-teacher interactions at the college level in the context of developing countries. My focus on mandatory courses, along with the limited ability of students to select teachers, allows me to attain a near-random setting to estimate the causal effects. I find that matching female students with female teachers significantly improves academic performance. Further, exposure to more female teachers improves female students' medium-term outcomes—such as cumulative GPA and the likelihood of finishing the degree with the regular cohort—and longer-term outcomes, such as the likelihood of enrolling in an economics master's program. I show that this improvement is not driven by female teachers' bias toward female students when grading.

My findings are consistent with previous studies that find same-gender teacher assignment matters at the college level. However, earlier studies that estimate female student-teacher interaction at the college level are in the context of developed countries, and to the best of my knowledge none focus on developing countries. The question of the role of female teachers in bridging the gender gap in educational attainment is much more salient in developing countries, in which stereotypes and prejudice are much more prevalent (Jayachandran, 2015) and gender gaps in college enrollment and attainment much larger (Ilie and Rose, 2016). In this paper, I document that—as is the case in developed countries—gender interaction also matters in developing countries, despite the differences in context. Prior research indicates that female students gain from gender matching in schools in developing countries (Muralidharan and Sheth, 2016; Rawal et al., 2010). I show that in a high-stakes setting such as college, this interaction continues to be beneficial for female students.

I document several other important results in my paper. First, gains from female student-teacher matching are higher in social science courses than math-intensive courses.

Second, I show that the interaction effects are higher for teachers with more experience. Third, below-median female students benefit more from having female teachers. Finally, exposure to female teachers solely in the first semester does not improve female students' long-term outcomes. Rather, continuous exposure to female teachers is what matters for improving female students' long-term academic achievement.

I also provide suggestive evidence that female teacher effectiveness in teaching female students is an important channel through which same-sex teacher assignment improves female students' academic achievement, in addition to the oft-cited role-model effect (Carrell et al., 2010; Hoffmann and Oreopoulos, 2009; Bettinger and Long, 2005). This finding has important policy implications. If the role model is the key mechanism, as suggested by prior literature, then the only way to benefit female students is to appoint more female teachers in college. On the other hand, if teachers' experience or qualifications play a role, which suggest that pedagogical style is at play, then female students can benefit if college authorities acquaint male teachers with relevant pedagogical techniques that render female teachers more effective in teaching female students. Further, educating male teachers about stereotype threats can reduce their gender stereotypes (Hill et al., 2010) and cause them to be more accessible to female students.

Finally, I show that the benefit of gender matching in college is not limited to short and medium-term outcomes but also extends to longer-term outcomes such as the probability of enrolling in an economics graduate program. This finding is of significant relevance. Female underrepresentation in graduate economics courses has attracted great attention recently (Lundberg and Stearns, 2019). My results suggest that appointing more qualified female teachers can be an important tool for reducing the gender gap in a male-dominated field such as economics without affecting the performance of males.

References

Ambady, N., Shih, M., Kim, A., and Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological science*, 12(5):385–390.

Aronson, J., Quinn, D. M., and Spencer, S. J. (1998). Stereotype threat and the academic underperformance of minorities and women. In *Prejudice*, pages 83–103. Elsevier.

Attiyeh, G. and Attiyeh, R. (1997). Testing for bias in graduate school admissions. *The Journal of Human Resources*, 32(3):524.

Bar, T. and Zussman, A. (2012). Partisan grading. *American Economic Journal: Applied Economics*, 4(1):30–48.

Bettinger, E. P. and Long, B. T. (2005). Do faculty serve as role models? The impact of instructor gender on female students. *American Economic Review*, 95(2):152–157.

Blank, R. M. (1991). The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from the American Economic Review. *American Economic Review*, 81(5):1041–1067.

Breda, T., Grenet, J., Monnet, M., and Van Effenterre, C. (2020). Do female role models reduce the gender gap in science? evidence from french high schools.

Burgess, S. and Greaves, E. (2013). Test scores, subjective assessment, and stereotyping of ethnic minorities. *Journal of Labor Economics*, 31(3):535–576.

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics*, 134(3):1163–1224.

Carrell, S. E., Page, M. E., and West, J. E. (2010). Sex and Science: How Professor Gender Perpetuates the Gender Gap*. *The Quarterly Journal of Economics*, 125(3):1101–1144.

Cornwell, C., Mustard, D. B., and Van Parys, J. (2013). Noncognitive skills and the gender disparities in test scores and teacher assessments: Evidence from primary school. *Journal of Human Resources*, 48(1):236–264.

Croizet, J.-C. and Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin*, 24(6):588–594.

Cunha, F., Elo, I., and Culhane, J. (2013). Eliciting maternal expectations about the technology of cognitive skill formation. Technical report, National Bureau of Economic Research.

Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *The Review of Economics and Statistics*, 86(1):195–210.

Egalite, A. J., Kisida, B., and Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, 45:44–52.

Falch, T. and Naper, L. R. (2013). Educational evaluation schemes and gender gaps in student achievement. *Economics of Education Review*, 36:12 – 25.

Filer, R. K. (1983). Sexual differences in earnings: The role of individual personalities and tastes. *Journal of Human Resources*, 18(1).

Foster, M. (1990). The politics of race: Through the eyes of african-american teachers. *Journal of education*, 172(3):123–141.

Freeman, C. E. (2004). Trends in educational equity of girls & women: 2004. nces 2005-016. *National Center for Education Statistics*.

Gershenson, S., Hart, C. M. D., Hyman, J., Lindsay, C., and Papageorge, N. W. (2018). The long-run impacts of same-race teachers. Working Paper 25254, National Bureau of Economic Research.

Goldin, C., Katz, L. F., and Kuziemko, I. (2006). The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic perspectives*, 20(4):133–156.

Goldin, C. and Rouse, C. (2000). Orchestrating impartiality: The impact of ”blind” auditions on female musicians. *American Economic Review*, 90(4):715–741.

Green, M. (2007). Science and engineering degrees: 1966-2004. *Virginia: National Science Foundation*.

Hanna, R. N. and Linden, L. L. (2012). Discrimination in grading. *American Economic Journal: Economic Policy*, 4(4):146–68.

Hill, C., Corbett, C., and St Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics.* ERIC.

Hinnerich, B. T., Ho¨glin, E., and Johannesson, M. (2011). Are boys discriminated in Swedish high schools? *Economics of Education review*, 30(4):682–690.

Hoffmann, F. and Oreopoulos, P. (2009). A professor like me: The influence of instructor gender on college achievement. *The Journal of Human Resources*, 44(2):479–494.

Ilie, S. and Rose, P. (2016). Is equal access to higher education in South Asia and sub-Saharan Africa achievable by 2030? *Higher Education*, 72(4):435–455.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *Annual Review of Economics*, 7(1):63–88.

Jones, E. B. and Jackson, J. D. (1990). College grades and labor market rewards. *The Journal of Human Resources*, 25(2):253.

King, S. H. (1993). The limited presence of African-American teachers. *Review of Educational Research*, 63(2):115–149.

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American educational research journal*, 32(3):465–491.

Lavy, V. (2008). Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment. *Journal of Public Economics*, 92(10):2083 – 2105.

Lindahl, E. (2016). Are teacher assessments biased? Evidence from Sweden. *Education Economics*, 24(2):224–238.

Lundberg, S. and Stearns, J. (2019). Women in economics: Stalled progress. *Journal of Economic Perspectives*, 33(1):3–22.

Luoto, J., McIntosh, C., and Wydick, B. (2007). Credit information systems in less developed countries: A test with microfinance in Guatemala. *Economic Development and Cultural Change*, 55(2):313–334.

Muralidharan, K. and Sheth, K. (2016). Bridging education gender gaps in developing countries: The role of female teachers. *Journal of Human Resources*, 51(2):269–297.

Neumark, D. and Gardecki, R. (1998). Women helping women? Role model and mentoring effects on female Ph.D. students in economics. *The Journal of Human Resources*, 33(1):220–246.

Nguyen, H.-H. D. and Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of applied psychology*, 93(6):1314.

Papageorge, N. W., Gershenson, S., and Kang, K. M. (2020). Teacher expectations matter. *Review of Economics and Statistics*, 102(2):234–251.

Rawal, S., Kingdon, G., et al. (2010). Akin to my teacher: Does caste, religious or gender distance between student and teacher matter? some evidence from india. Technical report, Department of Quantitative Social Science-UCL Institute of Education, University College London.

Rothstein, D. S. (1995). Do female faculty influence female students' educational and labor market attainments? *Industrial and Labor Relations Review*, 48(3):515–530.

Spencer, S. J., Steele, C. M., and Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of experimental social psychology*, 35(1):4–28.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American psychologist*, 52(6):613.

Summers, L. (2005). Diversifying the science and engineering workforce. In *Discussion presented at the NBER Conference on Diversifying the Science and Engineering Workforce, Cambridge, MA*.

Vincent-Lancrin, S. (2008). The reversal of gender inequalities in higher education: An on-going trend. *Higher education to*, 2030(1):265–298.

Wise, D. A. (1975). Academic achievement and job performance. *The American Economic Review*, 65(3):350–366.

Xie, Y. and Shauman, K. A. (2003). *Women in science*. Harvard university press.

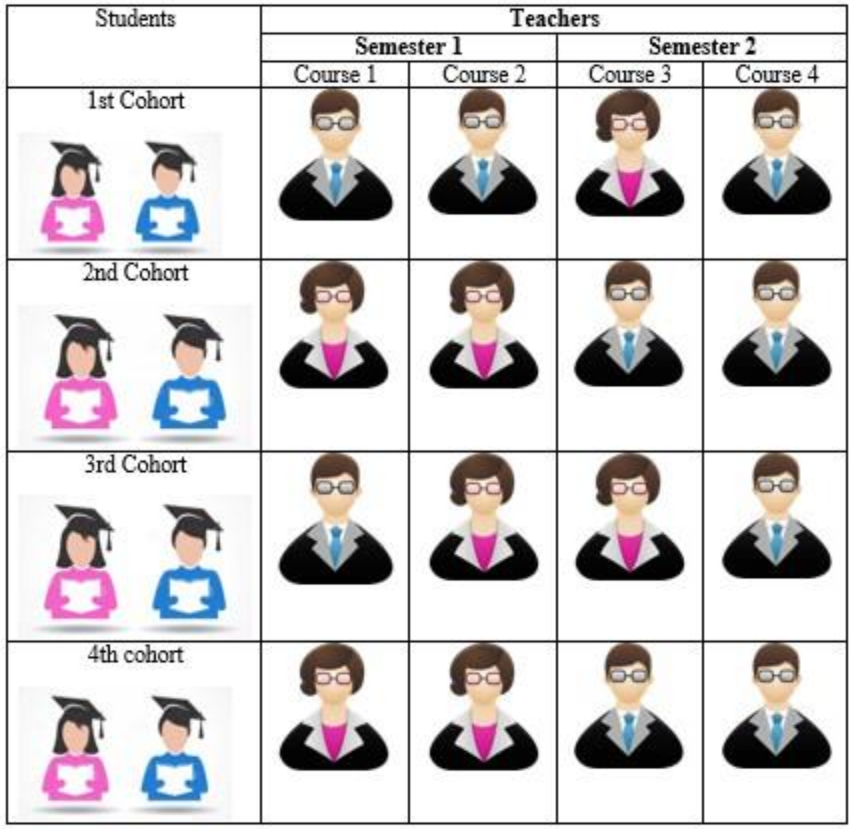


Figure 1: Hypothetical match between students and teachers across cohorts

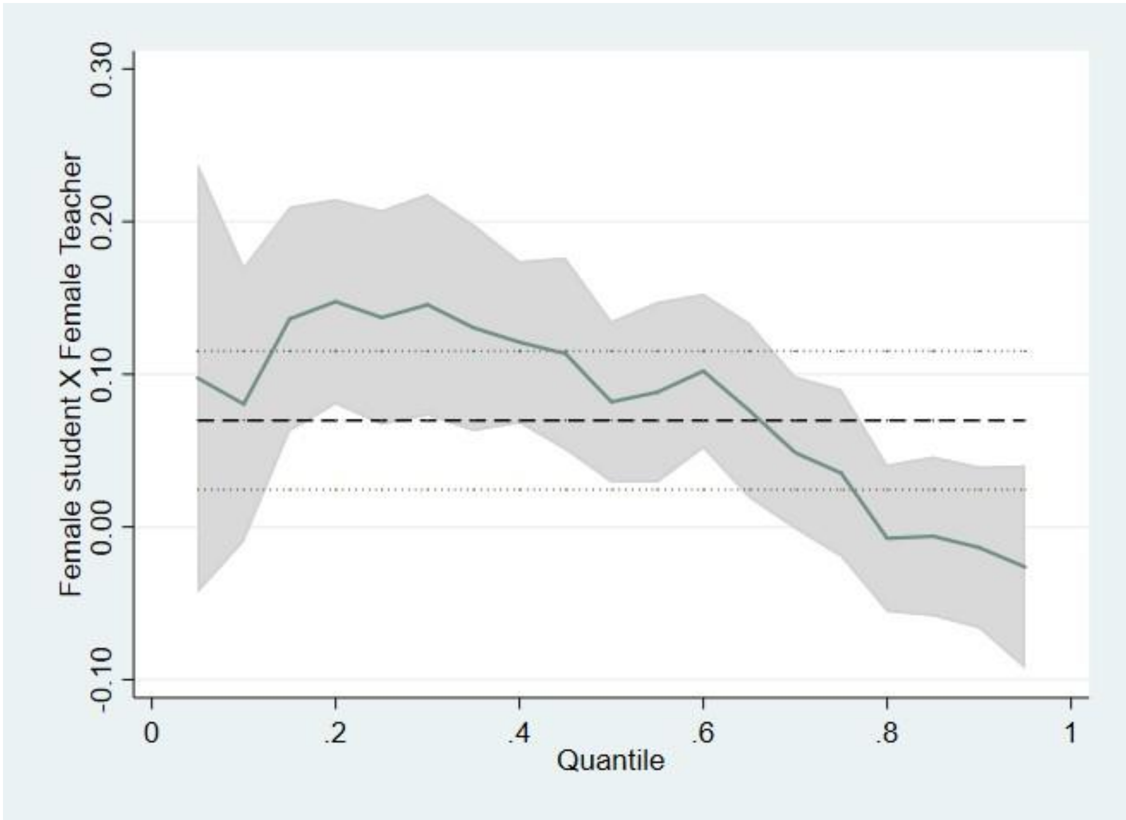


Figure 2: Estimates of the interaction effects at different quantiles of Normalized score
Notes: Estimates are from a quantile regression of student teacher interaction on normalized test scores.

Table I: Summary Statistics

	Number of observations	Male	Female
Student Profile			
Total students	1508	939 (62%)	569 (38%)
Average CGPA	1502	2.95	3.14
Average CGPA Conditional on Passing all courses	1280	3.18	3.27
Passed with Regular Batch	941	516 (55%)	425 (75%)
Number of years repeated			
Once	215	156	59
Twice	56	45	11
Thrice	6	4	2
Teacher Profile			
Total Teachers	66	46 (70%)	20 (30%)
Have a Terminal PhD Degree	30	20 (43%)	10 (50%)
Course offered			
No of courses taught by	402	246 (61%)	156 (39%)
Number of mandatory courses taught by	243	142 (58%)	101 (42%)
Number of courses taught by a instructor with a PhD Degree	212	133 (63%)	79 (37%)
Number of mandatory courses taught by a instructor with a PhD Degree	114	68 (60%)	46 (40%)
Number of first-year mandatory courses taught by	86	73 (85%	13 (15%)

Table II: Comparison of Mean Characteristics

	Male Teacher	Female Teacher	Difference	P-value	Observation
Panel A: Female Student					
Merit serial in entrance exam	215.6	210.9	4.73	0.42	13169
Merit score in entrance exam	147.4	147.5	-0.1	0.89	13169
High school degree was science	0.459	0.453	0.005	0.7	13169
Had quota	0.081	0.08	0.001	0.76	13169
HSC GPA	4.393	4.385	0.008	0.49	2987
SSC GPA	4.5	4.47	0.03	0.51	2987
Took math in HSC	0.333	0.317	0.016	0.56	3262
Took Statistics in HSC	0.141	0.147	-0.006	0.56	3262
Panel B: Male Student					
Merit serial in entrance exam	230.47	218.83	11.63	0.09	22438
Merit score in entrance exam	146.21	146.63	-0.43	0.56	22438
High school degree was science	0.49	0.48	0.01	0.23	22438
Had quota	0.068	0.065	0.003	0.36	22438
HSC GPA	4.211	4.21	0.001	0.93	5556
SSC GPA	4.44	4.43	0.01	0.54	5556
Took math in HSC	0.284	0.272	0.011	0.56	5844
Took Statistics in HSC	0.203	0.2	0.003	0.56	5844

Notes: Unit of observation is individual student-course. We have partial data (only two out of 11 cohorts) on the last four variables of each panel. HSC stands for Higher secondary Exam, and SSC stands for Secondary School exam. HSC and SSC exams are national level exam that takes place after 12 years and 10 years of schooling respectively. The P-values are adjusted for intra-classroom correlation.

Table III: Effect of Instructor's Gender on Students' Academic Performance

	Dependent variable: Standardized test score						
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female teacher x Female student	0.0681**	0.0494*	0.0877***	0.0869***	0.0862***	0.0864***	0.0935***
	(0.0272)	(0.0260)	(0.0246)	(0.0248)	(0.0248)	(0.0248)	(0.0256)
Student characteristics at entry	No	Yes	No	No	No	No	No
Student fixed effects	No	No	Yes	Yes	Yes	Yes	Yes
Teacher fixed effects	No	No	No	Yes	Yes	Yes	Yes
Course fixed effects	No	No	No	No	Yes	Yes	Yes
Teacher gender x Course fixed effects	No	No	No	No	No	Yes	Yes
Classroom fixed effects	No	No	No	No	No	No	Yes
Observations	31,613	27,853	31,613	31,613	31,613	31,613	31,613
R-squared	0.013	0.029	0.571	0.575	0.576	0.576	0.584

Notes: Unit of observation is individual student-course. Each column presents estimated coefficients from a linear fixed effect regression for students' academic performance. Dependent variables are standardized scores. Standard errors are corrected for classroom (cohort-semester-course) level clustering and are presented in parentheses. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table IV: Effect of Instructor's Gender on Students' Academic Performance by Subject Type

	Dependent Variable: Standardized Score	
	Math, Statistics, and Econometrics	Other courses
VARIABLES	(1)	(2)
Female Teacher x Female Student	0.0596*	0.123***
	(0.0346)	(0.0364)
Student fixed effects	Yes	Yes
Teacher fixed effects	Yes	Yes
Course fixed effects	Yes	Yes
Teacher gender x Course fixed effects	Yes	Yes
Classroom fixed effects	Yes	Yes
Observations	11,333	20,262
R-squared	0.675	0.571

Notes: Unit of observation is individual student-course. Each column presents estimated coefficients from a linear fixed effect regression for students' academic performance. Dependent variables are standardized scores. Standard errors are corrected for classroom (cohort-semester-course) clustering and are presented in parentheses. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table V: Effect of Instructor's Gender on Long-term Outcomes

	Bachelor Degree			Master's degree	
	Cumulative grade point average	Passed with regular cohort	Number of years repeated	Enrolled	Cumulative grade point average
VARIABLES	(1)	(2)	(3)	(4)	(5)
Female student	-0.0787	-0.0214	0.124	-0.411*	0.210
	(0.131)	(0.0987)	(0.116)	(0.233)	(0.414)
% of female teachers in all mandatory courses	0.0253***	0.0141***	-0.00750***	0.0134***	0.00279
	(0.00186)	(0.00140)	(0.00165)	(0.00249)	(0.00971)
% of female teachers in all mandatory courses x Female student	0.00549*	0.00459**	-0.00585**	0.0114*	-0.00433
	(0.00291)	(0.00219)	(0.00258)	(0.00598)	(0.0106)
Observations	1,496	1,496	1,496	654	524
R-squared	0.239	0.193	0.085	0.354	0.513
Male student mean	2.951	0.550	0.279	0.748	3.129

Notes: Unit of Analysis is individual student. The regression follows equation (4) as specified in the main text. Information on Master's degree was available only for the first four cohort of students. All the regressions include cohort fixed effects. Standard errors are corrected for cohort-level clustering and are presented in parentheses. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table VI: Effect of Instructor's Gender on Students' Academic Performance in Blind and Non-blind Exams

	Dependent Variable: Standardized Score	
	Non-blind exam	Blind exam
VARIABLES	(1)	(2)
Female Teacher x Female Student	0.0872***	0.0664**
	(0.0259)	(0.0286)
Student fixed effects	Yes	Yes
Teacher fixed effects	Yes	Yes
Course fixed effects	Yes	Yes
Teacher gender x Course fixed effects	Yes	Yes
Classroom fixed effects	Yes	Yes
Observations	31,930	31,571
R-squared	0.484	0.474

Notes: Unit of analysis is individual student-course. Each column presents estimated coefficients from a linear fixed effect regression for students' academic performance. Dependent variables are standardized scores. Standard errors are corrected for classroom (cohort-semester-course) clustering and are presented in parentheses. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table VII: Effect of Instructor's Gender on Students' Academic Performance with Blind Exam Interaction

	Dependent variable: Standardized test score	
VARIABLES	(1)	(2)
Female Teacher x Female Student	0.0926***	0.0861***
	(0.0263)	(0.0259)
Blind Exam	-0.00886	
	(0.00875)	
Female Student x Blind Exam	0.000788	
	(0.0220)	
Female Teacher x Blind Exam	0.0116	0.00298
	(0.0125)	(0.0125)
Female Teacher x Female Student x Blind Exam	-0.0322	-0.0218
	(0.0310)	(0.0300)
Student fixed effects	Yes	Yes
Teacher fixed effects	Yes	Yes
Course fixed effects	Yes	Yes
Teacher gender x Course fixed effects	Yes	Yes
Classroom fixed effects	Yes	Yes
Student x Exam type Fixed effects	No	Yes
Observations	63,502	63,499
R-squared	0.461	0.484

Notes: Unit of analysis is student-course-exam type. Each column presents estimated coefficients from a linear fixed effect regression for students' academic performance. Dependent variables are standardized scores. Standard errors are corrected for classroom (cohort-semester-course) clustering and are presented in parentheses. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table VIII: Teachers' Qualifications and Gender Interaction

	Normalized test score		
Teachers' characteristics	Rank	Experience	Has a Ph.D. degree
	(1)	(2)	(3)
Panel A: All compulsory courses			
Female student x Female teacher	-0.0962*	-0.0458	0.0861**
	(0.0575)	(0.0420)	(0.0343)
Female student x Female teacher x Teacher Characteristics	0.0766***	0.0103***	0.0264
	(0.0215)	(0.00259)	(0.0489)
Observations	31,499	31,499	31,499
Panel B: Math, Statistics, and Econometrics			
Female student x Female teacher	-0.188**	-0.125**	-0.0631
	(0.0888)	(0.0615)	(0.0710)
Female student x Female teacher x Teacher Characteristics	0.111***	0.0195***	0.217**
	(0.0394)	(0.00625)	(0.0911)
Observations	11,333	11,333	11,333
Panel C: Other courses			
Female student x Female teacher	-0.0339	0.0156	0.133***
	(0.0836)	(0.0591)	(0.0482)
Female student x Female teacher x Teacher Characteristics	0.0616**	0.00800**	-0.0301
	(0.0301)	(0.00335)	(0.0659)
Observations	20,148	20,148	20,148

Notes: Unit of analysis is individual student-course. Each column presents estimated coefficients from a linear fixed effect regression for students' academic performance. Dependent variables are standardized scores. Standard errors are corrected for classroom (teacher-course-cohort) clustering and are presented in parentheses. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table IX: Students' Qualifications and Gender Interaction

	Normalized test score			
Student Characteristics at entry	Merit Serial	Merit score	High school discipline is science	Admitted under quota
	(1)	(2)	(3)	(4)
Panel A: All compulsory courses				
Female student x Female teacher	0.0808***	0.118**	0.0739***	0.0836***
	(0.0248)	(0.0493)	(0.0250)	(0.0245)
Female student x Female teacher x Student Characteristics	3.13e-05	-0.000229	0.0313	0.0510
	(3.73e-05)	(0.000283)	(0.0298)	(0.0466)
Observations	31,613	31,613	31,613	31,613
Panel B: Math, Science, and Statistics				
Female student x Female teacher	0.0523	0.117	0.0479	0.0456
	(0.0361)	(0.0756)	(0.0316)	(0.0338)
Female student x Female teacher x Student Characteristics	7.51e-06	-0.000460	0.0142	0.128*
	(6.24e-05)	(0.000486)	(0.0489)	(0.0687)
Observations	11,333	11,333	11,333	11,333
Panel C: Other courses				
Female student x Female teacher	0.109***	0.105	0.107***	0.116***
	(0.0350)	(0.0658)	(0.0353)	(0.0349)
Female student x Female teacher x Student Characteristics	4.22e-05	9.00e-05	0.0251	0.0154
	(4.54e-05)	(0.000358)	(0.0327)	(0.0626)
Observations	20,262	20,262	20,262	20,262

Notes: Unit of analysis is individual student-course. Each column presents estimated coefficients from a linear fixed effect regression for students' academic performance. Dependent variables are normalized test scores. Standard errors are corrected for classroom (teacher-course-cohort) clustering and are presented in parentheses. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table X: Effect of Marginal Increase in Female Teachers on the Medium- and Longer-term Outcomes

	(1)	(2)	(3)	(4)	(5)
VARIABLES	CGPA	Passes with regular batch	Number of years repeated	Enroll in master's program	Masters CGPA
Female student	0.251	-0.279	0.0490	0.0777	2.702
	(0.253)	(0.188)	(0.225)	(0.382)	(8.987)
% of female teachers in all mandatory courses	0.0610***	0.0275***	0.00269	0.0442***	0.0913
	(0.00573)	(0.00425)	(0.00510)	(0.00945)	(0.315)
% of female teachers in all mandatory courses **squared**	-0.000508***	-0.000200***	-0.000150**	-0.000419**	-0.00113
	(7.61e-05)	(5.63e-05)	(6.77e-05)	(0.000179)	(0.00406)
% of female teachers in all mandatory courses x	-0.0134	0.0154*	-0.00311	-0.0257	-0.135
Female student	(0.0114)	(0.00846)	(0.0102)	(0.0226)	(0.461)
% of female teachers in all mandatory courses **squared**	0.000241*	-0.000110	-3.22e-05	0.000652*	0.00174
x Female student	(0.000132)	(9.75e-05)	(0.000117)	(0.000368)	(0.00588)
Observations	1,463	1,463	1,463	629	507
R-squared	0.259	0.214	0.095	0.171	0.030
Male student mean	2.945	0.556	0.290	0.755	3.124

Notes: Unit of Analysis is individual student. The regression is a modified version of equation (4) as specified in the main text. Information on Master1's degree was available only for the first four cohort of students. All the regressions include cohort-fixed effect. Standard errors are corrected for Cohort-level clustering and are presented in parentheses. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table XI: Teachers' Qualifications and Gender Interaction

	Dependent variable: Standardized test score							
	Rank			Teaching experience			PhD status	
VARIABLES	Lecturer	Assistant/ Associate Professor	Professor	<10 years	10-20 years	>20 years	Doesn't have a PhD degree	Has a PhD degree
Female Teacher x Female Student	-0.126**	0.144***	0.208***	-0.0403	0.195***	0.245***	0.0926***	0.120***
	(0.0596)	(0.0357)	(0.0562)	(0.0465)	(0.0400)	(0.0754)	(0.0353)	(0.0384)
Student fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Course fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher gender x Course fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Classroom fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,448	12,836	10,107	15,322	7,164	8,794	17,207	14,246
R-squared	0.667	0.635	0.583	0.633	0.628	0.630	0.623	0.578

Notes: Unit of analysis is individual student-course. Each column presents estimated coefficients from a linear fixed effect regression for students' academic performance. Dependent variables are standardized scores. Standard errors are corrected for classroom (cohort-semester-course) clustering and are presented in parentheses. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table XII: Effect of Instructor's Gender on long term-outcomes

	Bachelor Degree			Master's degree	
	Cumulative grade point average	Passed with regular cohort	Number of years repeated	Enrolled	Cumulative grade point average
VARIABLES	(1)	(2)	(3)	(4)	(5)
Female student	0.208***	0.237***	-0.169***	0.0268	0.0992**
	(0.0493)	(0.0362)	(0.0416)	(0.0456)	(0.0493)
% of female teachers in 1st year mandatory courses	0.00282	0.00486***	-4.94e-05	0.00176	0.00475
	(0.00223)	(0.00164)	(0.00188)	(0.00290)	(0.00405)
% of female teachers in 1st year mandatory courses x Female student	-0.00236	-0.00345**	0.00166	0.000197	-0.00415
	(0.00232)	(0.00171)	(0.00196)	(0.00239)	(0.00261)
Observations	1,485	1,485	1,485	650	524
R-squared	0.120	0.120	0.062	0.317	0.516
Male student mean	2.951	0.550	0.279	0.748	3.129

Notes: Unit of analysis is individual student. Each column presents estimated coefficients from a linear fixed effect regression for different long-term outcome variables. Information on Master's degree was available only for the first four cohort of students. All the regressions include cohort fixed effects. Standard errors are corrected for batch-level clustering and are presented in parentheses. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Appendix Figures and Tables

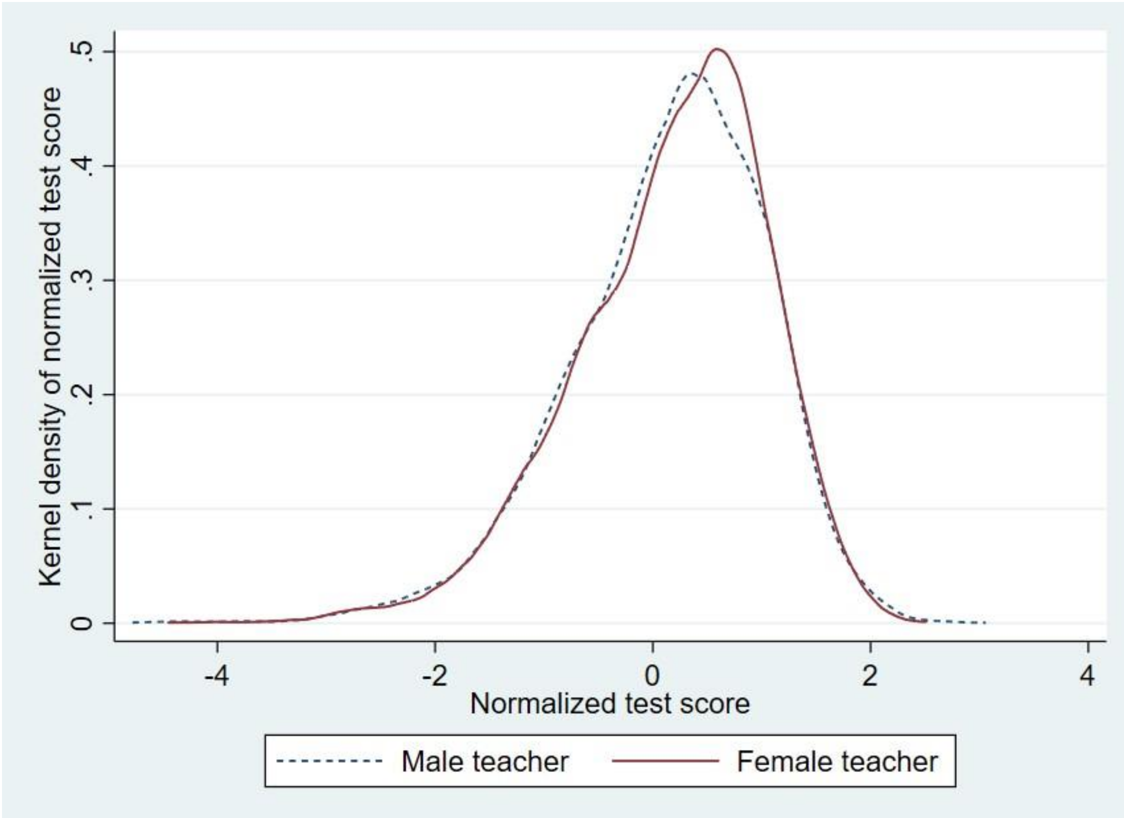


Figure A.1: Distribution of female students' test score by the gender of the teacher

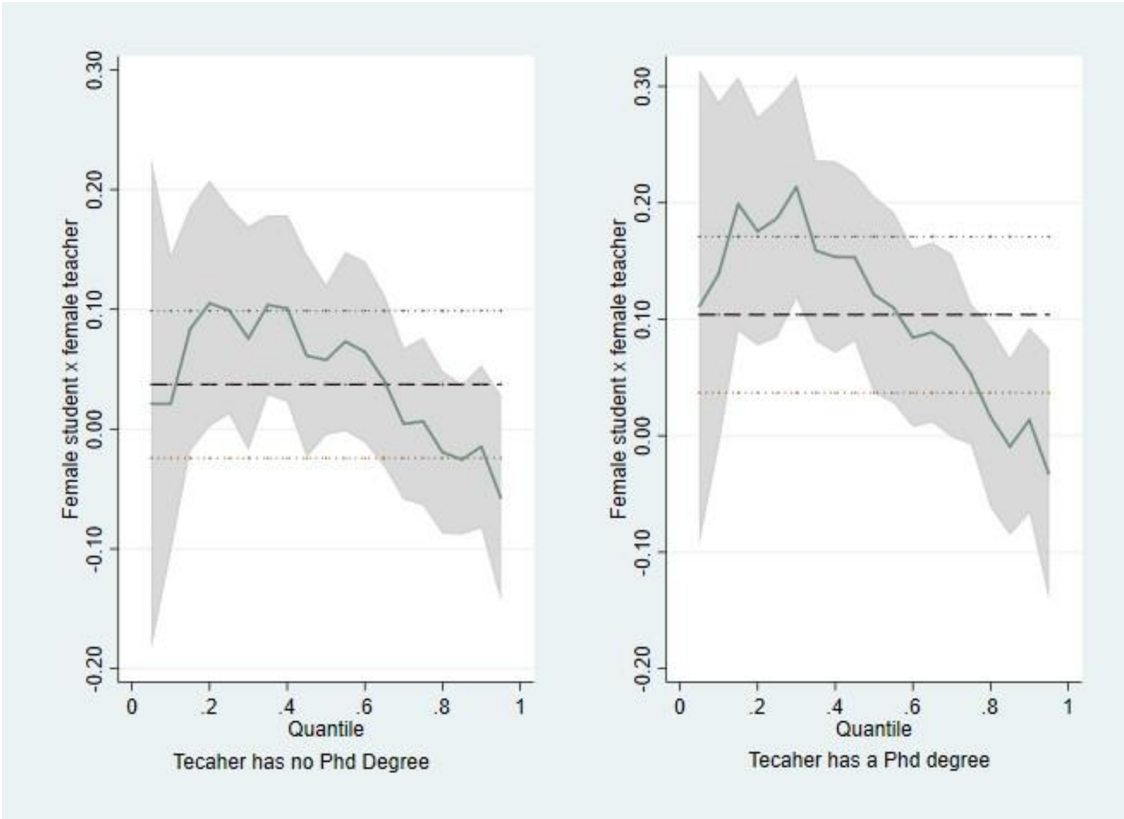


Figure A.2: Quantile Treatment Effect: By PhD Status of the Instructor

Table A.1: Grading System Used in the Department of Economics

Numerical Marks	**Grade**	**Grade Point**
80 to 100	A+	4.00
75 to 79	A	3.75
70 to 74	A-	3.50
65 to 69	B+	3.25
60 to 64	B	3.00
55 to 59	B-	2.75
50 to 54	C+	2.50
45 to 49	C	2.25
40 to 44	D	2.00
< 40	F	0.00

Note: Minimum GPA required for promotion is 2.00

Table A.2: Effect of Female Student Dummy on the Proportion of Female Teachers in Compulsory Courses in Different Semesters

	Proportion of female teachers in mandatory courses							
	1st semester	2nd semester	3rd semester	4th semester	5th semester	6th semester	7th semester	8th semester
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female student	-0.00446	-0.00205	-0.00477	0.00485	0.00205	0.00349	-0.000844	0.00106
	(0.00477)	(0.00769)	(0.00553)	(0.00281)	(0.00472)	(0.00525)	(0.00403)	(0.00499)
Constant	0.150***	0.168***	0.759***	0.602***	0.577***	0.731***	0.316***	0.373***
	(0.00182)	(0.00297)	(0.00218)	(0.00111)	(0.00191)	(0.00214)	(0.00169)	(0.00209)
Observations	1,483	1,429	1,265	1,239	1,097	1,082	928	691
R-squared	0.789	0.805	0.912	0.892	0.921	0.964	0.967	0.970

Notes: Unit of analysis is individual student. Each column presents estimated coefficients from a linear fixed effect regression of a dummy variable for female student in semester t on the proportion of teachers in compulsory courses in semester t. Each regression includes cohort fixed effect. Standard errors are corrected for cohort level clustering and are presented in parentheses. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.3: Effect of Female Student's Academic Performance on the Proportion of Female Teachers in Compulsory Courses in the Subsequent Semester

	Proportion of female teachers in mandatory courses						
	2nd semester	3rd semester	4th semester	5th semester	6th semester	7th semester	8th semester
VARIABLES	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CGPA of female students in the preceding semester	0.000762	0.00670	-0.00647	-0.00480	0.00854	0.00861	0.0112
	(0.0154)	(0.0154)	(0.00415)	(0.00667)	(0.0123)	(0.00775)	(0.00802)
Constant	0.163***	0.734***	0.627***	0.597***	0.707***	0.282***	0.335***
	(0.0485)	(0.0500)	(0.0134)	(0.0219)	(0.0406)	(0.0257)	(0.0264)
Observations	551	499	492	444	441	389	289
R-squared	0.877	0.929	0.926	0.972	0.988	0.991	0.992

Notes: Unit of analysis is individual student. Each column presents estimated coefficients from a linear fixed effect regression of female students' academic performance in semester $t-1$ on the proportion of teachers in compulsory courses in semester t. Each regression includes cohort fixed effect. Standard errors are corrected for cohort level clustering and are presented in parentheses. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.4: Effect of Instructor's Gender on Students' Academic Performance in Single Blind and Double-blind Exams

	Dependent Variable: Standardized Score	
	Double-blind exam	Single-blind exam
VARIABLES	(1)	(2)
Female Teacher x Female Student	0.0745***	0.0617**
	(0.0283)	(0.0274)
Student fixed effects	Yes	Yes
Teacher fixed effects	Yes	Yes
Course fixed effects	Yes	Yes
Teacher gender x Course fixed effects	Yes	Yes
Classroom fixed effects	Yes	Yes
Observations	14,050	17,520
R-squared	0.451	0.502

Notes: Unit of analysis is individual student-course. Each column presents estimated coefficients from a linear fixed effect regression for students' academic performance. Dependent variables are standardized scores. Standard errors are corrected for classroom (cohort-semester-course) clustering and are presented in parentheses. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.5: Effect of Instructor's Gender on Students' Academic Performance by Entry Score Quantile

	Dependent Variable: Standardized Score			
	1st Quantile	2nd quantile	3rd quantile	4th quantile
VARIABLES	(1)	(2)	(3)	(4)
Female Teacher x Female Student	0.0731*	0.117***	0.0561	0.0614*
	(0.0427)	(0.0364)	(0.0458)	(0.0335)
Student Fixed Effect	Yes	Yes	Yes	Yes
Cohort-Semester Fixed Effect	Yes	Yes	Yes	Yes
Teacher Fixed Effect	Yes	Yes	Yes	Yes
Course Fixed Effect	No	Yes	No	Yes
Observations	6,673	6,761	7,018	7,390
R-squared	0.526	0.564	0.600	0.644

Notes: Unit of analysis is student-course. Each column presents estimated coefficients from a linear fixed effect regression of students' academic performance at different entry score quantiles. Dependent variables are standardized scores. Standard errors are corrected for classroom (cohort-semester-course) clustering and are presented in parentheses. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Chapter 2: Unintended Consequences of a Well-Intentioned Policy: Impact of Credit on Child Labor in Bangladesh*

Md Amzad Hossain*

Abstract

Canonical models of credit suggest that relaxing credit constraints can increase human capital investment in children. However, when credit brings new business opportunities within reach in economies with many labor market frictions, increased access to credit might increase the use of child labor by increasing the opportunity cost of a child's time. In this study, I use data from a randomized controlled experiment to examine the effect of an agricultural credit expansion program in Bangladesh and find an increase in child labor. I present evidence that this increase in child labor is due to a rise in new opportunities for children to work in household self-employment activities. I also find that treated households with fewer working adults use more child labor and spend less on education. While I do not see any effect on schooling outcomes, the time budget survey reveals that children from treated areas spend significantly less time studying. Overall, these findings raise concerns about the unintended intergenerational consequences of easing credit constraints to increase self-employment.

Keywords: Agricultural microcredit, child labor, human capital, unitary model of household
JEL codes: G21, J13, O15, J21

* I am grateful to Sheetal Sekhri, Sandip Sukhtankar and Leora Friedberg for advising me and providing many helpful comments and discussions. I would also like to thank Eric Chyn, David McKenzie, Ahmed Mushfiq Mobarak, Jonathan Colmer, Craig McIntosh, Teevrat Garg, Lee Lockwood, Marcos Rangel, Jonas Hjort, and seminar participants at the University of Virginia for providing many helpful suggestions. Mohammad Abdul Malek, who was the principal investigator of the Randomized Control Trial I used for my study, provided me with data that made this project possible.

*University of Dhaka, Bangladesh, and University of Virginia, Charlottesville, VA-22903. Email: mh2vh@virginia.edu. Phone number: +1-434-466-1313

1 Introduction

Child labor is a widespread phenomenon throughout the developing world. According to the International Labor Organization (ILO), a total of 152 million children were engaged in child labor in 2017. The problem is most pervasive in Africa, Asia, and the Pacific region, where nine out of every ten child laborers are located. To make things worse, a significant share of child laborers (around 73 million) work in hazardous conditions that pose a serious threat to their safety, health, and moral development. Additionally, child labor can take significant time away from education and schooling of the children, and can, therefore, reduce educational attainment (Beegle et al., 2008). Hence, the incidence of child labor might have severe intergenerational consequences (Edmonds, 2007).

Child labor has often been viewed as a consequence of poverty.[17] It is often argued that reducing market inefficiencies can promote general economic development, which in turn, may reduce child labor. One such important market inefficiency is the lack of access to credit, which can adversely affect entrepreneurship and trap people in a vicious cycle of poverty (Augsburg et al., 2015). In addition, such credit constraints can hamper investment on human capital (Lochner and Monge-Naranjo, 2012), and increase the incidence of child labor (Eswaran and Kotwal, 1986; Dehejia and Gatti, 2005). This is because, if a child's time is valuable in home or market production, liquidity-constrained parents might start withdrawing their children from school and engaging them in production activities to smooth household consumption over time (Jacoby, 1994). Microcredit programs can address this constraint on entrepreneurial activity, and thus have the potential to reduce child labor by increasing household income (Dehejia and Gatti, 2005).

However, when there are other frictions in the labor market, as is the case in most developing economies, credit can also have unintended consequences in the form of worsened child

[17] See for example Krueger (1996), and Fallon and Tzannatos (1998).

outcomes. For instance, microcredit often encourages households to initiate new entrepreneurial activities or expand existing businesses, which require labor. Yet hiring labor is costly, due to moral hazard (e.g., shirking) and high supervisory costs. In the presence of such frictions, the separability of production and consumption decisions in agricultural households may no longer hold (Muller, 2014; Udry, 1996). Such frictions can increase the shadow price of children's time due to the relaxation of the liquidity constraint, which, in turn, can increase child labor. This is particularly true for households for which microcredit barely brings the entrepreneurial activity within reach (Augsburg et al., 2012).

Thus, the effect of microcredit on child labor is not clear ex-ante. While microcredit has the potential to lower child labor through income effect, it can increase the use of child labor through the substitution effect. The end result depends on the relative strength of these two effects. The empirical evidence, mostly observational, on the impact of microcredit on child labor and schooling has been mixed. Some studies in the developing country context report that access to credit can reduce child labor (e.g., Dehejia and Gatti, 2005; Jacoby, 1994;). Other studies find that microcredit increases child labor (e.g., Wydick, 1999; Hazarika and Sarangi, 2008; Augsburg et al., 2015). A number of other studies find no impact of microcredit on child labor and schooling (e.g., Tarozzi et al., 2015; Angelucci et al., 2015).

In this paper, I provide new experimental evidence in the context of an agricultural credit expansion program in Bangladesh. This microcredit program, known as Borgachashi Unnayan Prakalpa (BCUP), was initiated by the Bangladesh Rural Advancement Committee (BRAC), with financial support from the central bank of Bangladesh, to increase credit access to small farmers who were otherwise bypassed by formal credit institutions.[18, 19] The program provided loans ranging from USD 60 to USD 1500, for a variety of purposes (e.g.,

[18] In English, BCUP stands for Development Program for the Tenant Farmers.

[19] BRAC is the largest non-governmental development organization in the world.

crop production; machinery purchase; leasing lands; and rearing livestock), at an interest rate lower than that charged by other microfinance institutions.[20]

I use data from a randomized field experiment at the sub-district level to estimate the impact of agricultural microcredit on child labor.[21] The experiment was conducted in 40 sub-districts of Bangladesh, where the BCUP program had not yet been initiated as of 2012. Of these, 20 sub-districts were randomly selected to initiate the BCUP program (treatment) in 2012, while the remaining sub-districts served as controls. In my study, I use a baseline survey conducted in 2012, and a follow-up survey of the same set of eligible households conducted in 2014.[22] I rely on intent-to-treat (ITT) estimates to study the impact of the BCUP program on different outcomes including child labor. This study defines labor supplied by children below 14 years of age as child labor.[23]

I find that households in treatment areas are 20.1 percentage points more likely to take credit from the program, against a mean of zero for the control households. The average amount borrowed from the BCUP program is USD 77.8 higher for the treatment households. The local average treatment effects (LATE) estimates are much higher, USD 388. There is no effect, however, on borrowing from other sources, such as banks or co-operatives, NGOs and informal sources. Thus, I find no evidence of substitution or complementarity effects of the BCUP program on borrowing from other credit sources.

---

[20] The effective rate of interest charged by BCUP was approximately 19 percent on declining balance, compared to an average rate of 27 percent charged by other providers within Bangladesh. (Malek et al., 2015).

[21] A companion paper (Hossain et al., 2018) looked at the effect of agricultural credit on land productivity, the adoption of modern varieties, income, and household expenditure as part of overall impact assessment. The paper found treated households were more likely to adopt modern varieties of crops, and their land productivity was also higher. Further, the paper found that self-employment income increased due to the intervention, while wage income fell, and total income was unchanged. However, the paper did not document any effect on child labor.

[22] Eligibility is based on satisfying certain criteria. These are listed in Section 2.

[23]The ILO Minimum Age Convention, 1973 (No. 138) sets the general minimum age at 14 (12 for light work) where the economy and educational facilities are insufficiently developed. Since Bangladesh is a developing country, this study defines labor supplied by children below 14-year ages as child labor.

I then show that increased credit access encouraged the households in the treated subdistricts to engage more in self-employment activities, both farm, and non-farm. Households in the treated area borrowed significantly more land under fixed-rental contract. The total amount of land cultivated was approximately 8 percent higher for the treated households. Further, the number of non-farm self-employment activities operated by a household was 39 percent higher in the treated sub-districts, and the point estimate is statistically significant at the five percent level. Finally, the current market price of the business assets was also significantly higher for households in the treated sub-districts.

Increased self-employment activities necessitate the use of more labor. Since hiring external labor can be costly, households can opt to use child labor. My results suggest that microcredit indeed leads to an increase in child labor; Households in the treated sub-districts were 7.15 percentage points more likely to involve their children in household enterprises, compared to the control mean of 6.1 percent.[9] Given a 20 percent participation rate of the treated households in the BCUP program, and a 7 percent increase in child participation in household based economic activity, this implies participating households are 35 percent more likely to use child labor. Similarly, the probability that households used child labor in the endline, but not in the baseline, was nearly six percentage points higher for the treatment sub-districts.

I then explore the impact of increased access of microcredit on hours of child labor used. I find that children from households in the treated areas, on average, worked 0.34 more hours per week (43 percent higher) in household enterprises, compared to the control mean of 0.78 hours. This effect size is comparable to that estimated by Augsburg et al. (2015) – 0.53 more hours per week for teens aged between 16 and 19, against the control mean of 0.18 hours. Given only 20 percent of the treated households borrowed from the BCUP program, an ITT estimate of 0.34 hours implies that participating households were using around 1.70 hours of child labor more per week. The ITT and LATE estimates of microcredit, conditional upon ever using child labor, is much higher – approximately 0.88 hours and 4.40 hours respectively. Thus, the magnitude of the treatment effect on hours of child labor used seem economically meaningful.

I also observe heterogeneity in the treatment effect across households. I find that children from households with more adult members supply less labor: each additional member of working age (15-64 years) is associated with an approximate six percent reduction in child labor hours for the households in the treatment areas, compared to that of the control areas. I also find that households with more working members spend significantly more money on the education of their children. Each additional working member is associated with 10 percent more spending on education per year for the treated sub-districts, compared to that of the control ones. My results also indicate heterogeneity in the treatment effect by the gender of household head. Children from female-headed households in the treated sub-districts supply less labor. This is in line with development literature that suggests that increased women's bargaining power within the households leads to better outcomes for the children (Doss, 2013). Both of these heterogeneity results challenge the unitary model of household.

While I do not find any impact on school enrollment, the increase in child labor may result in lower school attendance (Psacharopoulos, 1997; Heady, 2003) or poorer test outcomes (Rosati and Rossi, 2003; Gunnarsson et al., 2006). While the data does not allow me to test these hypotheses, analysis using time budget survey reveals that children from the treated households spend around 3.54 hours less in studying per week, compared to the endline control mean of 20.69 hours. Overall, these findings raise concerns about the unintended inter-generational consequences of increasing credit access for the creation or expansion of self-employment activities.

One point to note that child labor outcomes were not the part of the original pre-analysis plan of the randomized experiment.[24] Thus, one potential concern is "data mining", e.g., reporting only the significant results, even though the experiment, in fact, has no effect. However, while the pre-analysis plan has some merits, imposing standards such as the pre-specification of analysis plan comes at a cost (See Olken (2015) for a detailed discussion). For instance, researchers often form new hypothesis by observing the realization of data, they did not expect to be previously. Thus, focusing too much on the pre-analysis plan might miss out on some important hypotheses. To show that data mining should not be a concern in this setting, I use a different nationally representative household survey to show that the results on child labor outcomes hold in other data too. It is very unlikely to get similar results with two different data sets by sheer chance. I also perform several sensitivity checks and show that results relating to child labor outcomes are robust to other specifications. Further, to control for the False Discovery Rate (FDR), I report the FDR- q values of the "treatment" indicator based on Benjamini and Hochberg's procedure (Anderson, 2008), and find that the results remain significant.[25]

This paper contributes to several strands of literature. First, this study present causal estimates of the effect of agricultural microcredit on child labor. Previous research has mostly relied on observational data and is not able to establish causality as clearly. To the best of our knowledge, only Augsburg et al. (2012) use experimental variation to see the effect of microfinance on teenage labor supply (16-19). In this paper, my focus is on the labor supply of children below 14 years, whom ILO defines as child labor. My study is also different from other studies because the credit programs in the other studies target either women or micro-enterprises, or are unconditional. These microcredit programs differ from conditional

[24] Detail about the experiment can be found in Malek et al. (2015). The primary analysis of the main outcomes of this experiment can be found in Hossain et al. (2018).

[25] False Discovery Rate is the expected proportion of rejections that are type I errors.

agricultural microcredit like ours, and may have different implications for child labor. Therefore, given the importance of agriculture in a developing country like Bangladesh, and given the labor-intensive nature of agricultural production in the developing countries, the impact of agricultural credit on child labor is worth examining.

Second, my paper contributes to intra-household bargaining literature. Earlier works suggest that women have different preferences compared to men, and the bargaining process often leads to better outcomes for the children, such as education, and health, when interventions are targeted towards women (Duflo, 2003; Hoddinott and Haddad, 1995). I show that this finding also holds for outcomes like child labor. I find that female-headed treated households use less child labor compared to male-headed households. My results also suggest that households do not function as a unitary body, and the effectiveness of microcredit programs can vary by gender.

Finally, my paper complements and extends the literature exploring the effect of different income shocks on human capital investment. For instance, Jacoby and Skoufias (1997) study how child labor, and thereby, child school attendance responds to seasonal variation in the income of the farm households in rural India and show that seasonal fluctuations in attendance act as a form of self-insurance. Edmonds and Theoharides (2019) document that a one-time large asset transfer program aimed at households with child labor can result in increased use of child labor, while several other papers show that employment-based safety-net programs can result in a decrease in human capital investment (Shah and Steinberg, 2015; Li et al., 2013). My paper differs from these studies on asset transfer programs in that, unlike asset transfer programs, the take-up of credit is lower (Banerjee et al., 2015b), and the amount of credit received under a microcredit program must be returned with interest. This higher interest rate could make the households myopic, which may induce parents to heavily discount the future return on human capital investment on their children (Islam and

Choe, 2013). Similarly, a microcredit intervention is also different from the employment-based safety-net interventions in their implications for the income and substitution effects.

The rest of the paper is organized as follows: Section 2 describes the randomized experiment and its design. Section 3 outlines the empirical methodology to estimate the impact of microcredit on child labor and to explore heterogeneity in the treatment effects. Section 4 discusses the results and describes the potential mechanism. I discuss various robustness and sensitivity tests in Section 5. Section 6 concludes the paper.

2 Program Description and Experimental Design

2.1 Background and Description of the BCUP Program

Credit constraints are perceived by the research community as well as policymakers to be a major barrier to economic growth in developing countries (Ayyagari et al., 2008; Beck and Demirguc-Kunt, 2006). These constraints impede investments that would propel growth. In particular, credit constraints affect marginal and small farmers (Reyes and Lensink, 2011), since they have limited assets to offer as collateral and the transaction costs of screening these applicants are higher (Barry and Robison, 2001). This lack of access to credit often leads to low agricultural productivity and low household income for the smallholders (Kumar et al., 2013). In Bangladesh, too, this lack of access to credit has posed a persistent difficulty in an effort to raise profitability for small farmers, particularly for tenant farmers (Faruqee, 2010; Hossain and Bayes, 2009), who comprised approximately 43.4 percent of the rural households in 2008 (Malek et al., 2015).

To mitigate this problem, the central bank of Bangladesh, known as Bangladesh Bank, granted BRAC a loan of USD 65 million at a five percent rate of interest only. The purpose of this loan was to allow BRAC to give loans to small farm households at an interest rate lower than other existing microcredit programs run by BRAC, or at the interest rate charged by other lending institutions. BRAC, therefore, initiated the Development Project for the Tenant

Farmers (BCUP) in October 2009, to provide credit support to small tenant farmers. By November 2014, BCUP had provided credit to over 274,000 borrowers.

The BCUP program provides loans for a variety of purposes, including producing crops, purchasing machinery, leasing land, and rearing livestock. Eligibility is based on satisfying the following criteria: farmers (a) have a National Identification Card; (b) are between 18 and 60 years of age; (c) have at most ten years of schooling; (d) have resided at the current place of residence for at least three years; (e) have landholdings under two acres (200 decimals); (f) are not members of other microfinance institutions; and (g) have expressed willingness to borrow from the BCUP program. BCUP typically offers loans for crop production, which range from USD 63 to USD 625. The program also offers loans for leasing land, with a maximum loan size of USD 750, and for purchasing machinery, with a maximum loan size of USD 1500. These loans must be paid off in equal monthly installments over one year. The effective rate of interest is 19 percent. This rate is considerably lower than the interest charged by other microcredit institutions in Bangladesh (those are on average, 27 percent (Malek et al., 2015)). In the case of failure to repay installments in due time, the borrowers must pay additional interest to the remaining installments.

2.2 Experimental Design

The experiment was conducted in 40 sub-districts (upzilas) of rural Bangladesh, where the BCUP program had not yet been initiated in 2012.[26] BRAC used a randomized experimental design to introduce the program. The unit of randomization was the BRAC branches (each branch corresponds to a unique sub-district). The randomized program was carried out at the sub-district level, with 20 sub-districts (out of 40) chosen randomly to implement the credit program under BCUP. The remaining 20 sub-districts served as the control

[26] These 40 sub-districts were identified by the BCUP program personnel based on agro-ecological conditions and level of infrastructural development.

group.[27] The 40 branches were pre-selected by the BRAC's credit program, and these branches are located in different parts of Bangladesh (Figure 1).

Once the randomization of the sub-districts was accomplished, six villages from each of the sub-districts within a radius of eight kilometers of the respective sub-district office were selected randomly. One potential worry is that there could be spillover effects or contamination from the intervention, which may bias the results. Therefore, geographic information of the sub-district offices was collected from the control and treatment subdistricts. As seen in Figure 1, most of the control sub-districts were quite far from the treatment sub-districts, with the notable exception of the sub-districts in the southern region. However, since the sample villages were within an eight-kilometer radius of the BRAC office, and the area of a sub-district is much higher, spillover are less likely to be a concern to this setting. In fact, when eight-kilometer buffer zones were constructed around the BRAC subdistrict offices in the southern region, using Geographic information system (GIS) mapping, none of them overlapped with one another (See Malek et al. (2015)). As a further precaution against the risk of spillover, for the sub-districts located in the southern part of the country, maps were circulated to the program officials, to make sure they operated in the treatment areas without accidentally overstepping into control areas.

2.3 Sample Selection

A census was carried out in the 240 villages in April 2012, and 31,322 households were surveyed, out of which, 7,563 households (nearly 24 percent of total) met the eligibility criteria for participating in the program. Out of all these eligible households, a total of 4,331 households were randomly selected for the household survey. This sample size was chosen to ensure sufficient power to detect changes in farm productivity, income, and food

[27] The only intervention in this experiment was BCUP credit. Initially, it was thought that the credit program would be bundled with agricultural extension service, which could not be materialized due to administrative issues. Thus, there was no other assistance that was exclusive to the treatment groups.

expenditure of the households, which were the primary outcomes of interest of BCUP program.[28] One implication of this is that the statistical power would be ex-ante sufficient for the incidence of child labor, the main outcome variable in this study, since the standard error of this variable is much lower to those of variables such as productivity and income.

A population-proportioned random sampling was followed to get the final number of sample households from the list of the eligible households, based on the concentration of the eligible households – areas where the concentration of eligible households was higher, were given more weight. This was done in two steps: In the first stage, the total sample needed at the sub-district level was chosen, using that sub-district's share out of all the eligible households as weight. In the second stage, the total sample needed at the village level was calculated, using that village's share out of all eligible households in the sub-district to which that village belongs. Eventually, a total of 2,164 households were selected for treatment, and the remaining 2,167 households served as control.

The baseline survey was conducted between June and August of 2012 (See Figure 2). The survey collected a wide range of information on households. These include information on demographics for all household members, household income from different farm and non-farm activities, land use and productivity, labor supplied in household enterprises by children aged between 5 and 14 years, household expenditure on different categories, and food security. For a sub-sample of these households (n = 1,248), a time budget survey was also conducted to get detailed information on the time allocation of the household members.

2.4 Baseline Sample Description and Balance

Nearly 99 percent of 4,331 households were surveyed in the baseline. The baseline allows testing the validity of the random assignment by comparing the household characteristics at

[28] See Malek et al. (2015) for details.

the baseline. To this end, I run regressions of the baseline outcome variables on the treatment dummy. Results are reported in Table 1, which suggest that the assignment of households between the treatment and control areas are random. The household characteristics are balanced across the two assignment groups. Of the 33 variables compared, only two variables (access to loan from NGOs, and access to any type of credit) is significant at the 5 percent level of significance, and one variable (sex of household head) is marginally significant at 10 percent. Further, these differences are not very large, and there is no systematic pattern in these differences between the control and treatment variables. All these suggest that selection bias is unlikely to pose a threat to this study.

The summary statistics depict a poorer overall socioeconomic status for the sampled households, compared to the national average. This is expected, however, because the BCUP eligibility criteria dictated that the household head could not be highly educated or household could not own more than two acres (200 decimals) of land. The sample households are larger, with 4.86 members on average, while the national average is 4.50 (BBS, 2011). A significant share (44 percent) of the household heads has no education. Most of the households are headed by the male; Female heads comprise only five percent of the total (Table 1).

Households from the treatment areas, on average, own 37.45 decimals (0.37 acre) of farming land; households from the control areas own an average of 38.71 decimals (0.38 acre). This is much less than the national average, which is 65.25 decimals (0.65 acre).[29] A significant share of the households leased land through the tenancy market. However, the prevalence of renting land to others was low. Both the treatment and control groups had very limited access to credit. Only 14 percent of the sample households had access to credit of any sort. The corresponding national average was 32.03 percent (BBS, 2011).

[29] Author's calculation from the 2010 Household Income and Expenditure Survey.

Household annual consumption expenditure per capita was USD 265 for the control households, whereas the corresponding 2010 national average was USD 366 (BBS, 2011). Per capita calorie intake and protein intake was about 2240 and 57.80 grams, respectively, whereas the corresponding national averages for the year of 2010 was 2318 and 62.66 grams, respectively (BBS, 2011).

2.5 End-line Survey and Attrition

The end-line survey was carried out in 2014, two years after the baseline survey. It was administered to the same households, to construct a two-round panel dataset. For all the outcome variables of interest, the same questions from the baseline survey were asked in the end-line survey. The survey was administered to 96 percent (4,141 households) of the original sample.

One potential worry is the differential attrition between the two groups of households, which may result if people choose to move into or out of the area because of the intervention. If attrition is different between the treatment and control groups, we cannot strictly compare these two groups. Therefore, I tested if there was differential attrition between the two groups between the pre-intervention and post-intervention surveys. I used the panel sample to study attrition. The results are reported in Table 2. As can be seen from Panel A of Table 2, the attrition rate was similar across the groups, and a p-value of the difference of 0.65 suggests that the difference is not statistically significant. I also check whether attrition changed the composition of the households in endline differently for the treatment and control households. To this end, I regress different household characteristics on the likelihood of attrition and the interaction of the likelihood of attrition and treatment. Panel B shows that the interaction of the probability of attrition and treatment is not significant. This suggests that the composition of the household remained the same, even after attrition. I also check whether Household characteristics can predict attrition. Appendix Table A1 shows that HH characteristics cannot predict attrition. Overall, the results provide suggestive evidence that differential attrition should not pose a threat to the interpretation of my results.

3 Empirical Specifications

To estimate the effect of BCUP on different outcomes, I focus on the intent-to-treat (ITT) estimates. To this end, I run ordinary least squares regressions of treatment assignment, and estimate the difference in the average outcome of the treatment and control households. I also estimate the Local Average Treatment Effects (LATE) of BCUP credit take-up by running a two-stage least squares regressions (with treatment assignment as an instrument for BCUP credit uptake). For both ITT and LATE estimates, I rely on the randomization of the treatment assignment to obtain causal interpretation.

To estimate the effect of treatment assignment, I run an OLS regression of the following form:

$$Y_{is,2014} = \beta_0 + \beta_1 Treatment_s + \beta_2 Y_{is,2012} + \epsilon_{is} \qquad (1)$$

Where *is* denotes household *i* from sub-district *s*, 2014 is the endline year, and 2012 is the baseline year. Y_i is the outcome of interests, for example, amount of expenditure on education by the households, or hours of labor supplied by children (5 years to 14 years). Treatment is a dummy variable taking a value of 1 if the sub-district *s* was selected for treatment. My main coefficient of interest is β_1, which measures the average difference in mean outcomes between the households of treatment and control groups. I include the baseline version of the outcome as an additional covariate in the regression. I do not need this covariate for a causal interpretation of β_1, since this covariate is unrelated to treatment assignment. However, these covariates have the potential to improve the precision of the estimates by controlling for chance differences between the two groups of households, in variables that may influence the outcomes (Taubman et al., 2014). In the empirical analysis, the standard errors are clustered at the branch (sub-district) level, since the randomization is at the branch level.

The ITT estimates measure the net effect of increased credit access. However, exploring the effect of credit itself on different outcomes can be of significant policy interest. Hence, I report the Local Average Treatment Effect (LATE) estimates in the Appendix Tables B1 to B4.[30]

I estimate heterogeneous treatment effects using the following model,

$$\begin{aligned} Y_{is,2014} = \beta_0 + \beta_1 Treatment_s + \beta_2 Y_{is,2012} + \beta_3 (X_{is,2012} \times Treatment_s) \\ + \beta_4 X_{is,2012} + \epsilon_{is} \end{aligned} \qquad (2)$$

where X_{is} is the heterogeneity variable e.g., number of members of working age or the sex of household head in household *i* from sub-district *s*. It is to note that, I use the baseline values for X_{is}, since interacting the treatment variable with X_{is} measured at endline would be endogenous. The parameter β_3 shows how the main effect varies with the number of the family's working members or gender of the household head.

4 Results and Discussions

4.1 Effect on Borrowing Behavior

I first document the impact of BCUP on borrowing behavior.[31] Panel A of Table 3 documents the intent-to-treat estimates of BCUP on the borrowing behaviors of the households. Column 1 shows that households in the treated sub-districts are 20.1 percentage points more likely to receive BCUP credit, compared to a mean of zero percentage points for the control sub-

[30] The empirical specification of LATE is discussed in Appendix Section B.1

[31] It is to note that Table 3 (credit access) and 4 (land use) were also in the companion paper (Hossain et al., 2018), albeit in slightly different forms: One, I used more variables in each of the tables to elicit the borrowing and land use behavior of the households. Two, I used USD as a unit of measurement in Table 3 of this paper instead of BDT as was in the case of the companion paper.

districts.[32] This number is comparable to those found in other impact evaluations of microcredit, despite the difference in contexts. For instance, the probability of borrowing is 9-12 percentage points higher for the treatment households in Morocco (Cr´epon et al., 2015), Mexico (Angelucci et al., 2015), and India (Banerjee et al., 2015a), due to microcredit introduction (India) or expansion (Morocco and Mexico) in the respective countries.

While a take-up rate of 20 percent is comparable to the other studies of microfinance, it is quite low considering the fact the all the households in the sample expressed interest to take the credit if offered. A number of demand and supply-side factors contributed to the low take-up.[33] One demand-side factor was the higher interest rate. In the census questionnaire, the households were asked whether they will take the credit if BRAC disburses credit with *easy terms*. Since the program could not fix the interest rate that would be charged by that time, the study team could not inform the households about the exact interest rate. While the 19 percent interest rate charged by BCUP was considerably lower than the market interest rate, it was still deemed higher to some of the households. A second demand-side factor is the occurrence of several natural shocks and price shocks for agricultural products during the time interval between the census and the implementation of the program. A qualitative survey conducted by BRAC suggests that such price and natural shocks made many previously identified eligible, willing farmers reluctant to take credit due to the restriction on BCUP credit to be used for agricultural purposes. Among the supply-side factors, the most important was the higher scrutiny in verifying eligibility by the program implementation committee than that by the study team. The study team used a simple questionnaire to find out the eligible households during the census to draw the sample for the study. However, the

[32] A total of 592 loans were taken from the BCUP program. Out of these 592 loans, 74 loans were taken during July-December 2012, 325 loans (54.9%) were taken in 2013, and 193 loans (32.6%) were taken in 2014. Households were allowed to take multiple loans. 233 households took one loan each, 174 households took two loans each, and 4 households took three loans each.

[33] A more detailed discussion can be found in Malek et al., 2015.

program staff collected more detailed information and validated those by cross-checking (e.g., verification of National Identity Card). Therefore, some households in our sample were not eligible to take loans. Unfortunately, the survey questionnaire did not collect information on why those households did not take BCUP credit or whether they were denied credit by BCUP. In Appendix Table A2, I show the determinants of BCUP credit take-up. It turns out that take-up is negatively correlated with HH head's age and female headship, number of working-age members in the HH and HH land ownership, but positively correlated with HH consumption expenditure.

The impact of BCUP on borrowing from other sources is reported in Columns 2-4 of Table 3. It is evident that there is no significant difference between treatment and control households in borrowing from banks or co-operatives (Column 2), or from NGOs (Column 3), or from informal sources (Column 4).[34] Overall there is no significant impact of the intervention on borrowing from any sources other than the BCUP program (Column 5). These results suggest that the intervention did not crowd-in or crowd-out borrowing from other sources.

With respect to the amount of borrowing, a similar overall picture is observed. Column 1 of Panel B of Table 3 shows a significant increase in the amount of credit taken from BCUP. The amount borrowed from BCUP is USD 77.80 more for the households in the treated areas. In Appendix Table B1, I document the corresponding estimates when I use the treatment assignment dummy as instruments for BCUP credit take-up. The resulting 2SLS estimates are much larger, indicating an increase of USD 388 in borrowing from BCUP for households that actually participated in the BCUP program. Again, as was the case with the probability of borrowing from other sources, there is no statistically significant difference in the amount borrowed from different other sources (Column 2-4 of Panel B of Table 3).

[34] Informal lender includes moneylenders, loans from friends/family, and buying goods/services on credit from seller.

4.2 Effect on Self-Employment Activities

One of the main aims of the BCUP program was to encourage credit-constrained farm households to engage in more self-employment activities. To the extent that the BCUP program eases credit needs, one would expect that credit-constrained farmers would engage more in self-employment activities. To see whether this is indeed the case, I explore the effect of credit on different proxies of self-employment activities. The impact of credit on the size of cultivated land is reported in Table 4. I find that households in the treated sub-districts cultivated approximately six decimals more land (around 7.6 percent more) compared to those of control sub-districts (Column 5), although this difference is not significant (p-value=0.24). The composition of total land rented under different tenancy arrangements is more telling. I find no significant difference in the amount of land rented under share-cropping arrangements (Column 1). However, Column 2 reveals a significant increase (6.5 decimals more land) in the amount of land rented in fixed-rental arrangements. This has important implications for labor demand for land cultivation. The Marshallian theory suggests that sharecropping leads to Pareto inefficient allocation of labor, because sharecroppers receive only a percentage of their marginal product of labor. To the extent that share-cropping uses sub-optimal labor, a shift from a share-cropping contracts to a fixed rental contracts implies an increase in the demand for labor.

Next, I show the impact of BCUP on non-farm self-employment activities. The results are reported in Table 5. Column 1 shows that the households from the treated areas are 6.34 percentage points more likely to participate in non-farm self-employment activities. The number of non-farm self-employment activities is also significantly higher in the treated areas, nearly 39 percent higher compared to the control areas (Column 2). I also find that the number of family members used for non-farm self-employment activities is 49.8 percent higher for the treatment sub-districts, and the point estimate is statistically significant (Column 3). On the contrary, there is no significant impact on the number of hired laborers for nonfarm self-employment activities (Column 4). Finally, the current market price of all assets engaged in non-farm self-employment activities is significantly higher for the households in the treated sub-districts. These results indicate that microcredit encouraged the households in the treated areas to participate more in non-farm self-employment activities.

4.3 Effect on Child Labor

From Tables 4 and 5, it is evident that the credit program increased both farm and non-farm self-employment activities. This necessitates an increase in the demand for labor. However, since hiring labor is costly, households may prefer internal labor to hired labor. This is not a problem, if the household has surplus adult labor, who are otherwise un(der)employed. However, in the absence of such surplus adult labor, an increase in the self-employment activities may lead to an increase in the use of child labor. Therefore, I inspect the impact of credit on the probability that a household uses child labor. The results are reported in Table 6. It is observed that households in the treatment area are 7.15 percentage points more likely to involve their children in self-employment activities, compared to the control mean of 6.1 percent (Column 1). Given a 20 percent participation rate of the treated households in the BCUP program, and a 7 percent increase in child participation in household based economic activity, this implies participating households are 35 percent more likely to use child labor. Further, the probability that households used child labor in endline, but not in the baseline, is 5.8 percentage points higher for the treated sub-districts.[35] This effect size is comparable to that of Edmonds and Theoharides (2019), who find that children not in child labor at baseline are nearly eight percentage points more likely to be economically active, due to a one-time asset transfer program in the Philippines.

[35] In Appendix Table A3, I use a Multinomial Logit specification, and find similar results.

In Column 3 of Table 6, I show the effect of credit on hours worked in household enterprises by children aged between five to 14 years.[36] I find that hours worked by children in household enterprises are significantly higher for the treatment group. Children in households from the treated sub-districts, on average, worked approximately 0.34 hours more each week (43 percent more), compared to the households from the control areas (Column 1). To put this into perspective, Augsburg et al. (2015) records an increase in labor supply of 0.53 more hours each week by teens aged between 16 and 19, against a control mean of 0.18 hours. Given only 20 percent of the treated households borrowed from the BCUP program, an ITT estimate of 0.34 hours implies that participating households were using around 1.70 hours of child labor more per week. The ITT estimate of microcredit, conditional upon ever using child labor, is much higher – approximately 0.88 hours. Given 20 percent actual participation in the BCUP program, this translates to a LATE estimate of 4.40 hours per week. Thus, the magnitude of the treatment effect on hours of child labor used seem economically meaningful.

4.4 Evidence from Time Use Survey

I cross-validate my findings on child labor using a time-budget survey. The time-budget survey was done for a subset of the original sample. I estimate the impact of BCUP on hours worked on different activities by children aged between five to 14 years.[37] These results are reported in Column 1 of Table 7. I observe that Children from the treated sub-districts, on

[36] Household enterprises include both self-employment farm activities such as poultry raising, livestock rearing, and fishery, and non-farm activities like agro-processing industries, wholesale and retail trading, storage and communication, transport and education, health industries and other service-related activities.

[37] Unlike the household data on time spent by children on self-employment activities, the time budget survey is at the individual level. I take the advantage of this by estimating the impact of credit on time spent in different activities at the individual level since it provides more information about how many household members work how much and what the potential margins are to shift and by how much. Since the individual level data is not a balanced panel, and as such does not allow me to take the baseline version of the outcome variable as a regressor, I run regressions of weekly hours supplied on treatment dummy, Post dummy, and the interaction of treatment dummy with Post dummy instead of estimating equation 1.

average, worked approximately 0.72 hours more each week in self-employment activities, compared to a control mean of 0.44 hours (Column 1). This effect is statistically significant at the 5 percent significance level. However, I do not observe any significant difference between children from treated and control households regarding hours spent on wage employment and household chores. I also explore the effect of the microcredit on different non-economic activities such as leisure, and study. I find that children from the treated households spend around 3.54 hours less in studying each week, compared to a mean of 20.69 hours per week for the control households. The effect is significant at 5 percent level of significance. I also conduct these analyses separately for male children (Panel B) and female children (Panel C). It is evident that the effects on time spent in self-employment and study were mainly driven by male children.

I then look at how the BCUP credit affected the labor supply of working-age population (15-64 years) of the households. Working-age males from the treated sub-districts worked significantly more hours on self-employment activities and significantly fewer hours on HH chores, while there was no significant difference between the treated and control areas in the hours spent in wage-earning activities (Appendix Table A4). The number of hours spent on self-employment activities was higher, albeit insignificantly, for the working-age females from the treated areas.

Overall, these findings suggest that the demand for labor in self-employment activities increased in the treated areas due to the intervention. The adult males contributed most to this increased demand (5.2 hours more per week). This increase in hours spent in self-employment activities came at the expense of hours spent on HH chores by adult males. Adult females also chipped in to contribute to the increased labor demand in self-employment activities, albeit insignificantly (0.97 hours more per week, SE (1.12)). Finally, children from the treated households spent more hours on self-employment activities. These results, coupled with the observation that households with fewer working members supply more child labor in self-employment activities, suggest that time is a binding constraint for the adult members of the households.

4.5 Heterogeneity Analysis

I found that microcredit led to an increase in self-employment activities, both farm, and nonfarm, for the treated households. Such an increase in self-employment activities necessitates more labor. This should not be a problem if the households have surplus working adults. But in the absence of such working-age members, families might opt to use child labor in the new self-employment activities created by the microcredit intervention.

I, therefore, test for heterogeneity in the treatment effect by the number of members of working-age (15-64 years) in the family. I find that the interaction of treatment assignment and the number of working-age members is both negative and significant for both of the child labor outcome variables – the likelihood that the household uses child labor and weekly hours supplied by the children. This implies that the treatment effect is much lower for households with more working members (Table 8). This is expected, since the households may not need to involve their children in the self-employment economic activity if there are more adult working members in the family. My findings thus suggest that household composition is an important determinant of household labor supply, implying non-separability in the consumption and production decisions of farm households. This is different from the findings of Benjamin (1992), who found evidence in favor of separability, i.e., labor allocation decisions of farm households do not depend on the structure of the household.

Earlier works suggest that women have different preferences compared to men, and the bargaining process often leads to better outcomes for the children, such as education, and health, when interventions are targeted towards women (Duflo, 2003; Hoddinott and Haddad, 1995). To this end, I explore the heterogeneity in the treatment effect by the gender of the household heads. Table 8 indicates that the treatment effect of credit on the likelihood that a household uses child labor is smaller for female-headed families. More precisely, when offered loan, the likelihood of using child labor is 12 percentage points lower for the female-headed households than the male headed households (Column 2). Similarly, labor supplied by children between 5 to 14 years is significantly lower for female-headed families than that of male-headed families in the treated sub-districts (Column 4). These results suggest that the effect of expanding credit access can vary by the gender of the household heads. However, one should interpret the results with a few caveats in mind. First, only 5 percent of the households are headed by females. Second, one could argue that the finding of children in female-headed households not increasing their labor supply is not necessarily driven by female preferences for education. Rather, it might well be a result of types of investments made by these households. While I cannot rule out this explanation, my findings that engagement in self-employment proxies are not significantly different for the male- and female-headed households (Appendix Table A5) does not support it either.

I also explore the heterogeneous impact of credit on child labor by baseline child labor use. This analysis would provide further information on whether the treatment effect are driven by households already using child labor or households that started using child labor after the intervention. Results are reported in Appendix Table A6. I find no evidence of heterogeneity in the treatment effect by baseline child labor use. This implies that the increase in child labor is not driven by the households that used child labor in the baseline. Rather the effect on child labor are driven by households that did not use child labor in the baseline but started using child labor following the roll-out of credit.

4.6 Education Expenditure and Schooling Outcome

I also try to explore how credit affected the education expenditure of households. Education expenditure is the summation of expenditure on the following items: (a) institutional (e.g., school) fees; (b) Books, exercise books, pen and pencils; (c) salary of private tutor; (d) school uniform; (e) other educational expenses. Table 9 shows the intent-to-treat estimate of microcredit on household expenditure on education. Column 1 shows that expanded access to microcredit is associated with lower spending on education, although the effect is not statistically significant. Column 2 reveals the heterogeneity in treatment effect. The interaction of treatment dummy and the number of working members is positive and significant. This means that the impact of expanded microcredit access on education expenditure is higher for households with more working members.

Finally, I see the impact of credit on school enrollment of children. More specifically, I explore the effect of credit on the likelihood that a child has stopped attending the school, or has never attended school. The results are reported in Table 10. I find no impact of credit on these indicators. These results should come as no surprise, because primary schooling is free and compulsory in Bangladesh. Therefore, credit constraints should not act as a barrier to school enrollment. However, while children are not being taken out of school to engage in household enterprises, they are trading-off time they could use for education with supplying labor for household enterprises. As I showed earlier, children from the treated subdistricts spend significantly less time in studying. This should negatively affect the quality of education, in terms of lower grades (Rosati and Rossi, 2003; Gunnarsson et al., 2006) or lower school attendance (Psacharopoulos, 1997; Heady, 2003). Unfortunately, the data does not include information on children's school results or attendance to verify these hypotheses.

5 Robustness

5.1 Multiple Hypothesis Testing and External Validity

Multiple hypothesis testing is another potential concern in my setting, especially because child labor was not the main outcome of the field experiment, and hence, was not part of the pre-analysis plan. To reduce the concern about the False Discovery Rate, I perform a number of tasks. First, in Appendix Table A7, along with different child labor outcomes, I report the primary outcomes of the experiment – productivity, adoption of modern varieties, income and expenditure, and following Anderson (2008), report the FDR- q values of the "treatment" indicator based on Benjamini and Hochberg procedure. The FDR-q values control for the *False Discovery Rate* (FDR), or the expected proportion of rejections that are type I errors. It is evident that child labor outcomes remain significant even after controlling for multiple hypothesis testing.

Second, I used a different household survey data-set (representative at the sub-district level, the unit of treatment in our experiment) to show that the findings on child labor hold in other data too. To this end, I use two rounds of the Bangladesh Sample Vital Registration Survey – 2012 and 2014. SVRS is the only national household survey in Bangladesh, which is representative at the sub-district level, which is the unit of treatment in our experiment. I confined my analysis to the 40 sub-districts which were part of the original randomized experiment. Employing a difference-in-difference estimation strategy using those two rounds of the pooled cross-section data, I find that treated households are 4.05 percentage points more likely to use child labor in self-employment activities (Appendix Table A8). The results are robust to the inclusion of different control variables. This result provides the strongest suggestive evidence that False discovery Rate is not a key driver of my main results, since it is highly unlikely that the two different data sets would produce the same findings purely by chance.

Third, I looked at other micro-finance studies to see whether those studies can back up my results. Of the very few RCTs that assessed the impact of microcredit on child labor as part of the overall impact assessment, the experiment in Ausberg et al. (2016) was the closest to ours regarding the type of credit offered. For instance, in both experiments, the borrowers were both male and female, loan size as a proportion of income was comparable, the interest rate was almost similar, and liability was individual. Ausberg et al. (2016) find that the treated households use 0.53 (SD 0.23) hours more of teenage (16 to 19 years) labor supply per week on business activities compared to the control households. For the comparable group of teens in my study, I find an increase of similar magnitude (0.49 hours of labor supply per week) for the treated households than the control ones (Appendix Table A9). The other RCTs were much different in terms of context and in the type of loans offered. For instance, the loans in Angelucci et al. (2015), and Banerjee et al. (2015a) were targeted towards females. Neither of these studies finds any significant impact of credit on child labor. This should not come as a surprise since earlier works suggest that women have different preferences compared to men, and the bargaining process often leads to better outcomes for the children, such as education, and health, when interventions are targeted towards women (Duflo, 2003; Hoddinott and Haddad, 1995). In my paper, too, I find that the treatment effect of credit on the likelihood of using child labor is lower for the female-headed households than the male-headed households. I summarize these discussions in Appendix Table A10 .

5.2 Other Robustness and Sensitivity Checks

I perform several robustness tests to allay concerns about the validity of my findings. First, as mentioned in the empirical framework, to estimate the intention-to-treat effect, I run regressions of the outcome variables on the random treatment assignment and the baseline version of the outcome variable. Such estimations can generate more power than difference-in-differences when outcomes are not strongly auto-correlated (McKenzie, 2012). If, however, the outcomes are strongly auto-correlated, say $\rho = .60$ or $\rho = .80$, this advantage tapers off. As seen in Tables 5-6, the auto-correlation between the outcome variables is very low. This suggests that the difference-in-difference (DiD) approach is not the right specification in my setting. Nevertheless, I use several other estimation strategies to show that my results are robust to other specifications too. First, in Appendix C, I report the ITT estimates using the difference-in-difference (DiD) approach, where I run regressions of the difference in outcome between the endline and the baseline on treatment assignment. It is evident that the DiD estimates are very similar to the estimates from our preferred specification, although the DiD estimates are less powered. Second, in Appendix D, I show simple regression estimates in which no control variables are included and find consistent results. Third, I control for the baseline variables using the Lasso method proposed by Belloni et al. (2014), and find consistent results (Appendix E).

Second, as seen in figure 1, some control sub-districts and treatment sub-districts are located adjacent to one another. In section 2, I argued that this should not be of much concern. However, as a robustness check, I run regressions excluding those sub-districts, which were located adjacent to one another. In results available on request, I find that the results are robust to the exclusion of these adjacent sub-districts.

Third, I cluster the stander error at the unit of randomization, which is sub-district in this study. One might be concerned that the number of clusters in this study (40: 20 treatment and 20 control sub-districts) is relatively small. However, Cameron et al. (2008) note that standard asymptotic tests can over-reject the null when there are only few (five to thirty) clusters. Since we have more than 30 clusters, this should not be a concern in this study's setting. Regardless, in Tables F1 to F4 in Appendix F, I adjust standard errors using the wild cluster bootstrap-t procedure (Cameron et al., 2008), and find similar inferences.

6 Conclusion and Policy Implications

In this paper, I investigate the impact of an agricultural credit expansion program on child labor, using a randomized field experiment conducted over 40 sub-districts in Bangladesh. My results suggest that household participation in microcredit programs increases child labor – Households that borrowed from the credit program, are 35 percent more likely to use child labor in their self-employment activities. The treatment effect of credit on child labor is more pronounced for households with fewer working adult members. The results from a time budget survey corroborate these findings. I argue that access to credit has opened new opportunities for children to work in household enterprises. I also find that, when offered loans, female-headed households use less child labor. This result aligns with the findings of development literature that women's bargaining power within households is associated with better outcomes for the children (Doss, 2013).

While I do not find any impact of credit on school enrollment, I find that children from the treated sub-districts spend significantly less time studying compared to the control ones. I also find that the education expenditure in households with a smaller number of working members is adversely affected by access to credit. These results imply that, although children are not being taken out of school to engage in household enterprises, the time they could use for their education is traded-off to supply labor for household enterprises. This should negatively affect the quality of education, in terms of lower grades or lower school attendance. While the data does not allow to verify these hypotheses, earlier literature suggests the increase in child labor may result in lower school attendance (Psacharopoulos, 1997; Heady, 2003) or poorer test outcomes (Rosati and Rossi, 2003; Gunnarsson et al., 2006). Even if most child laborers attend schools, hours worked still have significant consequences for educational attainment (Edmonds and Pavcnik, 2005). For instance, Beegle et al. (2009) finds a 35 percent decrease in schooling attainment due to a one standard deviation increase in hours worked for children attending school.

One implication of my results is that expanded access to credit can potentially influence a farm household's labor allocation decisions in a way that might negatively affect investment on human capital.[38] In the absence of any offsetting policy, these decisions can increase inequality in the long-run. Therefore, policymakers should be careful in introducing microcredit interventions to increase credit access to the credit-constrained household. A straight ban on child labor, however, may not be helpful, and may have perverse effects on child labor (Bharadwaj et al., 2013). Therefore, to implement the simultaneous goals of human capital formation and poverty reduction, microfinance interventions should be complemented by other policies (Islam and Choe, 2013). One such policy could be making access to microcredit conditional on the school requirement of the children within the household (Wydick, 1999). Reducing the interest rate or increasing the repayment period also have the potential to reduce child labor by allowing the marginal households to resort to hired labor instead of child labor (Islam and Choe, 2013).

Funding

I acknowledge the following financial support: Bangladesh Bank for sponsoring the BCUP program, the International Initiative for Impact Evaluation (3ie) for sponsoring the randomized control trial. These sponsors, however, did not play any other role aside from providing funding.

Declaration of Interest

I have no competing interest to declare.

[38] It might be possible that the experience gained from working in household enterprise may result in useful skill formation, which might increase future labor market return. However, existing evidence is mixed. For instance, Beegle et al. (2009) find that the net present discounted value of child labor is positive for higher level of discount rates (11.5 percent or higher). On the contrary, Emerson and Souza (2002) argue that entering the labor market earlier reduces the adult earnings significantly. One study goes even further claiming that "Child workers of today become the unskilled workers of tomorrow" (pp 24, Gupta (2001)).

Data Availability

This paper uses data from a randomized experiment, maintained by the principal investigator of the experiment, Mohammad Abdul Malek. Detail about the experiment can be found here. The primary analysis of the main outcomes of this experiment can be found here. This experiment was later registered into AEA RCT Registry by Palmen Nikolov (RCT ID: AEARCTR-0004460). The data can be obtained by filing a request directly to Mohammad Abdul Malek (E-mail: malekr25@gmail.com). The author of this paper is willing to assist (Md Amzad Hossain, e-mail: mh2vh@virginia.edu). I can make available my replication code and the data for replication by the journal if and when needed. I request that the data not be posted online and it only be shared with researchers for the express intent of replicating my results. To this end, I am willing to let researchers write to me asking for the replication data. I can share it with them if they sign an understanding that the data is only being used for replication purpose. However, if the journal can be the conduit of this information to the prospective replication seekers, I am happy to comply and submit my data files to any repository.

References

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association*, 103(484):1481–1495.

Angelucci, M., Karlan, D., and Zinman, J. (2015). Microcredit impacts: Evidence from a randomized microcredit program placement experiment by Compartamos Banco. *American Economic Journal: Applied Economics*, 7(1):151–82.

Augsburg, B., De Haas, R., Harmgart, H., and Meghir, C. (2015). The impacts of microcredit: Evidence from Bosnia and Herzegovina. *American Economic Journal: Applied Economics*, 7(1):183–203.

Augsburg, B., Haas, R. D., Harmgart, H., and Meghir, C. (2012). Microfinance, Poverty and Education. IFS Working Papers W12/15, Institute for Fiscal Studies.

Ayyagari, M., Demirgu¨¸c-Kunt, A., and Maksimovic, V. (2008). How important are financing constraints? The role of finance in the business environment. *The world bank economic review*, 22(3):483–516.

Banerjee, A., Duflo, E., Glennerster, R., and Kinnan, C. (2015a). The miracle of microfinance? Evidence from a randomized evaluation. *American Economic Journal: Applied Economics*, 7(1):22–53.

Banerjee, A., Karlan, D., and Zinman, J. (2015b). Six randomized evaluations of microcredit: Introduction and further steps. *American Economic Journal: Applied Economics*, 7(1):1– 21.

Barry, P. J. and Robison, L. J. (2001). Agricultural finance: Credit, credit constraints, and consequences. *Handbook of agricultural economics*, 1:513–571.

Barua, R. and Lang, K. (2009). School entry, educational attainment and quarter of birth: A cautionary tale of late. Technical report, National Bureau of Economic Research.

BBS (2011). Report of the household income and expenditure survey 2010. Technical report, Bangladesh Bureau of Statistics, Dhaka, Bangladesh.

Beck, T. and Demirguc-Kunt, A. (2006). Small and medium-size enterprises: Access to finance as a growth constraint. *Journal of Banking & finance*, 30(11):2931–2943.

Beegle, K., Dehejia, R., and Gatti, R. (2009).Why should we care about child labor? the education, labor market, and health consequences of child labor. *Journal of Human Resources*, 44(4):871–889.

Beegle, K., Dehejia, R. H., Gatti, R., and Krutikova, S. (2008). The consequences of child labor: evidence from longitudinal data in rural Tanzania. Policy Research Working Paper Series 4677, The World Bank.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, 81(2):608– 650.

Benjamin, D. (1992). Household composition, labor markets, and labor demand: Testing for separation in agricultural household models. *Econometrica*, 60(2):287–322.

Bharadwaj, P., Lakdawala, L. K., and Li, N. (2013). Perverse consequences of well intentioned regulation: Evidence from India's child labor ban. Technical report, National Bureau of Economic Research.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3):414–427.

Cr´epon, B., Devoto, F., Duflo, E., and Parient´e, W. (2015). Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco. *American Economic Journal: Applied Economics*, 7(1):123–50.

Dehejia, R. and Gatti, R. (2005). Child labor: The role of financial development and income variability across countries. *Economic Development and Cultural Change*, 53(4):913–932.

Doss, C. (2013). Intrahousehold bargaining and resource allocation in developing countries. *The World Bank Research Observer*, 28(1):52–78.

Duflo, E. (2003). Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in south africa. *The World Bank Economic Review*, 17(1):1–25.

Edmonds, E. V. (2007). Child labor. Working Paper 12926, National Bureau of Economic Research.

Edmonds, E. V. and Pavcnik, N. (2005). Child labor in the global economy. *Journal of Economic Perspectives*, 19(1):199–220.

Edmonds, E. V. and Theoharides, C. B. (2019). The short term impact of a productive asset transfer in families with child labor: Experimental evidence from the Philippines. Technical report, National Bureau of Economic Research.

Emerson, P. and Souza, A. P. (2002). From childhood to adulthood: the effect of child labor activities on adult earnings in brazil. *Vanderbilt University Department of Economics Working Paper*.

Eswaran, M. and Kotwal, A. (1986). Access to capital and agrarian production organisation. *The Economic Journal*, 96(382):482–498.

Fallon, P. and Tzannatos, Z. (1998). *Child Labor: Issues and Directions for the World Bank*. Directions in development. World Bank.

Faruqee, R. (2010). Microfinance for agriculture in Bangladesh: Current status and future potential. Working Paper 8, Institute of Microfinance (InM).

Gunnarsson, V., Orazem, P. F., and Sa´nchez, M. A. (2006). Child labor and school achievement in Latin America. *The World Bank Economic Review*, 20(1):31–54.

Gupta, M. R. (2001). Child labour, skill formation and capital accumulation: A theoretical analysis. *Keio economic studies*, 38(2):23–40.

Hazarika, G. and Sarangi, S. (2008). Household access to microcredit and child work in rural Malawi. *World Development*, 36(5):843–859.

Heady, C. (2003). The effect of child labor on learning achievement. *World Development*, 31(2):385–398.

Hoddinott, J. and Haddad, L. (1995). Does female income share influence household expenditures? evidence from coˆte d'ivoire. *oxford Bulletin of Economics and Statistics*, 57(1):77–96.

Hossain, M. and Bayes, A. (2009). *Rural Economy and Livelihoods: Insights from Bangladesh*. Dhaka, Bangladesh: AH Development Publishing.

Hossain, M., Malek, M., Hossain, M. A., Reza, M. H., and Ahmed, M. S. (2018). Agricultural microcredit for tenant farmers: Evidence from a field experiment in Bangladesh. *American Journal of Agricultural Economics*.

Islam, A. and Choe, C. (2013). Child labor and schooling responses to access to microcredit in rural Bangladesh. *Economic Inquiry*, 51(1):46–61.

Jacoby, H. G. (1994). Borrowing constraints and progress through school: Evidence from Peru. *The Review of Economics and Statistics*, 76(1):151–160.

Jacoby, H. G. and Skoufias, E. (1997). Risk, financial markets, and human capital in a developing country. *The Review of Economic Studies*, 64(3):311–335.

Krueger, A. B. (1996). Observations on international labor standards and trade. Working Paper 5632, National Bureau of Economic Research.

Kumar, C. S., Turvey, C. G., and Kropp, J. D. (2013). The impact of credit constraints on farm households: Survey results from India and China. *Applied Economic Perspectives and Policy*, 35(3):508–527.

Li, T., Sekhri, S., et al. (2013). The unintended consequences of employment-based safety net programs. *World Bank Economic Review (forthcoming)*.

Lochner, L. and Monge-Naranjo, A. (2012). Credit constraints in education. *Annu. Rev. Econ.*, 4(1):225–256.

Malek, M., Ahasan, A., Hossain, M., Ahmed, M., Hossain, M., and Reza, M. (2015). Impact assessment of credit programme for the tenant farmers in Bangladesh, 3ie grantee final report. *New Delhi: International Initiative for Impact Evaluation (3ie)*.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more t in experiments. *Journal of development Economics*, 99(2):210–221.

Muller, C. (2014). A test of separability of consumption and production decisions of farm households in Ethiopia.

Olken, B. A. (2015). Promises and perils of pre-analysis plans. *Journal of Economic Perspectives*, 29(3):61–80.

Psacharopoulos, G. (1997). Child labor versus educational attainment some evidence from Latin America. *Journal of population economics*, 10(4):377–386.

Reyes, A. and Lensink, R. (2011). The credit constraints of market-oriented farmers in Chile. *Journal of Development Studies*, 47(12):1851–1868.

Rosati, F. C. and Rossi, M. (2003). Children's working hours and school enrollment: Evidence from Pakistan and Nicaragua. *The World Bank Economic Review*, 17(2):283–295.

Shah, M. and Steinberg, B. M. (2015). Workfare and human capital investment: Evidence from India. Technical report, National Bureau of Economic Research.

Tarozzi, A., Desai, J., and Johnson, K. (2015). The impacts of microcredit: Evidence from Ethiopia. *American Economic Journal: Applied Economics*, 7(1):54–89.

Taubman, S. L., Allen, H. L., Wright, B. J., Baicker, K., and Finkelstein, A. N. (2014). Medicaid increases emergency-department use: Evidence from Oregon's health insurance experiment. *Science*, 343(6168):263–268.

Udry, C. (1996). Gender, agricultural production, and the theory of the household. *Journal of Political Economy*, 104(5):1010–1046.

Wydick, B. (1999). The effect of microenterprise lending on child schooling in Guatemala. *Economic Development and Cultural Change*, 47(4):853–869.

Table 1: Baseline Summary Statistics and Tests of Balance

	Baseline Statistics for the Control Group		Differences in Baseline Means between Treatment and Control Groups	
	Mean	SD	Difference	p-values
	(1)	(2)	(3)	(4)
Household Composition				
Number of HH members	4.77	1.69	0.176	0.330
Number of adults (>=15)	3.04	1.29	-0.01	0.95
Number of dependents (<15)	1.73	1.24	0.18	0.13
HH head is female	0.05	0.22	0.03*	0.08
HH head's age	44.57	12.09	0.72	0.30
HH head's years of schooling	3.02	3.35	0.19	0.56
Credit market participation				
Bank/Co-operatives	0.05	0.21	-0.01	0.19
NGO	0.09	0.3	-0.03*	0.05
Informal	0.04	0.19	-0.01	0.16
Any credit	0.17	0.38	-0.06***	0.01
Amount of credit Taken (in USD)				
Bank/Co-operatives	18.03	149.8	1.11	0.89
NGO	26.99	160.6	-10.39	0.10
Informal	21.95	224.3	-0.88	0.94
Any credit	66.98	328.3	-10.16	0.59
Amount of land (in Decimals)				
Share-cropped land	34.36	64.85	-5.37	0.31
Leased-in	7.64	44.73	5.49	0.41
Other rental arrangements	9.25	26.53	0.17	0.92
Total Rented in	51.25	78.6	0.29	0.97
Owned land	38.71	51.83	-1.25	0.70
Cultivated land	89.95	88.92	-0.96	0.92
Non-farm self-employment activities				
HH participates	0.23	0.42	-0.00	0.98
Number of activities	0.26	0.51	0.00	0.97
HH uses child labor	0.16	0.37	0.01	0.80
Number of hours of supplied by children	0.78	2.76	-0.02	0.91
Annual expenditure (in USD)				
Food expenditure per capita	151.8	59.3	2.15	0.74
Non-food expenditure per capita	113.7	86.9	-8.90	0.19
Education expenditure per children	29.4	62.9	0.95	0.81
Schooling Outcome				
Number of children never attended schools	0.02	0.15	0.0004	0.006
Number of children stopped attending schools	0.08	0.3	-0.001	0.15

Notes: Data from baseline (2012) survey. Sample size is n = 4,141, of which 2,072 assigned to treatment and 2,069 assigned to control. Columns 1 and 2 report statistics for households in the control areas. Column 3 shows the difference between the mean for households in the treatment area and the means in Column 1. Column 4 shows p-values for the test of equality of means, robust to intra-cluster correlation. The number of clusters (sub-districts) is 40. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) percent level. All figures expressing monetary values are in US Dollar. Unit of land is in decimal, where 100 decimals=1 acre. Informal lenders include moneylenders, loans from friends or family, and buying goods or services on credit from sellers.

Table 2: Endline Attrition

Panel A: End line attrition in treatment vs. control	
Found in the end line survey, in control	0.964
Found in the end line survey, in treatment	0.961
p-value of difference	0.648

Panel B: Compositional changes in sample at endline

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	HH size	Number of dependents	Age of head	Education of head	Head is female	Owned land	Total cultivated land
Treatment	0.176	0.181	0.724	0.178	0.0323*	-1.251	5.492
	(0.178)	(0.116)	(0.694)	(0.320)	(0.0175)	(3.221)	(10.96)
HH not found at endline	-0.545***	-0.0511	-1.675	0.436	0.0916*	3.550	5.451
	(0.148)	(0.137)	(1.819)	(0.327)	(0.0524)	(7.826)	(16.54)
Treatment x HH not found at endline	0.254	-0.0369	-3.020	-0.103	-0.0787	1.887	-7.528
	(0.250)	(0.194)	(2.255)	(0.573)	(0.0662)	(10.27)	(19.26)
Constant	4.766***	1.726***	44.57***	3.032***	0.0512***	38.71***	90.26***
	(0.130)	(0.0963)	(0.484)	(0.223)	(0.00796)	(2.345)	(6.413)
Observations	4,301	4,301	4,301	4,301	4,301	4,301	4,301

Notes: Panel A reports the percentage of households contacted for end line, among those on listing sheets based on the baseline survey. Panel B shows that there were no compositional changes in sample due to attrition. “HH not found at endline” is an indicator for households that were surveyed in baseline, but could not be found during the endline survey. Cluster-robust standard errors (at the sub-district level) in parentheses. There are 40 clusters. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) percent level.

Table 3: Impact of BCUP on Credit Market Participation

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	BCUP	Bank/ Cooperative	NGO	Informal	Any credit other than BCUP	Any credit including BCUP
Panel A: Probability of take-up						
Treatment	0.201***	0.000	-0.017	-0.002	-0.015	0.145***
	(0.027)	(0.009)	(0.029)	(0.014)	(0.032)	(0.031)
Lag of the dependent variable	0.000	0.099***	0.352***	0.073***	0.238***	0.258***
	(0.000)	(0.030)	(0.034)	(0.023)	(0.026)	(0.027)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	0	0.0353	0.203	0.0387	0.268	0.268
Panel B: Amount Borrowed (in USD)						
Treatment	77.99***	5.28	0.51	18.16	25.04	103.09**
	(11.06)	(10.82)	(18.02)	(29.77)	(43.50)	(40.43)
Lag of the dependent variable	0.00	0.60**	0.29**	0.10	0.35**	0.35**
	(0.00)	(0.28)	(0.11)	(0.06)	(0.14)	(0.13)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	0	25.38	81.90	27.47	134.8	134.8

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4,141. The table presents the coefficient of a "treatment" dummy in a regression of each variable on treatment (Equation 1 in the text). Errors are clustered at the branch (sub-district) level. The dependent variables in Columns 1-6 of Panel A are defined as follows: a dummy for whether the household had an outstanding loan from BCUP (Column 1), or from banks or co-operatives (2), or from Non-Government Organizations (NGOs) such as Grameen Bank, BRAC programs other than BCUP, and other NGOs (3), or from informal sources such as money lenders or other individuals such as family and friends (4), or if a household had a loan from any source other than BCUP (5), or if a household had a loan from any source including BCUP (6) . The dependent variables in Columns 1-6 of panel B are the amounts corresponding to the loans defined in the column headers. The Endline control means reported at the bottom of each panel are calculated for the control areas randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level. All figures expressing monetary values are in USD. Informal lenders include moneylenders, loans from friends or family, and buying goods or services on credit from sellers.

Table 4: Impact of credit on Amount of Cultivated Land (in Decimal)

	Rented-in Land				Own land	Total cultivated land
	Share-cropping	Fixed rental	Others	Total ((1)+(2)+(3))		Column (4) + Column (5)
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-2.01	6.58**	0.18	5.89	0.12	5.98
	(2.85)	(3.00)	(1.88)	(4.04)	(2.43)	(5.33)
Lag of the dependent variable	0.46***	0.68***	0.40***	0.64***	0.68***	0.67***
	(0.05)	(0.07)	(0.05)	(0.06)	(0.04)	(0.05)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	26.82	7.421	10.39	44.64	34.17	78.80

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "treatment" dummy in a regression of each variable on treatment (Equation 1 in the text). Errors are clustered at the branch (sub-district) level. The dependent variables in Column 1 and 2 shows the amount of land rented-in under share-cropping arrangement (Column 1) and under fixed-rental arrangement (column 2). Column 3 shows total amount of rented-in land under any type of tenancy arrangement. The dependent variable in column 4 shows the amount of owned cultivated land. The dependent variable in column 5 shows the amount of total cultivated land, which is the summation of owned cultivated land and rented-in land. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. In all cases the unit of measurement of land size is in Decimals (1 decimal is equal to 1/100 acre). Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table 5: Impact of Credit on Non-farm Self-Employment Activities

	(1)	(2)	(3)	(4)	(5)
VARIABLES	HH participates in non-farm self-employment activities	Number of non-farm self-employment activities	Number of family labor	Number of hired labor	Current market price of all business assets (in USD)
Treatment	0.0634***	0.07**	0.11**	0.01	179.78**
	(0.0230)	(0.03)	(0.05)	(0.03)	(80.84)
Lag of the dependent variable	0.475***	0.46***	0.40***	0.01	0.03*
	(0.0357)	(0.04)	(0.05)	(0.01)	(0.02)
Observations	4,141	4,141	4,141	4,141	4,141
Endline control mean	0.174	0.189	0.215	0.0715	645.1

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "treatment" dummy in a regression of each variable on treatment (Equation 1 in the text). Errors are clustered at the branch (sub-district) level. Business outcomes are aggregated at the household level when the households have more than one business. The outcome variables are set to zero when the household does not have a business. Non-farm self-employment is defined as self-employment in non-agricultural activities like agro-processing industries, wholesale and retail trading, storage and communication, transport and education, health industries and other service-related activities. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table 6: Impact of Credit on Different Proxies of Child Labor (5-14) in Self-Employment Activities

	(1)	(2)	(3)
VARIABLES	HH employed child labor	HH did not use child labor in baseline, but used it in endline	Number of hothe worked by children
Treatment	0.07**	0.06*	0.34***
	(0.03)	(0.03)	(0.12)
Lag of the dependent variable	0.04***	0.00	0.01
	(0.02)	(0.00)	(0.01)
Observations	4,141	4,141	4,141
Endline control mean	0.0609	0.0459	0.203

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. Column 1 presents the coefficient of a "treatment" dummy in a regression of the probability that household employs child labor on treatment. Column 2 presents the coefficient of a "treatment" dummy in a regression of the probability that household employed child labor in endline, but not in baseline. Errors are clustered at the branch (sub-district) level. The dependent variable is a dummy variable taking a value of 1 if household employs child labor and 0 otherwise. The Endline control means are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table 7: Impact of Credit on Hours Worked by Children (5-14 Years) on Different Activities: Time Budget Survey

		Economic activities		Non-economic activities		
Wage/Salaried employment		Self-employment	HH chores	Study	Leisure	Other
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All children						
Treatment x Post	-0.36	0.72**	-0.15	-3.54**	1.73	1.59
	(0.39)	(0.33)	(0.53)	(1.37)	(1.29)	(1.50)
Observations	2,841	2,841	2,841	2,841	2,841	2,841
Endline control mean	0.713	0.596	1.858	20.69	29.33	114.8
Share of total hours	0.4%	0.4%	1.1%	12.3%	17.5%	68.3%
Panel B: Male children	-0.63					
Treatment x Post		1.64***	0.36	-4.70**	2.86	0.47
	(0.77)	(0.64)	(0.54)	(2.10)	(1.96)	(2.27)
Observations	1,313	1,313	1,313	1,313	1,313	1,313
Endline control mean	0.993	0.807	0.563	20.33	31.13	114.2
Share of total hours	0.6%	1.2%	0.3%	12.1%	18.5%	68.0%
Panel C: Female children						
Treatment x Post	-0.18	-0.04	-0.55	-2.47	0.65	2.59
	(0.31)	(0.28)	(0.86)	(1.80)	(1.69)	(2.01)
Observations	1,528	1,528	1,528	1,528	1,528	1,528
Endline control mean	0.471	0.413	2.984	21.01	27.77	115.4
Share of total hours	0.3%	0.2%	1.8%	12.5%	16.5%	68.7%

Notes: The analysis is at the *individual* level. Data are from 2012 and 2014 Time-budget surveys. This time budget survey was done for a sub-sample of the original sample. Each column presents the coefficient of a *Treatment* × *Post* dummy in a regression of weekly hours supplied by children aged between 5 and 14 on treatment dummy, Post dummy, and the interaction of treatment dummy with Post dummy. Errors are clustered at the branch (sub-district) level. The endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. Other non-economic activities include sleep, rest, taking care of younger siblings or sick persons, etc. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table 8: Heterogeneity in the Impact of Credit on the Use of Child Labor (5-14)

	HH used child labor		Weekly hours supplied by children	
VARIABLES	(1)	(2)	(3)	(4)
Treatment	0.07**	0.03	0.80***	0.46**
	(0.03)	(0.02)	(0.29)	(0.19)
Treatment x Number of working-age members	-0.02**		-0.14*	
	(0.01)		(0.07)	
Treatment x HH head is female		-0.12***		-1.26***
		(0.04)		(0.29)
Observations	1,262	1,262	1,262	1,262
Endline control mean	0.0940	0.0940	0.755	0.755

Notes: Data are from 2012 and 2014 Time budget surveys. The time budget survey was done for a sub-sample of the original sample. The analysis is at the household level. The table presents the coefficients of the "treatment" dummy and the interaction of the "treatment" dummy with the heterogeneity variable from regressions specified in Equation 2 in the text. The variables *Number of working-age members* and *HH head is female* are baseline values. The level variables are also included in the model. The dependent variable in column (1) is a dummy variable taking a value of 1 if household employs child labor in self-employment activities and 0 otherwise. The endline control means are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table 9: Impact of Credit on Annual Education Expenditure Per Child

	Education expenditure per child (in USD)		
VARIABLES	(1)	(2)	(3)
Treatment	-3.97	-20.20**	-4.93
	(7.54)	(7.63)	(7.69)
Lag of dependent variable	0.51***	0.50***	0.51***
	(0.07)	(0.07)	(0.07)
Treatment x Number of working-age members		5.35***	
		(1.81)	
Treatment x HH head is female			11.53
			(8.63)
Observations	4,141	4,141	4,141
Endline control mean	53.99	53.99	53.99

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. Column 1 presents the coefficient of a "treatment" dummy in a regression of each variable on treatment (Equation 1 in the text). Column 2 and 3 presents the coefficients obtained using regression Equation 2 in the text. The variables *Number of working-age members* and *HH head is female* are baseline values. These variables are also included in the model in level forms. Errors are clustered at the branch (sub-district) level. Education expenditure is yearly. The endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. All figures expressing monetary values are in USD. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table 10: Impact (ITT) of Credit on Child Schooling (Between 5 to 14 Years)

	Number of children never attended school			Number of children stopped attending school		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.008	0.008	-0.009	-0.001	0.006	-0.001
	(0.006)	(0.016)	(0.007)	(0.010)	(0.019)	(0.010)
Lag of the dependent variable	0.239***	0.240***	0.239***	0.265***	0.265***	0.265***
	(0.055)	(0.055)	(0.055)	(0.028)	(0.028)	(0.028)
Treatment x Number of working-age members		-0.006			-0.002	
		(0.005)			(0.006)	
Treatment x HH head is female			0.003			0.012
			(0.019)			(0.029)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	0.0295	0.0295	0.0295	0.0575	0.0575	0.0575

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Column 1 and 4 present the coefficient of a "treatment" dummy in a regression of each variable on treatment (Equation 1 in the text). Column 2, 3, 5, and 6 presents the coefficients obtained using regression Equation 2 in the text. The variables *Number of working-age members* and *HH head is female* are baseline values. These variables are also included in the model in level forms. Errors are clustered at the branch (sub-district) level. The endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

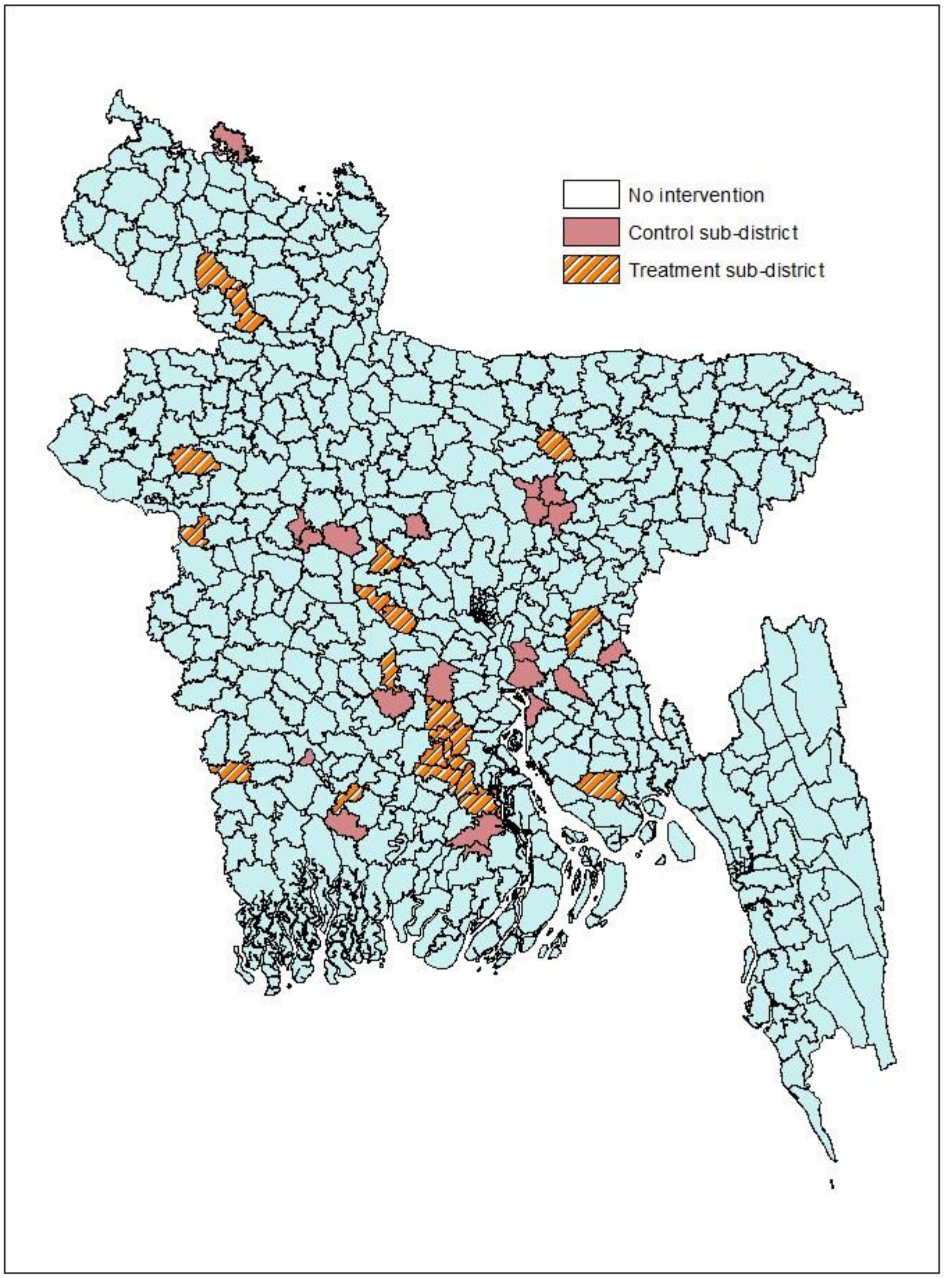


Figure 1: Study Area

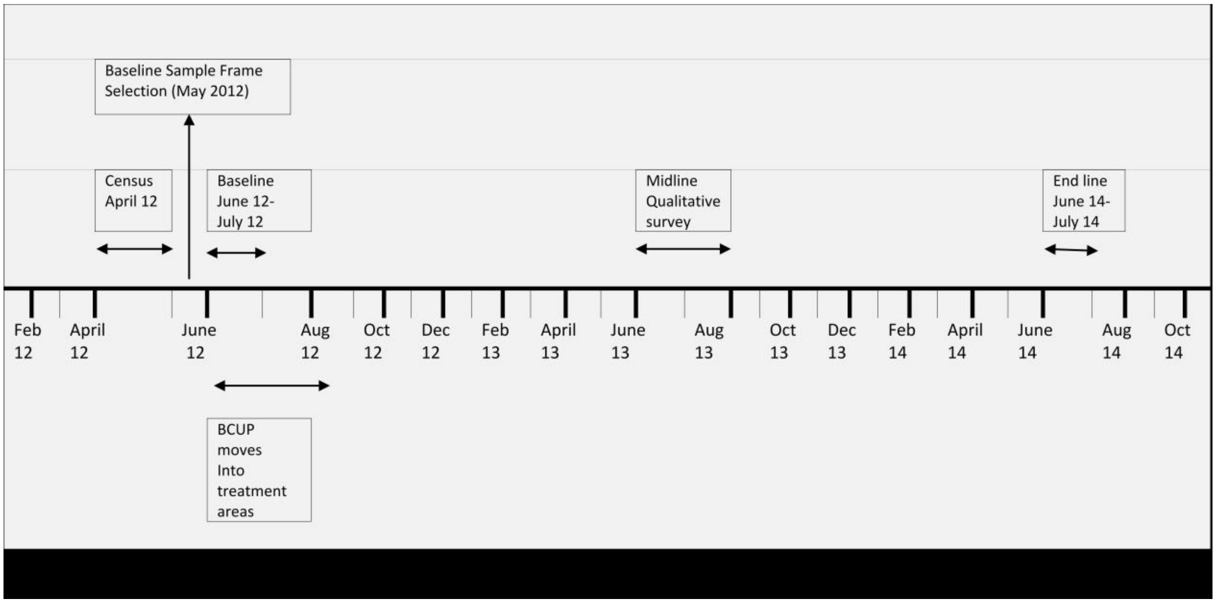


Figure 2: Timeline of intervention and data collection

Supplementary Appendix

A Appendix Figures and Tables

Table A1: Determinants of Endline Attrition

	(1)	(2)
VARIABLES	**HH Not found at the endline**	**HH Not found at the endline**
Treatment	0.00263	0.00307
	(0.00727)	(0.00777)
HH head is female		0.0273
		(0.0177)
HH head's years of schooling		0.000890
		(0.000877)
HH size		-0.00211
		(0.00217)
Number of rooms		0.00190
		(0.00384)
HH has electricity connection		0.00247
		(0.00523)
HH has concrete wall		0.00189
		(0.0101)
Amount of owned land (in decimals)		4.48e-05
		(6.47e-05)
Amount of cultivated land (in decimals)		1.09e-05
		(2.84e-05)
HH income		-1.90e-08
		(2.51e-08)
Constant	0.0359***	0.0345***
	(0.00524)	(0.0111)
Observations	4,301	4,301

Notes: "HH not found at endline" is an indicator for households that were surveyed in baseline, but could not be found during the endline survey. Amount of total cultivated land is the summation of owned cultivated land and rented-in land. In all cases the unit of measurement of land size is in Decimals (1 decimal is equal to 1/100 acre). Cluster-robust standard errors (at the branch level) in parentheses. There are 40 clusters. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) percent level.

Table A2: Determinants of Credit Uptake from the BCUP Program

	(1)
VARIABLES	Takes loan from BCUP program
Number of members of working age (15-64)	-0.00725
	(0.00854)
HH head is female	-0.0648*
	(0.0346)
Age of HH head	-0.00371***
	(0.000740)
Maximum years of schooling in HH	0.00170
	(0.00191)
Have outstanding credit	0.0884*
	(0.0460)
Owned land	-0.000274
	(0.000190)
Cultivated land	-2.48e-05
	(8.36e-05)
HH expenditure per capita (in USD)	0.000428***
	(0.000147)
Per day per capita calorie intake	2.35e-05
	(5.36e-05)
Per day per capita protein intake	0.00126
	(0.00175)
Distance to market	0.000428
	(0.00853)
Distance to upzila Sadr	0.0190**
	(0.00766)
House has concrete floor	-0.00808
	(0.0343)
HH has sanitary toilet	-0.0472**
	(0.0212)
Total income	-5.02e-09
	(8.64e-08)
Number of cows	-0.0833***
	(0.0222)
Number of goats	0.0366
	(0.0390)
Number of chickens	-0.0170
	(0.0266)
HH has water pump	-0.0322
	(0.0302)
Observations	2,072

Notes: Data from 2012 and 2014 surveys. The dependent variable is a binary variable taking a value of 1 if the household borrows from BCUP, and 0 otherwise. The independent variables are baseline figures. The analysis is at the household level. The sample is restricted to the treated households only. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table A3: Multinomial Logit specification of the impact of credit on child labor

	(1)	(2)	(3)
VARIABLES	Stopped using child labor	Unchanged	Started using child labor
Treatment	-0.00697	-0.0509*	0.0578**
	(0.0283)	(0.0261)	(0.0291)
Observations	4,141	4,141	4,141

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. The table presents the marginal effects from a Multinomial Regression of credit on a categorical variable with three categories: (a) Household used child labor in the baseline but not in endline; (b) Household status of using child labor is unchanged (Base category); and (c) Household did not use child labor in the baseline, but started using child labor in the endline. Errors are clustered at the branch (sub-district) level. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table A4: Impact of Credit on Hours Worked by Adults (15-64 years) on Different Activities: Time Budget Survey

	Economic activities			Non-economic activities		
	Wage/Salaried employment	Self-employment	HH chores	Study	Leisure	Other
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All adults (15-64 years)						
Treatment	0.46	2.61*	-1.90	0.78	0.81	-2.75
	(1.30)	(1.37)	(1.39)	(0.52)	(1.13)	(2.87)
Observations	3,376	3,376	3,376	3,376	3,376	3,376
Endline control mean	10.18	10.66	25.67	1.760	11.65	108.1
Share of total hours	6.1%	6.3%	15.3%	1.0%	6.9%	64.3%
Panel B: Male adults (15-64 years)						
Treatment	1.32	5.20***	-3.32**	0.59	1.67	-5.46*
	(2.17)	(1.91)	(1.26)	(0.70)	(1.61)	(2.81)
Observations	1,595	1,595	1,595	1,595	1,595	1,595
Endline control mean	19.96	16.15	6.140	1.802	16.87	107.1
Share of total hours	11.9%	9.6%	3.7%	1.1%	10.0%	63.8%
Panel B: Female adults (15-64 years)						
Treatment	0.71	0.97	-2.65	0.94*	0.62	-0.57
	(0.65)	(1.12)	(2.21)	(0.47)	(1.10)	(3.21)
Observations	1,781	1,781	1,781	1,781	1,781	1,781
Endline control mean	0.930	5.464	44.12	1.721	6.703	109.1
Share of total hours	0.6%	3.3%	26.3%	1.0%	4.0%	64.9%

The analysis is at the *individual* level. Data are from 2012 and 2014 Time-budget surveys. This time budget survey was done for a sub-sample of the original sample. Each column presents the coefficient of a *Treatment* × *Post* dummy in a regression of weekly hours supplied by adults aged between 15 and 64 on treatment dummy, Post dummy, and the interaction of treatment dummy with Post dummy. Errors are clustered at the branch (sub-district) level. The endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. Other non-economic activities include sleep, rest, taking care of children or sick persons, etc. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table A5: Impact of Credit on Non-farm Self-Employment Activities by the Sex of Household Head

	(1)	(2)	(3)	(4)	(5)
VARIABLES	HH participates in non-farm self-employment activities	Number of non-farm self-employment activities	Number of family labor	Number of hired labor	Current market price of all business assets (in USD)
Treatment	0.0655***	0.07**	0.11**	0.01	188.04**
	(0.0237)	(0.03)	(0.05)	(0.03)	(87.90)
Lag of the dependent variable	0.471***	0.46***	0.40***	0.01	0.03*
	(0.0354)	(0.04)	(0.05)	(0.01)	(0.02)
Treatment x HH head is female	0.00247	0.01	0.01	-0.00	-32.47
	(0.0278)	(0.03)	(0.05)	(0.03)	(110.14)
Observations	4,141	4,141	4,141	4,141	4,141
R-squared	0.249	0.27	0.19	0.00	0.01
Endline control mean	0.174	0.189	0.215	0.0715	645.1

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "treatment" dummy in a regression specified in Equation 2 in the text. The variable *HH head is female* is the baseline value, and also included in the model in level form. Errors are clustered at the branch (sub-district) level. Observations with inconsistent amount of assets are dropped in column 5. Business outcomes are aggregated at the household level when the households have more than one business. The outcome variables are set to zero when the household does not have a business. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table A6: Heterogeneity in Child labor by Baseline Child Labor Use

	(1)	(2)
VARIABLES	HH employed child labor	Number of hours worked by the children
Treatment	0.0698**	0.323***
	(0.0334)	(0.119)
HH employed child labor in baseline	0.0396*	0.221**
	(0.0201)	(0.102)
Treatment x HH employed child labor in baseline	0.0104	0.0845
	(0.0317)	(0.159)
Observations	4,141	4,141

Notes: Data from 2012 and 2014 Time-budget surveys. This time budget survey was done for a sub-sample of the original sample. The analysis is at household level. Column 1 and 2 present the coefficients obtained from the regression specified in Equation 2 in the text. Errors are clustered at the branch (sub-district) level. The endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table A7: Multiple Hypothesis Testing: Impact of Credit on Different Outcomes

	Productivity			Adoption of Modern Varieties			Income					Child Labor		Time spent by children on		
VARIABLES	Aman yield	Boro yield	Aggregate yield	Aman HYV	Aman hybrid	Boro - hybrid	Farm income	Wage income	Business income	Total income	Total expenditure	HH uses child labor	Weekly number of hours	Wage/ salaried employment	Self-employment	Study
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Treatment	0.82**	0.33	0.41*	0.15**	0.05***	0.08***	34.94	-36.05	61.00**	75.06	60.31	0.07**	0.34***	-0.36	0.72**	-3.54**
	(0.31)	(0.25)	(0.24)	(0.06)	(0.01)	(0.02)	(26.81)	(35.54)	(28.78)	(103.53)	(125.51)	(0.03)	(0.12)	(0.39)	(0.33)	(1.37)
P-value	0.01	0.20	0.10	0.02	0.00	0.00	0.20	0.32	0.04	0.47	0.63	0.04	0.01	0.36	0.03	0.01
BH-q value	0.032	0.267	0.16	0.054	0.001	0.001	0.267	0.394	0.072	0.502	0.63	0.072	0.032	0.412	0.069	0.032
Constant	0.46***	0.80***	1.58***	0.11***	0.00	0.02*	143.90***	313.24***	96.28***	1,023.55***	815.78***	0.05***	0.19***	0.21	1.14***	33.06***
	(0.15)	(0.28)	(0.26)	(0.03)	(0.00)	(0.01)	(18.97)	(22.43)	(19.16)	(119.41)	(105.93)	(0.01)	(0.04)	(0.21)	(0.17)	(0.73)
Observations	4,141	4,141	4,141	4,141	4,141	4,141	4,141	4,141	4,141	4,141	4,141	4,141	4,141	2,841	2,841	2,841

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. Standard errors are clustered at the branch (sub-district) level. Aggregate yield in column 3 is calculated as total production divided by total land cultivated in the Amon and Boro rice seasons. Amon is the rain-fed monsoon production period, in which rice is seeded during April1–May and harvested in November1–December. Boro is the irrigation intensive dry-season rice production period, in which rice is seeded during December1–February and harvested in April1–May. Columns (14) - (16) are estimated using time budget survey. This time budget survey was done for a sub-sample of the original sample. BH-q values are False Discovery Rate (FDR)-q values based on Benjamini and Hochberg procedure. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table A8: Impact of Credit on the Probability that Household Uses Child Labor (5-14) in Self-Employment Activities: Evidence from SVRS Survey

	(1)	(2)
VARIABLES	HH used child labor	HH used child labor
Treatment	-0.0112	-0.00839
	(0.0152)	(0.0135)
Post	-0.0283**	-0.0224**
	(0.0131)	(0.0107)
Treatment x Post	0.0405**	0.0342*
	(0.0198)	(0.0174)
Baseline control mean	0.04	0.04
Socio-economic controls	Yes	Yes
Observations	19,892	19,892

Notes: Data from 2012 and 2014 rounds of Sample Vital Registration Surveys (SVRS). Column 1 and 2 present the coefficients from a Difference-in-differences regression of the dependent variable on the treatment dummy, post dummy and the interaction of the two. Errors are clustered at the branch (sub-district) level. The dependent variable is a dummy variable taking a value of 1 if a household employs child labor in self-employment activities and 0 otherwise. Socio-economic controls include age and sex of the household head, and indicator variables for household having electricity, sanitary latrine and piped water supply. The baseline control mean is calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table A9: Impact of Credit on Hours Worked by Teens (15-19 years) on Different Activities: Time Budget Survey

	(1)	(2)	(3)
VARIABLES	Wage employment	Self-employment	HH Chores
Panel A: Adolescent (15-19 years)			
Treatment	-0.074	0.485*	0.329
	(0.302)	(0.269)	(0.345)
Lag of the dependent variable	0.124	0.102	0.066*
	(0.102)	(0.070)	(0.038)
Observations	1,263	1,263	1,263
Endline control mean	0.511	0.514	1.100
Panel B: Male Adolescent (15-19 years)			
Treatment	-0.074	0.561**	-0.067
	(0.302)	(0.246)	(0.110)
Lag of the dependent variable	0.124	0.104	0.003
	(0.102)	(0.078)	(0.005)
Observations	1,263	1,263	1,263
Endline control mean	0.511	0.309	0.165
Panel C: Female Adolescent (15-19 years)			
Treatment	0.000	-0.061	0.406
	(0.000)	(0.094)	(0.339)
Lag of the dependent variable	0.000	0.002	0.046
	(0.000)	(0.009)	(0.036)
Observations	1,263	1,263	1,263
Endline control mean	0	0.205	0.935

Notes: Data from 2012 and 2014 Time-budget surveys. This time budget survey was done for a sub-sample of the original sample. The analysis is at household level. Column 1 presents the coefficient of a "treatment" dummy in a regression of weekly hours supplied by children aged between five and 14 on treatment (Equation 1 in the text). Column 2 presents the coefficient of a "treatment" dummy in the regression Equation 2 in the text. Errors are clustered at the branch (sub-district) level. The endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit.
Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table A10: Summary of the Findings on the Impact of Microcredit on Child Labor from Other RCTs

	Characteristics of microcredit	Findings on child labor	Comment
Bosnia: (Ausberg et al., 2015)	*Gender of borrowers:* Male, Female *Targeted to Microentrepreneurs? :*Yes *Loan size as a proportion of income:* 9 percent *Interest rate:* 22 Percent (Market: 27.3% APR) *Liability:* Individual	Treated households use 0.53 (SD 0.23) hours more of teenage (16 to 19 years) labor supply per week on business activities than the control households.	For the comparable group of people in my study, I find an increase of similar magnitude (0.49 hours of labor supply per week) for the treated households than the control ones (Appendix Table XX)
Etheopia (Tarozzi et al., 2015)	*Gender of borrowers:* Male, Female *Targeted to Microentrepreneurs?* Yes *Loan size as a proportion of income:* 118 percent *Interest rate:* 12 Percent (Market: 24.7% APR) *Liability:* Individual	Find small and not significant impacts in the number of hours worked for children aged between 10-15 years.	While there are many similarities between this experiment with ours, the loan size in this experiment is considerably larger than ours. This liquidity could prompt households to use hired labor instead of child labor. For instance, in our study, the treatment effect of credit on child labor is much smaller for households with a higher baseline income
Mexico: (Angelucci et al., 2015)	*Gender of borrowers: Female* *Targeted to Microentrepreneurs?* Yes *Loan size as a proportion of income:* 6 percent *Interest rate:* 110 Percent APR (Market: 145% APR) *Liability:* Joint	The 95% confidence interval for the variable "fraction of children working" is (-0.020, 0.005), ruling out even small positive effects on child labor.	The loans in these two experiments were targeted towards females. Neither of these studies finds any significant impact of credit on child labor. This should not come as a surprise since earlier works suggest that women have different preferences compared to men, and the bargaining process often leads to better outcomes for the children, such as education, and health, when interventions are targeted towards women (Duflo, 2003; Hoddinott and Haddad,1995). In my paper, too, I find that the treatment effect of credit on the likelihood of using child labor is 5.9 percentage points lower for the female-headed households than the male-headed households.
India (Banerjee et al., 2015)	*Gender of borrowers:* Female *Targeted to Microentrepreneurs?:* No *Loan size as a proportion of income:* 22 percent *Interest rate:* 24 Percent (Market: 15.9% APR) *Liability:* Joint	Find no difference in the number of hours worked by girls or boys aged 5 to 15	

Table A11: Definition of Important Variables Used in the Paper

Variable	Definition
Child labor	The ILO Minimum Age Convention, 1973 (No. 138) sets the general minimum age at 14 (12 for lightwork) where the economy and educational facilities are insufficiently developed. Since Bangladesh is a developing country, this study defines labor supplied by children below 14 year ages as child labor
Number of working members	Number of members of working age (15-64 years)
Informal lenders	Informal lenders include moneylenders, loans from friends or family, and buying goods or services on credit from sellers
Share cropping	A system of agriculture or agricultural production in which a landowner leases his/her land to a tenant in return for a share of the crop produced on the land
Fixed rental contract	A system of agriculture or agricultural system wherein the landlord leases out his land to the tenant for cultivation for a fixed rent
Farm self-employment	Self-employment in agricultural activities
Non-farm self-employment	Self-employment in non-agricultural activities like agro-processing industries, wholesale and retail trading, storage and communication, transport and education , health industries and other service related activities
Household enterprise	Household enterprise includes both farm and non-farm employment. In this study household enterprise and self-employment activities have been used alternatively
Education expenditure	Education expenditure is the summation of expenditure on the following items: (a) institutional (e.g., school fees); (b) Books, exercise books, pen and pencils; (c) salary of private tutor; (d) school uniform; (e) other educational expenses

B LATE Estimates

B.1 Effect of BCUP Credit Uptake (Local Average Treatment Effect (LATE))

The ITT estimates measure the net effect of increased credit access. However, exploring the effect of credit itself on different outcomes can be of significant policy interest. To this end, I run a regression of the following form:

$$Y_{is,2014} = \pi_0 + \pi_1 BCUP_{is} + \pi_2 Y_{is,2012} + \epsilon_{is}$$

BCUP is a binary indicator variable, taking a value of 1 if anyone in the household received a loan from the BCUP program during the period of study, and 0, otherwise. I estimate the above equation by means of instrumental variable regression, running the first stage equation of the following form:

$$BCUP_{is,2014} = \delta_0 + \delta_1 Treatment_s + \delta_2 Y_{is,2012} + \eta_{is}$$

in which *Treatment* is the excluded instrument. The coefficient on BCUP from the instrumental variable estimation, π_1, is the local average treatment effect of BCUP credit. π_1 can be interpreted as the causal effect of credit among the subset of individuals who take credit upon being selected for treatment assignment, but who would not take credit if they were not selected for treatment assignment (i.e., the compliers), provided some assumptions are satisfied. The first condition for a valid instrument is the relevance condition, i.e., *Treatment* and *BCUP* are strongly correlated, which is indeed the case in this setting.[39]The second condition is the exclusion restriction condition. One implication of this condition is that there are no externalities between the compliers and the non-takers in the treated sub-districts. Such externalities would violate the exclusion restriction condition required for identification using instrumental variables (Barua and Lang, 2009). I report the LATE estimates in the Appendix Tables B1 to B5.

[39] The F-statistic from the F test of excluded instrument is 54.49 and can be rejected at the 1 percent level of significance.

Table B1: LATE Estimates of Impact of BCUP on Credit Market Participation

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	BCUP	Bank/ Cooperative	NGO	Informal	Any credit other than BCUP	Any credit including BCUP
Panel A: Probability of take-up						
BCUP credit uptake	1	0.002	-0.086	-0.012	-0.076	0.714***
	(0)	(0.045)	(0.140)	(0.070)	(0.155)	(0.131)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	0	0.0353	0.203	0.0387	0.268	0.268
Panel B: Amount Borrowed (in USD)						
BCUP credit uptake	388.44***	26.29	2.56	90.47	124.71	513.43**
	(26.37)	(53.73)	(88.52)	(152.46)	(221.85)	(225.44)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	0	25.38	81.90	27.47	134.8	134.8

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "BCUP credit uptake" dummy in an instrumental variable regression of each variable on BCUP credit uptake using treatment as an instrument for BCUP uptake. Errors are clustered at the branch (sub-district) level. There are 40 such clusters. The dependent variables in Columns 1-6 of Panel A are defined as follows: a dummy for whether the household had an outstanding loan from BCUP (Column 1), or from banks or co-operatives (2), or from Non-Government Organizations (NGOs) such as Grameen Bank, BRAC programs other than BCUP, and other NGOs (3), or from informal sources such as money lenders or other individuals such as family and friends (4), or if a household had a loan from any source other than BCUP (5), or if a household had a loan from any source including BCUP (6) . The dependent variables in Columns 1-6 of panel B are the amounts corresponding to the loans defined in the column headers. The Endline Control Mean reported at the bottom of each panel are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level. All figures expressing monetary values are in USD. Informal lender includes moneylenders, loans from friends/family, and buying goods/services on credit from seller.

Table B2: LATE Estimates of the Impact of credit on Amount of Cultivated Land (in Decimal)

	Rented-in Land				Own land	Total cultivated land
	Share-cropping	Fixed rental	Others	Total ((1)+(2)+(3))		Column (4) + Column (5)
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
BCUP credit uptake	-9.97	32.71**	0.92	29.35	0.59	29.80
	(14.42)	(15.57)	(9.23)	(19.95)	(11.97)	(26.50)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	26.82	7.421	10.39	44.64	34.17	78.80

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "BCUP credit uptake" dummy in an instrumental variable regression of each variable on BCUP credit uptake using treatment as an instrument for BCUP uptake. Standard errors are clustered at the sub-district level. There are 40 such clusters. The dependent variables in Column 1 and 2 shows the amount of land rented-in under share-cropping arrangement (Column 1) and under fixed-rental arrangement (column 2). Column 3 shows total amount of rented-in land under any type of tenancy arrangement. The dependent variable in column 4 shows the amount of owned cultivated land. The dependent variable in column 6 shows the amount of total cultivated land, which is the summation of owned cultivated land and rented-in land. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. In all cases the unit of measurement of land size is in Decimals (1 decimal is equal to 1/100 acre). Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level

Table B3: LATE Estimates of the Impact of Credit on Non-farm Self-Employment Activities

	(1)	(2)	(3)	(4)	(5)
VARIABLES	HH participates in non-farm self-employment activities	Number of non-farm self-employment activities	Number of family labor	Number of hired labor	Current market price of all business assets (in USD)
BCUP credit uptake	0.315***	0.35**	0.54**	0.06	893.50**
	(0.110)	(0.14)	(0.23)	(0.14)	(422.05)
Observations	4,141	4,141	4,141	4,141	4,141
Endline control mean	0.174	0.189	0.215	0.0715	645.1

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141.The table presents the coefficient of a "BCUP credit uptake" dummy in an instrumental variable regression of each variable on BCUP credit uptake using treatment as an instrument for BCUP uptake. Errors are clustered at the branch (sub-district) level. There are 40 such clusters. Observations with inconsistent amount of assets are dropped in columns 5. Business outcomes are aggregated at the household level when the households have more than one business. The outcome variables are set to zero when the household does not have a business. The Endline control means are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table B4: LATE Estimates of the Impact of Credit on Different Proxies of Child Labor (5-14 Years) in Self-Employment Activities

	(1)	(2)	(3)
VARIABLES	HH employed child labor	HH did not use child labor in baseline, but used it in endline	Number of hothe worked by children
BCUP credit uptake	0.36**	0.29*	1.69***
	(0.17)	(0.15)	(0.63)
Observations	4,141	4,141	4,141
Endline control mean	0.0609	0.0459	0.203

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "BCUP credit uptake" dummy in an instrumental variable regression of each variable on BCUP credit uptake using treatment as an instrument for BCUP uptake. Column 1 presents the coefficient of a "treatment" dummy in a regression of the probability that household employs child labor on treatment. Column 2 presents the coefficient of a "treatment" dummy in a regression of the probability that household employed child labor in endline, but not in baseline. Errors are clustered at the branch (sub-district) level. There are 40 such clusters. The Endline control means are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

C Difference-in-Differences (DiD) Estimates

Table C1: DiD Estimates of Impact of BCUP on Credit Market Participation

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	ΔBCUP	ΔBank/ Cooperative	ΔNGO	ΔInformal	ΔAny credit other than BCUP	ΔAny credit including BCUP
Panel A: Probability of take-up						
Treatment	0.201***	0.014	0.002	0.011	0.027	0.187***
	(0.027)	(0.013)	(0.028)	(0.012)	(0.035)	(0.032)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	0	0.0353	0.203	0.0387	0.268	0.268
Panel B: Amount Borrowed (in USD)						
Treatment	77.99***	4.83	7.87	18.95	31.66	109.65***
	(11.06)	(11.24)	(17.24)	(27.57)	(42.41)	(39.73)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	0	25.38	81.90	27.47	134.8	134.8

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "treatment" dummy in a regression of the change in each outcome variable between the endline and baseline on the treatment dummy. Errors are clustered at the branch (sub-district) level. There are 40 such clusters. The dependent variables in Columns 1-6 of Panel A are defined as follows: a dummy for whether the household had an outstanding loan from BCUP (Column 1), or from banks or co-operatives (2), or from Non-Government Organizations (NGOs) such as Grameen Bank, BRAC programs other than BCUP, and other NGOs (3), or from informal sources such as money lenders or other individuals such as family and friends (4), or if a household had a loan from any source other than BCUP (5), or if a household had a loan from any source including BCUP (6) . The dependent variables in Columns 1-6 of panel B are the amounts corresponding to the loans defined in the column headers. The Endline Control Mean reported at the bottom of each panel are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level. All figures expressing monetary values are in USD. Informal lender includes moneylenders, loans from friends/family, and buying goods/services on credit from seller.

Table C2: DiD Estimates of the Impact of credit on Amount of Cultivated Land (in Decimal)

	ΔRented-in Land				ΔOwn land	ΔTotal cultivated land
	ΔShare-cropping	ΔFixed rental	ΔOthers	ΔTotal ((1)+(2)+(3))		Column (4) + Column (5)
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.89	4.81*	0.08	5.79	0.51	6.30
	(2.41)	(2.84)	(1.64)	(3.86)	(2.50)	(5.07)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	26.82	7.421	10.39	44.64	34.17	78.80

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "treatment" dummy in a regression of the change in each outcome variable between the endline and baseline on the treatment dummy. Errors are clustered at the branch (sub-district) level. There are 40 such clusters. The dependent variables in Column 1 and 2 shows the amount of land rented-in under share-cropping arrangement (Column 1) and under fixed-rental arrangement (column 2). Column 3 shows total amount of rented-in land under any type of tenancy arrangement. The dependent variable in column 4 shows the amount of owned cultivated land. The dependent variable in column 6 shows the amount of total cultivated land, which is the summation of owned cultivated land and rented-in land. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. In all cases the unit of measurement of land size is in Decimals (1 decimal is equal to 1/100 acre). Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level

Table C3: DiD Estimates of the Impact of Credit on Non-farm Self-Employment Activities

	(1)	(2)	(3)	(4)	(5)
VARIABLES	ΔHH participates in non-farm self-employment activities	ΔNumber of non-farm self-employment activities	ΔNumber of family labor	ΔNumber of hired labor	ΔCurrent market price of all business assets (in USD)
Treatment	0.0638*	0.07	0.08	0.05	760.11
	(0.0332)	(0.04)	(0.06)	(0.07)	(571.40)
Observations	4,141	4,141	4,141	4,141	4,141
Endline control mean	0.174	0.189	0.215	0.0715	645.1

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141.The table presents the coefficient of a "BCUP credit uptake" dummy in a regression of the change in each outcome variable between the endline and baseline on the treatment dummy. Errors are clustered at the branch (sub-district) level. There are 40 such clusters. Observations with inconsistent amount of assets are dropped in column 5. Business outcomes are aggregated at the household level when the households have more than one business. The outcome variables are set to zero when the household does not have a business. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table C4: DiD Estimates of the Impact of Credit on Different Proxies of Child Labor (5-14 Years) in Self-Employment Activities

	(1)	(2)
VARIABLES	ΔHH employed child labor	ΔNumber of hours worked by the children
Treatment	0.065	0.358
	(0.051)	(0.245)
Observations	4,141	4,141
Endline control mean	0.0609	0.203

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "treatment" dummy in a regression of the change in each outcome variable between the endline and baseline on the treatment dummy. Column 1 presents the coefficient of a "treatment" dummy in a regression of the probability that household employs child labor on treatment. Column 2 presents the coefficient of a "treatment" dummy in a regression of the probability that household employed child labor in endline, but not in baseline. Errors are clustered at the branch (sub-district) level. There are 40 such clusters. The dependent variable is a dummy variable taking a value of 1 if household employs child labor and 0 otherwise. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

D Regressions Using Only Endline Data

Table D1: Impact of BCUP on Credit Market Participation Using Endline Data Only

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	BCUP	Bank/ Cooperative	NGO	Informal	Any credit other than BCUP	Any credit including BCUP
Panel A: Probability of take-up						
Treatment	0.201***	-0.001	-0.028	-0.003	-0.029	0.130***
	(0.027)	(0.009)	(0.031)	(0.015)	(0.033)	(0.033)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	0	0.0353	0.203	0.0387	0.268	0.268
Panel B: Amount Borrowed (in USD)						
Treatment	77.99***	5.94	-2.52	18.07	21.50	99.49**
	(11.06)	(12.17)	(18.29)	(30.25)	(44.99)	(41.85)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	0	25.38	81.90	27.47	134.8	134.8

Notes: Data from 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "treatment" dummy in a regression of each outcome variable on the treatment dummy. Errors are clustered at the branch (sub-district) level. There are 40 such clusters. **No controls have been used.** The dependent variables in Columns 1-6 of Panel A are defined as follows: a dummy for whether the household had an outstanding loan from BCUP (Column 1), or from banks or co-operatives (2), or from Non-Government Organizations (NGOs) such as Grameen Bank, BRAC programs other than BCUP, and other NGOs (3), or from informal sources such as money lenders or other individuals such as family and friends (4), or if a household had a loan from any source other than BCUP (5), or if a household had a loan from any source including BCUP (6) . The dependent variables in Columns 1-6 of panel B are the amounts corresponding to the loans defined in the column headers. The Endline Control Mean reported at the bottom of each panel are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level. All figures expressing monetary values are in USD. Informal lender includes moneylenders, loans from friends/family, and buying goods/services on credit from seller.

Table D2: Impact of credit on Amount of Cultivated Land (in Decimal) Using Endline Data Only

	Rented-in Land				Own land	Total cultivated land
	Share-cropping	Fixed rental	Others	Total ((1)+(2)+(3))		Column (4) + Column (5)
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-4.48	10.30	0.25	6.08	-0.74	5.34
	(4.91)	(6.36)	(2.30)	(7.25)	(3.48)	(9.48)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	26.82	7.421	10.39	44.64	34.17	78.80

Notes: Data from 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "treatment" dummy in a regression of each outcome variable on the treatment dummy. Errors are clustered at the branch (sub-district) level. There are 40 such clusters. **No controls have been used.** The dependent variables in Column 1 and 2 shows the amount of land rented-in under share-cropping arrangement (Column 1) and under fixed-rental arrangement (column 2). Column 3 shows total amount of rented-in land under any type of tenancy arrangement. The dependent variable in column 5 shows the amount of owned cultivated land. The dependent variable in column 6 shows the amount of total cultivated land, which is the summation of owned cultivated land and rented-in land. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. In all cases the unit of measurement of land size is in Decimals (1 decimal is equal to 1/100 acre). Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level

Table D3: Impact of Credit on Non-farm Self-Employment Activities Using Endline Data Only

	(1)	(2)	(3)	(4)	(5)
VARIABLES	HH participates in non-farm self-employment activities	Number of non-farm self-employment activities	Number of family labor	Number of hired labor	Current market price of all business assets (in USD)
Treatment	0.0630***	0.07**	0.12**	0.01	159.63**
	(0.0241)	(0.03)	(0.05)	(0.03)	(80.36)
Observations	4,141	4,141	4,141	4,141	4,141
Endline control mean	0.174	0.189	0.215	0.0715	645.1

Notes: Data from 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "treatment" dummy in a regression of each outcome variable on the treatment dummy. Errors are clustered at the branch (sub-district) level. There are 40 such clusters. **No controls have been used.** Observations with inconsistent amount of assets are dropped in column 5. Business outcomes are aggregated at the household level when the households have more than one business. The outcome variables are set to zero when the household does not have a business. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table D4: Impact of Credit on Different Proxies of Child Labor (5-14 Years) in Self-Employment Activities Using Endline Data Only

	(1)	(2)
VARIABLES	HH employed child labor	Number of hours worked by the children
Treatment	0.072**	0.339***
	(0.033)	(0.121)
Observations	4,141	4,141
Endline control mean	0.0609	0.203

Notes: Data from 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "treatment" dummy in a regression of each outcome variable on the treatment dummy. Errors are clustered at the branch (sub-district) level. There are 40 such clusters. **No controls have been used.** Column 1 presents the coefficient of a "treatment" dummy in a regression of the probability that household employs child labor on treatment. Column 2 presents the coefficient of a "treatment" dummy in a regression of the probability that household employed child labor in endline, but not in baseline. The dependent variable is a dummy variable taking a value of 1 if household employs child labor and 0 otherwise. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

E Regression Using Lasso Controls

Table E1: Impact of BCUP on Credit Market Participation Using Lasso Controls

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	BCUP	Bank/ Cooperative	NGO	Informal	Any credit other than BCUP	Any credit including BCUP
Panel A: Probability of take-up						
Treatment	0.199***	-0.001	-0.011	-0.003	-0.014	0.147***
	(0.027)	(0.009)	(0.028)	(0.014)	(0.031)	(0.030)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	0	0.0353	0.203	0.0387	0.268	0.268
Panel B: Amount Borrowed (in USD)						
Treatment	76.40***	5.94	0.09	18.07	24.34	104.06***
	(10.63)	(12.02)	(17.23)	(29.86)	(40.50)	(37.34)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	0	25.38	81.90	27.47	134.8	134.8

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "treatment" dummy in a regression of each outcome variable on the treatment dummy. I control for the baseline variables using the Lasso method proposed by Belloni et al. (2014). Errors are clustered at the branch (sub-district) level. There are 40 such clusters. The dependent variables in Columns 1-6 of Panel A are defined as follows: a dummy for whether the household had an outstanding loan from BCUP (Column 1), or from banks or co-operatives (2), or from Non-Government Organizations (NGOs) such as Grameen Bank, BRAC programs other than BCUP, and other NGOs (3), or from informal sources such as money lenders or other individuals such as family and friends (4), or if a household had a loan from any source other than BCUP (5), or if a household had a loan from any source including BCUP (6) . The dependent variables in Columns 1-6 of panel B are the amounts corresponding to the loans defined in the column headers. The Endline Control Mean reported at the bottom of each panel are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level. All figures expressing monetary values are in USD. Informal lender includes moneylenders, loans from friends/family, and buying goods/services on credit from seller.

Table E2: Impact of credit on Amount of Cultivated Land (in Decimal) Using Lasso Controls

	Rented-in Land				Own land	Total cultivated land
	Share-cropping	Fixed rental	Others	Total ((1)+(2)+(3))		Column (4) + Column (5)
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-1.84	6.58**	0.17	6.39	-0.06	5.89
	(2.74)	(2.96)	(1.83)	(4.01)	(2.43)	(5.53)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	26.82	7.421	10.39	44.64	34.17	78.80

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "treatment" dummy in a regression of each outcome variable on the treatment dummy. I control for the baseline variables using the Lasso method proposed by Belloni et al. (2014). The dependent variables in Column 1 and 2 shows the amount of land rented-in under share-cropping arrangement (Column 1) and under fixed-rental arrangement (column 2). Column 3 shows total amount of rented-in land under any type of tenancy arrangement. The dependent variable in column 5 shows the amount of owned cultivated land. The dependent variable in column 6 shows the amount of total cultivated land, which is the summation of owned cultivated land and rented-in land. Errors are clustered at the sub-district level. There are 40 such clusters. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. In all cases the unit of measurement of land size is in Decimals (1 decimal is equal to 1/100 acre). Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level

Table E3: Impact of Credit on Non-farm Self-Employment Activities Using Lasso Controls

	(1)	(2)	(3)	(4)	(5)
VARIABLES	HH participates in non-farm self-employment activities	Number of non-farm self-employment activities	Number of family labor	Number of hired labor	Current market price of all business assets (in USD)
Treatment	0.0596***	0.07**	0.10**	0.01	146.89**
	(0.0219)	(0.03)	(0.05)	(0.03)	(62.96)
Observations	4,141	4,141	4,141	4,141	4,141
Endline control mean	0.174	0.189	0.215	0.0715	645.1

Note: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4141. The table presents the coefficient of a "treatment" dummy in a regression of each outcome variable on the treatment dummy. I control for the baseline variables using the Lasso method proposed by Belloni et al. (2014). Errors are clustered at the branch (sub-district) level. There are 40 such clusters. Observations with inconsistent amount of assets are dropped in column 5. Business outcomes are aggregated at the household level when the households have more than one business. The outcome variables are set to zero when the household does not have a business. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table E4: Impact of Credit on Different Proxies of Child Labor (5-14 Years) in Self-Employment Activities Using Lasso Controls

	(1)	(2)	(3)
VARIABLES	HH employed child labor	HH did not use child labor in baseline, but used it in endline	Number of hothe worked by children
Treatment	0.06**	0.05**	0.31***
	(0.03)	(0.02)	(0.12)
Observations	4,141	4,141	4,141
Endline control mean	0.0609	0.0459	0.203

Note: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4,141. The table presents the coefficient of a "treatment" dummy in a regression of each outcome variable on the treatment dummy. I control for the baseline variables using the Lasso method proposed by Belloni et al. (2014). Column 1 presents the coefficient of a "treatment" dummy in a regression of the probability that household employs child labor on treatment. Column 2 presents the coefficient of a "treatment" dummy in a regression of the probability that household employed child labor in endline, but not in baseline. Errors are clustered at the branch (sub-district) level. There are 40 such clusters. The dependent variable in column (1) is a dummy variable taking a value of 1 if household employs child labor and 0 otherwise. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

F Regression Using Wild Cluster Bootstrap-t Procedure

Table F1: Impact of BCUP on Credit Market Participation Using Wild Cluster Bootstrap-t Procedure

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	BCUP	Bank/ Cooperative	NGO	Informal	Any credit other than BCUP	Any credit including BCUP
Panel A: Probability of take-up						
Treatment	0.201***	0.000	-0.017	-0.002	-0.015	0.145***
	(0.00)	(0.98)	(0.54)	(0.898)	(0.646)	(0.00)
Lag of the dependent variable		0.099***	0.352***	0.073***	0.238***	0.258***
		(0.002)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
R-squared	0.112	0.011	0.062	0.004	0.038	0.056
Endline control mean	0	0.0353	0.203	0.0387	0.268	0.268
Panel B: Amount Borrowed (in USD)						
Treatment	77.99***	5.28	0.51	18.16	25.04	103.09**
	(0.00)	(0.63)	(0.96)	(0.65)	(0.63)	(0.02)
Lag of the dependent variable		0.6	0.29**	0.10	0.35**	0.35**
		(0.18)	(0.04)	(0.22)	(0.03)	(0.01)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
R-squared	0.07	0.17	0.01	0.01	0.05	0.05
Endline control mean	0	25.38	81.90	27.47	134.8	134.8

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4,141. The table presents the coefficient of a "treatment" dummy in a regression of each outcome variable on the treatment dummy (Equation 1 in the text). **Figures in the parentheses are p-values obtained using wild cluster bootstrap-t procedure (Cameron et al., 2008).** P-values are adjusted for sub-district level clustering. There are 40 such clusters. The dependent variables in Columns 1-6 of Panel A are defined as follows: a dummy for whether the household had an outstanding loan from BCUP (Column 1), or from banks or co-operatives (2), or from Non-Government Organizations (NGOs) such as Grameen Bank, BRAC programs other than BCUP, and other NGOs (3), or from informal sources such as money lenders or other individuals such as family and friends (4), or if a household had a loan from any source other than BCUP (5), or if a household had a loan from any source including BCUP (6) . The dependent variables in Columns 1-6 of panel B are the amounts corresponding to the loans defined in the column headers. The Endline Control Mean reported at the bottom of each panel are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level. All figures expressing monetary values are in USD. Informal lender includes moneylenders, loans from friends/family, and buying goods/services on credit from seller.

Table F2: Impact of credit on Amount of Cultivated Land (in Decimal) Using Wild Cluster Bootstrap-t Procedure

	Rented-in Land				Own land	Total cultivated land
	Share-cropping	Fixed rental	Others	Total ((1)+(2)+(3))		Column (4) + Column (5)
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-2.01	6.58**	0.18	5.89	0.12	5.98
	(0.514)	(0.042)	(0.932)	(0.19)	(0.954)	(0.28)
Lag of the dependent variable	0.46***	0.68***	0.40***	0.64***	0.68***	0.67***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	4,141	4,141	4,141	4,141	4,141	4,141
Endline control mean	26.82	7.421	10.39	44.64	34.17	78.80

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4,141. The table presents the coefficient of a "treatment" dummy in a regression of each outcome variable on the treatment dummy (Equation 1 in the text). **Figures in the parentheses are p-values obtained using wild cluster bootstrap-t procedure (Cameron et al., 2008).** P-values are adjusted for sub-district level clustering. There are 40 such clusters. The dependent variables in Column 1 and 2 shows the amount of land rented-in under share-cropping arrangement (Column 1) and under fixed-rental arrangement (column 2). Column 3 shows total amount of rented-in land under any type of tenancy arrangement. The dependent variable in column 4 shows the amount of owned cultivated land. The dependent variable in column 5 shows the amount of total cultivated land, which is the summation of owned cultivated land and rented-in land. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. In all cases the unit of measurement of land size is in Decimals (1 decimal is equal to 1/100 acre). Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level

Table F3: Impact of Credit on Non-farm Self-Employment Activities Using Wild Cluster Bootstrap-t Procedure

	(1)	(2)	(3)	(4)	(5)
VARIABLES	HH participates in non-farm self-employment activities	Number of non-farm self-employment activities	Number of family labor	Number of hired labor	Current market price of all business assets (in USD)
Treatment	0.0634**	0.0704**	0.107**	0.0113	80.71*
	(0.02)	(0.03)	(0.03)	(0.69)	(0.08)
Lag of the dependent variable	0.475***	0.465***	0.399***	0.00796	0.0211
	(0.00)	(0.00)	(0.00)	(0.69)	(0.02)
Observations	4,141	4,141	4,141	4,141	4,141
R-squared	0.249	0.27	0.19	0.00	0.01
Endline control mean	0.174	0.189	0.215	0.0715	645.1

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4,141. The table presents the coefficient of a "treatment" dummy in a regression of each outcome variable on the treatment dummy (Equation 1 in the text). **Figures in the parentheses are p-values obtained using wild cluster bootstrap-t procedure (Cameron et al., 2008).** P-values are adjusted for sub-district level clustering. There are 40 such clusters. Observations with inconsistent amount of assets are dropped in column 5. Business outcomes are aggregated at the household level when the households have more than one business. The outcome variables are set to zero when the household does not have a business. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Table F4: Impact of Credit on Different Proxies of Child Labor (5-14 Years) in Self-Employment Activities Using Wild Cluster Bootstrap-t Procedure

	(1)	(2)	(3)
VARIABLES	HH employed child labor	HH did not use child labor in baseline, but used it in endline	Number of hothe worked by children
Treatment	0.07*	0.06*	0.34***
	(0.06)	(0.054)	(0.006)
Lag of the dependent variable	0.04***		0.01
	(0.02)		(0.11)
Observations	4,141	4,141	4,141
Endline control mean	0.0609	0.0459	0.203

Notes: Data from 2012 and 2014 surveys. The analysis is at household level. Total sample size is 4,141. The table presents the coefficient of a "treatment" dummy in a regression of each outcome variable on the treatment dummy (Equation 1 in the text). **Figures in the parentheses are p-values obtained using wild cluster bootstrap-t procedure (Cameron et al., 2008).** P-values are adjusted for sub-district level clustering. There are 40 such clusters. The dependent variable in column (1) is a dummy variable taking a value of 1 if household employs child labor and 0 otherwise. The Endline control mean are calculated for the control areas that were randomly assigned not to receive BCUP credit. Asterisks denote statistical significance at the 10(*), 5(**) or 1(***) % level.

Chapter 3: Access to Colleges, Human Capital, and Empowerment of Women*

Sheetal Sekhri

Md Amzad Hossain

Pooja Khosla

Abstract

Women's enrollment in colleges in developing countries is staggeringly low. Due to cultural norms and safety concerns, women may not be allowed to enroll. Alternatively, local colleges might be in short supply. Using the variation in college construction in India under a college construction grant policy, we assess whether increased access to colleges in home districts improves human capital outcomes for women. We find an improvement in educational and other development outcomes. The educational gains accruing to women are twice as large as men. The benefits accrue to the rural women but do not spillover to the neighboring districts.

JEL codes: O15, I25, J16

1 Introduction

College enrollment in developing countries, especially in South Asia and Sub-Saharan Africa, remains strikingly low despite many policy efforts. Data from UNESCO's Institute for

* Sheetal Sekhri is an Assistant Professor in the Department of Economics at the University of Virginia. Md Amzad Hossain is a Ph.D. candidate in the Department of Economics at the University of Virginia, and an Assistant Professor in the Department of Economics at the University of Dhaka. Pooja Khosla is a Graduate Student in the Department of Economics at the University of Virginia. The corresponding author is Sheetal Sekhri (ssekhri@virginia.edu). We are grateful to Professor Sachidanand Sinha at Jawaharlal Nehru University, New Delhi for helping us compute the gross enrollment ratios for the Indian districts for 2001 and sharing the methodology which formed the basis for district eligibility assessment for the government grant program. The paper has benefited from comments by Asim Khwaja. The usual caveats apply. IRB approval was not obtained for this study, which is limited to secondary analysis of existing data. We do not have any conflicts of interest. The data and the replication materials of this study can be found here. An additional Online Appendix with supplemental materials can be found at this link.

Statistics show that the *Gross Enrollment Ratio* (GER) in higher education in low-income countries stood at 10 percent in 2012.[1] More saliently, while in the developed countries more girls were enrolled in colleges than boys, there was a significant gender gap in the college enrollment rates in both South Asia and Sub-Saharan Africa – with 81 and 64 girls enrolled for 100 boys, respectively (Ilie and Rose, 2016). Supply-side constraints are often considered a major factor in giving rise to this pattern and, therefore, a popular policy prescription to combat this issue has been the construction of more local colleges. In developing countries, concerns about the safety of women and girls are particularly acute. These concerns and restrictive gender norms about limiting exposure of women significantly constrain women's mobility (Jayachandran, 2015; Muralidharan and Prakash, 2017; Dean and Jayachandran, 2019; Seki et al., 2020; Jayachandran, 2020). As a result, girls traditionally do not attend educational institutions that are too far (Sutton, 1998). Relaxing these supply constraints by constructing more colleges in the educationally lagging regions, for example, might benefit girls and help narrow these gender gaps.

A different school of thought identifies lower demand for education, and misplaced expectations and prejudices as the main impediments to increasing girls' access to higher education (see for example, Jones (2008)). Children in developing countries often have to perform household chores and/or engage in the farm enterprise. Since these household or economic activities are different for boys and girls, the opportunity cost of attending colleges can be higher for girls (Sutton, 1998). There is also a perception that the returns to higher education for women are low: young age has a premium in the marriage market[2], labor market opportunities are worse for women[3], and girls typically move to the husband's house after they get married which may not allow natal families to benefit from investments in girls.[4] These enduring perceptions can preclude girls from attending colleges despite an adequate supply of colleges.[5]

Our study informs this debate by evaluating the consequences of a massive college expansion policy in India. With the objective of promoting access to higher education and making it more inclusive and equitable, the University Grant Commission (UGC) of India (the apex body that regulates higher education) introduced an innovative program in 2008 to provide financial assistance by way of grants for constructing new model colleges in the 374 'Educationally Backward Districts' (EBD) of India. These were identified as the districts where the GER in higher education was lower than the national average in 2001. We explore the variation in college expansion generated by this program to isolate the causal impact of college access on the human capital. We also examine returns in the labor market, marriage market outcomes such as husband quality, and other development outcomes such as the age at marriage for women.

In Appendix Figure A.1 Panel (a), we plot the decadal increase in total number of colleges between 1950-51 and 2010-11. Until 2001, there was a steady and slow increase in college growth. By contrast, college growth accelerated in the decade between 2001 to 2011. In Panel (b), we plot the total number of colleges by year since 1983. The data reveals a sharp jump in the total number of colleges in 2009-10. Our identification strategy harnesses this temporal trend change. Analogous to this, in Appendix Figure A.2, we observe a steady increase in college enrollment until 2001. From the aggregate statistics that we plot and examine, two facts about higher education in India in the last three decades stand out. First, while the liberalization of the Indian economy in 1991 resulted in an uptick in college construction and enrollment in the 1990s, a notable precipitous change in enrollment occurred during the period 2001-2011. Second, women's enrollment in colleges was dismally low despite increases in overall enrollment before 2001. However, the decade between 2001 and 2011 saw an escalation in women's enrollment. In this paper, we first try to establish a causal increase in women's college enrollment due to an increase in the number of local colleges post 2008 and then attempt to highlight its effects.

The key identification challenge while estimating the impact of college access is that college placement is non-random. To address this issue, we exploit the fact that exposure to college expansion was different across regions and birth cohorts. Districts with GER lower than the national average in 2001 were eligible for the college construction grants. We utilize this spatial variation and the policy-induced temporal variation for identification. Average district GER, and not gender-specific GER, was used to ascertain eligibility. We identify the eligible districts by computing the 2001 district-level GER using the Census of India 2001. We employ the same algorithm that formed the basis for the government program. The eligible districts are then considered treated. We rely on a trend-break model for identification.[6] To identify the causal impact of the college expansion program, we compare the outcomes in the 374 eligible districts (treatment) with those in the 266 districts not eligible for the grants (control), before and after the expansion. We use a long district-by-year panel (from 1985 to 2017) for estimating the changes in college construction. We then employ cohort-analysis for estimating the overall and post-policy trends in outcomes by treatment status. We use birth years 1965-1998 (ages 18-51) for our analysis. Our empirical strategy controls for district and year fixed effects that absorb any confounding time-invariant district characteristics and year-specific shocks common across all districts, respectively. We also account for an overall secular trend (albeit negative) in the treated districts. Our identification rests on estimating post-policy trend changes from this overall trend in treated districts relative to the controls. We identify both slope and intercept changes.

Two stylized facts inform our choice of the identification strategy. First, because there are only 640 districts in India (less so in 2001) and most of the expansion happened in districts with very low GER to begin with (in the tails of the GER distribution as clearly observed in Appendix Figure A.3, rather than close to the grant eligibility cutoff), we harness the panel dimension of the data for identification instead of relying on a regression discontinuity design exclusively even though a sharp threshold was used for determining eligibility. Second, the districts with low GER (EBDs) were not necessarily similar to the high GER districts. The low GER districts exhibited a striking negative trend in the education outcomes prior to the policy which was followed by a precipitous increase in college construction after the policy change. This makes the use of a differences-in-differences strategy less suitable. We present the results from these alternative methods and demonstrate that these are not viable.

Our reduced-form analysis first establishes that there was a change in college expansion in treated areas right after the policy was implemented. College construction was less rapid in the treated areas before the policy but this trend reversed immediately after. Having established that the policy indeed increased the supply of colleges, we turn to examining the consequences for human capital using a panel of cohorts born between 1965 and 1998 using

National Family Health Survey (NFHS-4), a nationally representative survey, conducted in 2015-2016. We find that years of education, college enrollment, and college completion for girls exhibit similar trends in the treated areas: much slower growth prior to the policy and a trend reversal with an acceleration in the post-policy period. Our estimates indicate a 11 percent gain in years of schooling over the baseline control mean. Men too benefited from the program but the benefit to women is substantially larger. Thus, the program was successful in narrowing the gender gap in higher education. Our heterogeneity analysis reveals that the education benefits accrued to women in rural areas, albeit from wealthy families.

In addition to changes in college education, we also find evidence for other improvements in the welfare of girls. Expansion lowered the prevalence of child marriage among girls. In the treated areas, we observe a decline in fertility, albeit these effects are not significant. Labor market outcomes for married women improved significantly. Men's labor force participation, on the other hand, was not impacted. We do detect suggestive evidence of changes in spousal quality of women in the marriage market. On examining the characteristics of the spouses in the post-period with the NFHS data, we find an increase in the likelihood of employment of the husband driven by non-agricultural work.[7] Our findings reveal that better college access can increase college participation and completion among women in developing countries and improve their labor market outcomes. This occurs concurrently with a delay in marriage age, which not only affects the welfare of women (Field and Ambrus, 2008; Jensen and Thornton, 2003), but that of their children as well (Sekhri and Debnath, 2014).

In a series of robustness checks, we bolster our estimation by showing that our results are not sensitive to trends in economic growth (proxied by night time luminosity) and poverty. Structural time-series trend-break analysis indicates a break in trend in college construction and educational outcomes in 2009. We also rule out political changes or migration as alternative explanations. We find evidence of heterogeneity in college education outcomes by place of residence – benefits mostly accrue to rural residents. We do not find evidence of spatial spillovers. Women in adjacent neighboring districts that did not receive a new college did not see an improvement in educational outcomes.

Our paper complements and extends three strands of literature. First, it relates to a growing literature examining the effect of educational infrastructure investments in developing countries on education outcomes of girls. Some of the existing studies in this vein have focused on newly constructed schools. For example, Burde and Linden (2013) examine the effects of local village-schools on girls' educational outcomes in Afghanistan and find very large effects on enrollment and performance. Kazianga et al. (2013) look at primary school construction with complementary girl-friendly interventions and find similar results. Kim et al. (1999) and Barrera-Osorio et al. (2011) evaluate the effect of granting larger subsidies to girls in the case of newly constructed private schools. Other papers have examined improvements in already existing elementary schools such as toilet construction, electrification, and adding girls-only sections (Adukia, 2017; Meller and Litschig, 2016) targeted towards improving enrollment of girls. We extend and complement this literature in two ways. First, we innovate by studying gender-based benefits of college expansion in developing countries where college access for women is limited and document that relaxing the supply constraints increases college participation and graduation of girls. Prior literature has only focused on schools. Second, unlike these papers, we go beyond the immediate educational benefits of colleges, and provide evidence for improvements in labor and marriage market outcomes.[8]

The second strand of papers debate whether social norms and traditional institutions impede women and girls from reaping the benefits of expanding educational opportunities by way of improved infrastructure or global economic integration. Evidence is mixed. While Cheema et al. (2018) and Jacoby and Mansuri (2011) indicate a *boundary effect* precluding women from utilizing improved educational facilities, Munshi and Rosenzweig (2006) show that traditional institutions disadvantage boys whereas girls gain from higher economic returns to education resulting from economic integration. Others like Alderman et al. (2001) find that longer distances lower the probability of parents sending girls to private school by more vis-a`-vis boys. Ashraf et al. (2020) implicate cultural practices as an important determinant of whether primary-school construction benefits girls or not. Our paper contributes to this literature by exploring the geographic spillovers of college construction. We find that benefits of colleges do not cross district boundaries.

The third strand of literature we complement relates to the geographic proximity of colleges and returns to educational attainment. A number of papers focusing on the United States (US) have used the presence (or absence) of a local college or distance from college as instruments to estimate the returns to education (Card, 1993; Kane and Rouse, 1995; Carneiro et al., 2011; Carneiro and Heckman, 2002; Doyle and Skinner, 2016). The majority of these papers have focused on the rapid higher education expansion phase in the US during the 1960s and 1970s. The idea behind this approach is that proximity to college improves access and reduces costs as students can commute as opposed to being in residence at college (Kerr, 1991).[9] We contribute to this literature in a variety of ways. First, we evaluate the benefits of college proximity in a developing country. There are no prior estimates of the benefits of college proximity for developing countries where supply constraints are still binding. Second, unlike most of the previous works, our focus is on access to higher education for women.[10] In our study, we examine a variety of social outcomes for women in addition to educational outcomes and labor market outcomes. We also shed light on benefits in the marriage market.[11]

Finally, our paper also contributes to the debate concerning investment in higher education in developing countries. International donor agencies like the World Bank have, in the past, considered investment in higher education as equity-detracting and have espoused allocation of scarce government resources to primary education as it provides higher social return than higher education (Psacharopoulos and Patrinos, 2004; World Bank, 2000). We find that constructing colleges not only increases college enrollment and improves labor market outcomes of women but also reduces child marriage and fertility. To the extent that the confluence of these factors improves the welfare of women, college expansion can have gender-based distributional effects and can aid in reducing gender disparities in developing countries.

The rest of the paper is organized as follows. Section II provides the cultural context and background information on establishing the model colleges in the educationally backward districts in India. Section III describes the data. We discuss the empirical methodology to identify the causal impact of the college construction program in section IV. Section V reports the empirical results and the falsification tests, and section VI concludes.

2 Background

While India's knowledge economy is fueled by its massive college-educated workforce, college education still remains elusive to its masses. In 2016-17, with a higher education GER of only 25.2 percent, India trailed behind Asian countries like China which had a GER of 43.39 percent. By way of comparison, the GER for the US was as high as 85.8 percent.[12] In addition to lagging behind many other economies, India faces another challenge of uneven access across states. In the past six decades, there has been a visible and substantial growth in colleges in India but it has not been spatially uniform. While some states have seen massive increases, others have not kept pace. College density, defined as the number of colleges for every one hundred thousand people in the age group of 18-23 years, varied from 7 in the state of Bihar to 59 in Telangana compared to an all-India average of 28 (according to the All-India Survey on Higher Education, 2015-16). In 2016-17, the GER in higher education ranged from 14.9 percent in Bihar to 46.9 percent in Tamil Nadu.

Indian policymakers opine that proximity to colleges can go a long way in addressing this disparity. The 11th five-year plan of India emphasized increasing access to higher education in educationally backward districts. According to the 2001 Census of India, 374 districts had a gross enrollment ratio in higher education lower than the national average of 12.4 percent. The apex higher education regulating body, the University Grants Commission (UGC) of India, along with the Ministry of Human Resource Development (MHRD) initiated a program to provide financial assistance for establishing new model colleges in these 374 educationally backward districts in 2008.[13]

To receive financial assistance under this program, the college had to be set up on, or after, 1st January 2008 in any of these 374 eligible districts. While setting up model colleges, preferences were given to the following areas: (a) hilly or border regions populated with a higher official share of minority and tribal population, (b) areas with no college within a radius of 10 kilometers, and (c) rural areas with reasonably good transportation facilities.

The model colleges established under this scheme had to be either a constituent unit or permanently/temporally affiliated with a university covered under Section 12B of the UGC Act.

Colleges established under this program received grants from the state as well as the central government/government-funded bodies. The state government had to provide land to establish the model college. Two-thirds of the non-recurring and entire recurring expenditures were also to be funded by the state government with a provision for further appreciation in capital expenditure in the future. The affiliating universities were responsible for ensuring that the funds provided to the colleges were being utilized for infrastructural development. While establishing the model colleges was a collaborative project between the UGC and the state governments, the latter were allowed to collaborate with non-profit bodies or to enter into partnership arrangements with a for-profit organization under a *public-private partnership* arrangement clause.

In 2013, this program was subsumed under the new centrally sponsored 'Rashtriya Uchchatar Shiksha Abhiyan' (RUSA). The primary objective of RUSA was to improve access, equity, and quality in higher education through the planned development of higher education at the state level. The perception was that this would be an instrument to harmonize the national program for funding state universities and colleges through a single over-arching umbrella body. RUSA's key actions and funding areas included the construction of new model colleges (general and professional) as before. In addition, the policy was also geared towards upgradation of existing autonomous colleges to universities, conversion of colleges into cluster universities and provision of infrastructure grants to colleges and universities.[14]

3 Data

We use several datasets to conduct our empirical analysis. There are two main sources of data pertaining to college expansion and eligibility, educational and marriage market outcomes, and employment. To these sources, we add the demographic data from the Census of India and poverty and night light (average luminosity) data to control for covariates in our empirical specifications. Below we describe these data in detail.

3.1 College Expansion and Eligibility

Data on the year of establishment for all colleges comes from the All-India Survey on Higher Education (AISHE). This is an annual web-based survey conducted by the Ministry of Human Resource Development for all institutions engaged in higher education and is available from 2011 onwards. This survey collects information about the year of establishment, programs offered, teachers, infrastructure, and so on. We use 2018 AISHE data for our analysis, which contains an exhaustive list of all the colleges of India, their location, and their year of establishment.[15] We construct a panel dataset of the number of new colleges (net of closures) constructed in each district by year between 1985 to 2017. The data is self-reported. [16]

Information about the treatment districts was obtained from the University Grant Commission of India. Out of the 593 districts in the country (as per the 2001 census), 374 districts with GER lower than 12.4 percent were identified as educationally backward districts. These districts were eligible for the grant. For our analysis, we consider these districts as treatment districts and the remaining as controls. Hence, our estimation will yield Intent-to-Treat (ITT) estimates. In order to compute the 2001 GER for all the districts, we rely on the methodology used by the government to establish eligibility. We use Table C-10 from 'C–Series: Social and Cultural Tables' of the 2001 Census of India to obtain district-wise gross enrollment in higher education. This is the total population enrolled in colleges in India. We combine this with district-wise population in the 18-23 age group. The latter is imputed using the age- and district-wise population data from the 2001 Census. This data provides population information for ages 18, 19 and 20-24. We assume a uniform distribution of the population in the 20-24 age group to get an approximation for the population in the 20-23 age group and combine it with that for ages 18 and 19.[17]

3.2 Educational and Other Outcomes

We use household data from the fourth round of the National Family and Health Survey (NFHS-4), conducted in 2015-16. By this time, Indian districts had expanded to 640 in total. The survey covers all the 640 districts of India spanning the 29 states and six union territories. NFHS has four broad modules: (a) household survey, (b) woman's survey, (c) man's survey, and (d) bio-markers. We rely on the woman's survey for most of our analysis. The woman's module collected data on 699,686 women from about 568,200 households on various demographic and socio-economic characteristics such as education, marital status, age of marriage, autonomy in decision making, and domestic violence. It also covered a variety of health-related outcomes, such as fertility, knowledge and use of contraceptives, and child health. In addition, for a subset of women, the survey also collected information on the respondent's employment and, if married, their husband's education and employment.

Since our identification utilizes spatial variation in exposure to college construction, and the NFHS does not identify the districts in which females resided at different ages, we limit our analysis to individuals who have never moved. Thus, our primary sample for the educational attainment analysis consists of 141,641 women. Among these, 75.54 percent of the women are unmarried. The NFHS administered employment and spouse characteristic questions to only a subsample of women. So, our sample size for women's employment outcomes is 24,378 and that for husbands' education and employment outcomes for married women is 6,158. Information on wages or earnings is not collected by the NFHS.

3.3 Other Ancillary Data

District-level economic and demographic data (for example, literacy rate, unemployment rate, gender ratio, percentage of population residing in rural areas) come from the 2011 Census of India. We use 2011-12 poverty estimates by Mohanty et al. (2016), who calculated these by pooling the 66th and 68th rounds of the consumption expenditure surveys carried out by the National Sample Survey Office (NSSO). Night light data comes from the National Geophysical Data Center (NGDC) at the National Oceanic and Atmospheric Administration (NOAA).[18] Finally, we use two rounds of Indian Human Development Survey (IHDS) to see the effect of the grants policy on in-migration and out-migration.

3.4 Summary Statistics

We report the summary statistics of the outcome variables for the married and unmarried women in our sample (Panel A), married women's husbands (Panel B) and, men (Panel C) in Table 1. The summary statistics highlight the poor socio-economic condition of women. The average years of schooling are less than nine years, with only 15 percent of the women enrolling in colleges. Women's participation in the labor force is very low, at only 27 percent.

Panel A of Appendix Table A.1 compares women in the treated and control districts with respect to several outcome variables for the cohorts born before 1990 (these cohorts were already 19 years or older at the time of the policy reform and, hence, unlikely to be affected by the reform) in our sample. We can see that women in treated districts have worse educational outcomes and a higher likelihood of child marriage. Trailing educational attainment is also reflected in the 2001 population census (Panel B of Appendix Table A.1).

4 Empirical Strategy

The main identification challenge for our analysis is that colleges are not randomly located across regions. Local economic or political conditions, for example, can influence college investment as well as the human capital and status of women. In order to address this endogeneity issue, we rely on the policy-generated sharp variation in college expansion across India. As mentioned before, the government's 'model college' policy provided college construction grants. This college expansion subsidy program commenced in 2008 and was targeted based on GER in higher education in 2001 in the various districts. The government used the national average GER in 2001 as a threshold for allocating the subsidy. Districts with GER below this threshold were eligible for these grants. There was a total of 374 such districts. The program was based on the overall GER, and not the gender-specific GER. Appendix Figure A.4 highlights the gap in the total number of colleges between eligible and ineligible district over time. Starting from roughly 1,100 fewer colleges in 1985 in the control districts, the gap widens until 2009 and reaches over 3,100 fewer colleges relative to the treatment districts. We observe this gap narrowing after the policy change. Appendix Figure A.5 shows the map of the eligible districts to highlight the spatial variation. We utilize these two sources of variation (temporal and spatial) to identify the causal estimates of access to colleges on development outcomes of women.[19] We compare the 374 eligible districts to the remaining 266 districts of India before and after this policy change.

We define the variable *Treatment* as[20]

$$Treatment = \begin{cases} 1, & if\ I(GER < National\ Average_{2001}) = 1 \\ 0, & otherwise \end{cases}$$

4.1 Event Study

In order to estimate the effects of the grant policy, we begin by first estimating the following event-study equation:

$$(1) \qquad C_{it} = \alpha_i + \beta_t + \sum_{l=1985}^{2017} \gamma_l \times (\delta_l \times Treatment_i) + \varepsilon_{it}$$

where C_{it} is the number of new colleges constructed in district i in year t. The district fixed effects, α_i, control for time-invariant unobserved determinants of college construction. The inclusion of time fixed effects β_t control for yearly shocks common to all districts and allow us to non-parametrically adjust for overall trends in college construction. δ_l represents the year indicators and γ_l are the estimated coefficients.

Figure 1 plots the estimated γ_l coefficients from Equation 1 (blue circles in solid line).[21] Recall that treated districts were the educationally backward districts with lower GER in 2001 and we expect these to have lower college access. The event study graph bears out our conjecture about college expansion in India. It seems quite evident from the figure that the treated districts were on a differential trend relative to the control districts as far as college construction was concerned. Treated districts experienced a slower college expansion before 2009. However, in 2009, we observe a sharp trend reversal, which coincides with the year when the policy became effective. While the event study clearly shows that after the policy change, the low GER districts experienced a much more rapid college expansion, we also confirm a statistically significant and negative pre-trend. This violates the identifying assumption for differences-in-differences (DID) approach which requires parallel pre-trends in order to yield unbiased estimates. This motivates the use of a trend-break approach which allows for selection on observables and, relative to the DID approach, requires a less stringent identifying assumption. Differential pre-trends are not a source of bias, though attribution requires that no other variables exhibit similar trends and a break at the time of the policy change.

4.2 Trend-Break Estimation

The empirical model is as follows:

$$(2)\qquad C_{it} = \alpha_i + \beta_t + \gamma_1 \times Treatment_i \times [t - 1985] + \gamma_2 \times Treatment_i \times [t - 2008] + \gamma_3 \times Treatment_i \times Post + \varepsilon_{it}$$

where, $[t - 1985]$ and $[t - 2008]$ are linear time trends denoting the overall trend for the entire sample period and post-trend after the policy change, respectively. We interact both trends with the treatment indicator. γ_1 and γ_2 are the respective coefficients.[22] *Post* is a dummy variable taking a value of 1 for the years after the policy change (2009 and after) and 0 otherwise. This variable enters the regression interacted with treatment status. The coefficient on this interaction, γ_3, captures the intercept change post-policy. We cluster the standard error at the district level to account for the possible serial correlation in errors. For college expansion, we use a panel spanning 1985-2017. We conduct sensitivity checks to verify that the results are not sensitive to the period chosen.[23]

We report the results of the trend-break model estimating Equation 2 in Column 1 of Table 2.[24] We can see that treated districts get 0.05 fewer colleges each year before the policy change. But annual college construction reversed the trend in 2009 and the post-trend became 0.2 more colleges in the treated districts each year relative to the overall trend. The post-trend is positive and significant implying that we can reject $\gamma_2 = 0$ in Equation 2. We also report the F-statistics for testing $\gamma_1 + \gamma_2 = 0$ (F-stat 1) and the joint significance of γ_1 and γ_2 (F-stat 2). These tests are indicative of a statistically significant trend-break. The implied year-specific treatment effects from this trend-break estimation are depicted by the maroon squares in the solid line in Figure 1.

To put these estimates in perspective, in 1985, the control districts on average had 14.5 colleges compared to an average of 7.3 colleges in the treatment district – a gap of 7.2 colleges. Our estimates from Table 2 suggest that this gap widened by 0.05 colleges each year prior to the policy. Over the period 1985 - 2009, this translates to approximately 1.2 fewer colleges constructed in the eligible districts on average relative to the ineligible district – widening the baseline gaps by roughly 13 percent. After the policy, however, this trend reversed with 0.192 more colleges constructed in the treated districts relative to the untreated districts each year. Over the period 2009-2017, this translates to roughly 1.54 more colleges constructed in the average treated districts. This implies that the policy was successful in generating a more rapid expansion of colleges in the treated areas thereby decelerating the gap in the two areas in the number of total colleges. In the absence of the policy, the baseline gaps would have substantially widened.

4.2.1 Validating the Structural Break: Supremum Wald Test and Using Continuous GER Measure in the Trend-break Models

To bolster our analysis, we also use structural trend-break methods in time-series: (i) to test whether the year-specific coefficients exhibit a structural break in 2008, and (ii) without imposing the year of break, estimate the year in which these coefficients exhibit the structural break. To this end, we perform a supremum Wald test.[25] The results indicate strong support for our identification design: (1) we reject the null hypothesis of no structural break in 2008 at the 1 percent significance level, and (2) the data reveals a structural break in the year 2008, precisely when the policy came into effect.

To show that the trend reversal observed in Table 2 is driven by the policy change, we augment the empirical analysis by conducting the trend-break estimation as a function of the 2001 district GER (instead of the treatment indicator). Using the base year 1985, Appendix Table A.2 reassuringly reveals that the districts with higher GER in 2001 had a higher college expansion trajectory relative to the educationally backward districts until2008. This trend reversed in 2009.[26]

4.3 Evidence from RDD

Though the use of a sharp threshold to identify eligible districts warrants using a regression discontinuity design, there are four reasons why we do not rely on it as our main identification strategy. One, there is a relatively small number of districts in our data, so we are not adequately powered for this analysis (640 districts). Second, preferential treatment was accorded to some areas (described in the background section) which undermines the use of this approach. Third, the educationally backward districts (those with GER below the national average) exhibited a statistically significant differential trend relative to the other districts prior to 2008. Finally, the expansion after the policy happened in the tails of the GER distribution. That is, in districts with very low GER.

In Appendix Figure A.3, we plot the growth rate of colleges between 2008 and 2018 (number of new colleges constructed in a district as a share of the respective district's total number of colleges in 2008) on either side of the GER cutoff of 12.4. We observe that the plot for the grant eligible districts is steeper than the control districts and is downward sloping, indicating that the highest growth of colleges occurred in treated districts with the lowest GERs, that is, in the tails of the distribution rather than in the neighborhood of the 12.4 threshold. We do notice a small jump at the threshold. In order to test for significance and sensitivity, we carry out an RDD analysis using the optimal bandwidth with varied polynomial control functions for each year between 2008-2018 for the cumulative number of colleges in the district in the specific year normalized by the total population in lakhs (that is, colleges per 100,000 individuals). Our empirical analysis verifies that this estimate is modest and not statistically significant. The results are summarized in Appendix Table A.3.

5 Main Reduced Form Results and Discussion

Having established that the policy led to a rapid expansion of colleges in the districts with low GER in 2001, we now turn to examining the effects of this policy driven college expansion on education and other development outcomes. We take the reduced form approach and estimate a trend-break model substantiating it with evidence from an event-study.

5.1 Educational Outcomes

Our reduced-form analysis indicated that there was a sharp trend reversal in college expansion after 2009. Districts with low GER had a sharper increase in college expansion due to the college grants. We now turn to evaluate how this affected educational and other development outcomes for women. To this end, we compare the cohorts of girls of college-going age before and after the policy reform in a similar trend-break model. Our sample period includes birth years from 1965 to 1998. Individuals born between 1991 and 1998 were 18 years or younger at the time of the policy change. Since our survey data is from 2016, we are able to use eight cohorts that are younger than the college eligible age of 18. The youngest cohort is 11 years old in 2009 and 18 years old in 2016.[27] We rely on reduced-form regressions of the following form:

$$(3)\quad \begin{aligned} Y_{whit} = \alpha_i + \beta_t + \lambda_1 \times Treatment_i \times [t-1965] + \lambda_2 \times Treatment_i \\ \times [t-1990] + \lambda_3 \times Treatment_i \times Post + \theta \\ \times X_{hit} + \varepsilon_{it} \end{aligned}$$

We index an individual girl by *w*, household by *h*, district by *i*, and the individual's birth year by *t*. *Y* is the outcome variable of interest. *Post* is a dummy variable taking a value of 1 if the individual is born in 1991 or later, and 0 otherwise. Analogous to Equation 2, [t−1965] and [t−1990] capture the overall and post-policy trends, respectively. X_{hit} denotes a vector of household characteristics such as household assets, number of family members, rural/urban dummy, and caste. We cluster the standard errors at the district level.

5.1.1 Educational Outcomes for Women

Table 3 reports the results of the trend-break analysis estimating Equation 3, along with the standard errors. Panel A and Panel B present the results with and without the socioeconomic controls, respectively.[28] The number of years of schooling (Column 1), college enrollment (Column 4), college completion (Column 5), and number of years in college (column 6) reflect an analogous change in trend as college expansion. Prior to the policy change, these variables were trending negatively for the educationally backward districts (treatment districts). However, after the policy change, we observe a trend reversal and the F-tests reported at the bottom of the table show that the trend reversal is indeed significant.

Similar to Figure 1, we present the results of the estimation graphically in Figure 2 using the specification with controls. As before, the blue circles in solid line depict the estimates from an analogous event-study analysis and the maroon squares in the solid line reflect the implied year-specific treatment effects from the trend-break estimation. For both the number of years of schooling [Panel (a)] and college enrollment [Panel (b))], we see a remarkably similar pattern: a negative trend for cohorts born in or before 1990 (19 years or older in 2009, and thus not likely to benefit from the grant program), followed by a sharp trend reversal for cohorts born after 1990.[29]

Our estimates from Panel A reveal that girls in the treated cohorts experienced an annual increase of 0.103 years of schooling after the policy, relative to the overall trend. In terms of magnitude, this is equivalent to an annual gain of 0.023 standard deviations. Our results translate to an average increase of 0.82 years of schooling or approximately 11 percent increase over the baseline control mean. These estimates are much higher compared to those documented for the US. For instance, Doyle and Skinner (2016) find an average increase of 0.56 years of education for individuals with a four-year college in their county at age 17. Similarly, Card (1993) finds positive effects on education attainment of 0.32 to 0.38 years for those growing up near a college. The increase in years of schooling is driven by both increases in college enrollment and an increased enrollment in schools. In Appendix Table A.4, we show that there is an overall annual increase of 0.103 in years of schooling (column 1), 0.038 years of increase in schooling for individuals who did not attend college, and 0.067 years increase in college.[30] This increase of 0.067 translates to an increase of 0.53 years in college which is comparable to the US as documented by Doyle and Skinner (2016). College enrollment and completion increased annually by 2.1 and 1.7 percentage points, respectively, for the treated cohorts after the policy change, relative to the overall trend, whereas the number of years in college increased annually by 0.066 years for the treated cohorts.[31] Overall, the local college construction policy led to economically large and significant improvements in girls' educational outcomes.

As mentioned in Section III.B, our primary sample includes women who have always stayed in the same place. We did this to ensure we know the women's region of birth and to abstain from pooling migrants in the sample. That said, we also report the trend-break estimates for female's educational outcomes for the full sample (i.e., all females regardless of how long they have lived in current place of residence). Table 4 shows that college enrollment and completion increased annually by 1.42 and 1.44 percentage points, respectively, for the treated cohorts after the policy change, relative to the overall trend. The number of years in college increased annually by 0.052 years for the treated cohorts. These estimates are statistically significant albeit lower than the estimates obtained from the restricted sample, i.e., the non-mover sample. This is largely driven by a smaller effect on those who moved.

In order to discuss the plausibility of our estimates, we discuss the increase in capacity generated by the local colleges and the increase in enrollment in colleges in Appendix A.1 and A.2 and distil the discussion in Appendix Table A.5. The main take-away is that the approximated increase in enrollment is less than the approximated increase in capacity.

5.1.2 Educational Outcomes for Men

We report the results of the trend-break analysis for men in Table 5. As before, Panel A and Panel B present the results with and without the socio-economic controls, respectively. From Panel A, we can see that, prior to the policy, boys in treated districts attained 0.021 more years of education each year relative to boys in the control districts. Results from both the panels, with and without controls, are similar. After the policy, educational outcomes for boys in the treated districts improved further. There was a statistically significant annual increase of 0.06 in years of education and a 0.6 percentage points annual increase in the probability of enrolling in college in the post-policy period. Even though men benefit from this policy-induced college expansion, benefits to women are much larger in magnitude. In treated areas, women caught up with men in terms of years of education and their likelihood of enrolling in college outpaced men statistically significantly. These results are in line with existing evidence on the effects of gender-neutral K-12 school infrastructure improvements. Such interventions benefit both boys and girls but gains for girls are disproportionately larger. Analogous to women, we also estimate the effect using the full sample of men. The results are summarized in Appendix Table A.6.

5.2 Trends in Other Potential Confounders and Other Robustness Tests

5.2.1 Economic Growth, Poverty, and State Policies

To allay concerns about other potential confounding variables exhibiting the same trends, and thus leading to a spurious relationship between college expansion and changes in human capital, we show that our results are robust to a variety of controls. In Table 2, we control for local economic conditions proxied by trends in average night-time luminosity. The concern is that the economic growth trajectory might have reversed direction for an alternative reason, which causes the observed trend-reversal in college expansion. However, Column 2 of Table 2 reveals that the results remain unchanged. Related to this, we also control for trends in poverty. We use different measures of poverty for 2011 – poverty gap (Column 3), squared poverty gap (Column 4), and headcount ratio (Column 5) – all interacted with the time trend. The trend reversal remains significant regardless of the poverty measure used. Finally, state education policies can alter the college expansion in different ways. To rule out these as drivers of our results, we include state-specific time trends in the specification and get consistent results (Column 6).

Analogously, we show that these variables do not affect the estimates for the number of years of schooling and college enrollment. We document the results in Appendix Table A.7 (Panels A and B). We include the trends in night lights (average luminosity as a proxy for GDP) in Column 2 and the various measures of poverty (poverty gap, squared poverty gap, and headcount ratio, respectively) in Columns 3-5. The results remain unchanged.

5.2.2 Sensitivity to Sample Period

As a robustness test, we test whether our results are sensitive to different sample periods. We report the results in Appendix Table A.8. Our estimates remain statistically significant for a range of sample periods: 1981-2017, 1985-2017, 1990-2017, and 1995-2017.

5.2.3 Political Outcomes

Political changes could also affect outcomes, which are then attributed to college access via expansion. Since there was an election for the national government in 2009 and the political representation could have changed around this period, we investigate if there is a change in the local representation affiliated with the party forming the federal government. Such affiliation can bring in more resources (Asher and Novosad, 2017) and affect local labor markets. In so far as night lights absorb the local growth trends, our results are robust to including these (as shown in Appendix Table A.7).

We also directly assess whether there is a change in the local representation that can potentially affect resource allocation. In India, the local state assemblies nest within districts. In the 2009 elections, a coalition of parties formed the national government as no single party secured the majority. Hence, we compute the proportion of local state assembly elected representatives from each district that align with either the alliance forming the national government or the major party in the national government, before and after the policy reform, using the state assemblies' political representation and the national government's political representation data from 2005 to 2017.[32] We report the results from a trend-break model in Appendix Table A.9. One caveat is that we have political outcome variables starting in 2004. We do not see a trend-break in the post-policy period.

We also estimate a differences-in-differences specification and summarize the results in the first two columns of Appendix Table A.10. Conditional on district fixed effects, the proportion of local elected representatives affiliated with the national government did not change in the treated areas post-policy. The results are robust to including district-specific linear trend (Column 1) or state-specific linear trend (Column 2).

5.2.4 Migration

Migration for education is negligible among women in India though marital migration is substantial. Women do move to their husband's place of residence after getting married. It is possible that the policy change induces some women to attend college after being married into a treated district or women choose a spouse endogenously in a treated district to benefit from the newly founded colleges. If this were the case, our estimates would be attenuated due to out-migration from the control districts to the treated districts. In such a case, we would be estimating a lower bound of the treatment effect. That said, very few women in India continue to study in colleges after getting married. In the NFHS, only 0.9 percent females married before 18 years of age enroll in college and typically 18 is the age of entry into college. Among those married before 21 years, which would be the college completion age, 0.5 percent are enrolled in colleges. While we do not have data on the year of college enrollment, completion or the district where college was completed to allay this concern more concretely, we conduct a number of tests in order to rule out such attenuation bias. [33]

First, we consider in-migration. Since we do not have annual data on in-migration, we are unable to estimate a trend-break model. Instead, we rely on two waves of a household survey panel (India Human Development Surveys) from 2005 and 2012 to carry out this analysis. We investigate whether the probability of having migrated to the current residence in the last five years changes in the treated districts after the policy reform. We include household fixed effects in our specification reported in Column 3 of Appendix Table A.10 as we observe the same households in 2005 and 2012. The probability of in-migration in the last five years is negligible and statistically insignificant. Column 2, using district fixed effects, also shows the same results. Hence, there is no evidence of in-migration in the treated districts around the time of the policy that could be biasing our results.[34]

Next, we consider out-migration. For this analysis, we used the 2012 IHDS tracking sheets. The sheets track individual members of each household from the 2005 round who have moved to another household or migrated for any reason including education and employment between 2005 and 2012.[35] Using this data, we apply a trend-break to the probability of female migration for education or employment (our outcome variables of interest). The trend-break results are reported in Appendix Table A.11. Reassuringly, we do not detect female out-migration for education or work. We also report the event study estimates in Appendix Figure A.6 and find no effect of the grant policy on out-migration for education or employment. These results suggest that out-migration of females due to education is not confounding our main findings.

5.3 Heterogeneity by Rural Status and Wealth

5.3.1 Rural and Urban Status

We explore the impact of the grants program on the educational outcomes of girls across rural and urban locations using our sample described in the data section. Table 6 suggests that the gains from the program with respect to the number of years of schooling are more than three times as large in rural areas than the urban areas – an annual increase of 0.125 in rural areas compared to an annual increase of 0.037 in urban areas. The Chow test for statistical equivalence of the two estimates indicates that the difference is statistically significant. We also observe a larger trend reversal in the rural areas compared to urban areas with respect to college enrollment and college completion.[36] In the sample, 77.51 percent of the women in treated areas reside in the rural areas. In the absence of well-developed transportation infrastructure, distance to colleges is a much larger constraint for women in rural areas. Hence, availability of a local college option in the district affects them to a much larger extent.

5.3.2 Wealth

Next, we look at the heterogeneity in the effect of the program on college enrollment across different wealth quintiles of the households.[37] Results using the sample described in the data section are reported in Table 7. The effect on college enrollment does vary across wealth groups: the college enrollment of girls monotonically increased with the wealth level of the households except for the richest quintile. Women from richer households seem to be the primary beneficiaries of the program. While this program does benefit women, the poorest women are still not able to access colleges. This might be indicative of binding financial constraints and higher opportunity cost of time.

5.4 Spillovers

In South Asia, women are precluded from crossing village and district boundaries to access education or skill-development opportunities (Cheema et al., 2018; Jacoby and Mansuri, 2011). Enduring social norms are the main drivers of this *boundary effect*. With rapid globalization, however, economic incentives have favored more compelling education opportunities for girls (Munshi and Rosenzweig, 2006), often challenging traditional norms. In light of this backdrop, it is unclear whether there would be spillovers for women across geographic boundaries. On the one hand, crossing districts might be especially challenging. So, the women in geographic neighbors of districts receiving colleges may not benefit from this expansion in infrastructure. On the other hand, given rapidly rising returns to education in India, access to nearby colleges may also benefit the women in close proximity who live in the neighboring untreated districts.

We explore this by splitting the control districts into two sets: neighbors and non-neighbors. We then estimate a modified version of Equation 3 in which we add interactions of an indicator for neighboring controls with the overall trend and a post-trend, in addition to the original interactions of treatment indicator with these trends, respectively.[38] The excluded category is the non-neighboring control districts. Our results in Table 8 using our main sample reveal that there is no statistically significant difference between the neighboring controls and the non-neighboring controls in terms of educational attainment after the policy change.[39] The interaction of neighboring controls with post-policy trend is small and insignificant. In a comparison of treated districts with non-neighboring controls, we continue to find statistically significant trend changes post-policy. Thus, our evidence points towards the existence of a *boundary effect*. While women from rural areas benefit within a district, women from neighboring districts do not reap any benefits of increased proximity to colleges.

5.5 Benefits of Education: Labor and Marriage Market Outcomes

Table 9 documents the results of the effect of the grants policy on female labor force participation using the NFHS data. The probability that a female from the married sample participates in the labor force trends positively for the treated districts after the policy change and the coefficients are statistically significant. As in Figure 2, we present the results of the estimation using our main sample graphically in Appendix Figure A.7. For cohorts born in or before 1990 (19 years or older in 2009), we see a relatively flat trend in labor force participation. However, for the cohorts born after 1990, there is a sharp upward trend.[40] When we look at the heterogeneity in the effect of the grants policy on employment across rural and urban location (Columns 2 and 3), we find that the increase in employment is driven by rural women.[41] This is in line with our earlier findings that rural women in the treated districts benefited more in terms of educational attainment after the policy change, compared to their urban counterparts. We do not observe any effects for males (refer to Appendix Table A.14)

We examine the changes in the traits of husbands of the females in our main sample (described in the data section) after the policy change. In Table 10, we examine the husband's education in terms of the number of years of schooling and enrollment in college for all married women (Panel A) and for married women with at least a secondary education (Panel B). [42] There is a positive trend in the number of years of schooling (Column 1) after the policy change, albeit it is not significant. The probability that the husband had attended college was trending negatively in the treated districts prior to the policy (Column 2). However, the trend reversed in 2009, and it is significant (at the 10 percent level) for women with at least a secondary education. The policy led to improvements in the overall pool of potential husband's in the treated districts. The husband's probability of being employed (Column 3) also trends positively post the policy change. This effect seems to be driven by more husbands participating in non-agricultural work (Column 5) although the effects are not significant. There are two plausible channels for changes in husband quality. First, since the policy affected the educational outcomes of men as well, the average quality of the pool of men in marriage markets improved. The other possibility is that better educated women were able to match with better spouses due to improved bargaining. While we cannot completely rule out the first channel, we think the evidence weighs in favor of the second channel. This is because we find that the likelihood of women with better educational outcomes (at least secondary school education) to marry men who are enrolled in college increased after the policy change (Panel B, Column 2 of Table 10).

5.6 Other Development Outcomes

We are interested in examining whether this improvement in educational outcomes generates other synergistic improvements in development outcomes of interest. Social norms can influence women's labor market participation. However, there can be other empowering aspects of women's lives that change for the better due to access to colleges.

5.6.1 Marriage Age and Child Marriage

A vast literature focuses on the benefits of delayed age of marriage of girls (Field and Ambrus, 2008; Jensen and Thornton, 2003). In Table 11, we document the results of our trend-break specification for child marriage using our main sample. The age of marriage exhibited a negative trend in the treated districts for cohorts eligible for college prior to the policy reform. However, the trend reversed after the policy, albeit it is not statistically significant. The legal age of marriage in India is 18 and yet a large proportion of women are married at younger ages. Column 2 of Table 11 indicates that the proportion of girls getting married before the age of 18 exhibited an upward trend before the policy in the treated districts. However, after the policy, this reversed direction. The likelihood of being married before 18 years of age fell by 0.42 percentage points annually in the post-policy period relative to the overall trend. There was also a statistically significant negative intercept shift. In terms of magnitude, this is an annual gain of 8.6 percent of a standard deviation. We also report these estimates separately for rural and urban regions. Not surprisingly, we find that the trend reversal in under-age marriages is driven by the rural sample. These findings are consistent with heterogeneity results for education and labor market outcomes reported above.

5.6.2 Fertility

We then examine the effect of college expansion on actual and desired fertility. Earlier literature suggests that increased level of female education increases the age of marriage and reduces fertility (Currie and Moretti, 2003; Breierova and Duflo, 2004). We report the effect of expanded college access on fertility in Appendix Table A.15. We can see that treated districts experienced a faster decline in the number of children ever born (Column 1) and in the number of living children (Column 2) compared to those of the control districts post the policy change, albeit the effects are not statistically significant. However, there was a statistically significantly negative intercept shift after the policy change indicating a reduction in fertility.[43] The availability of contraceptives in the US increased professional education for women and raised the age at first marriage (Goldin and Katz, 2002). In the Indian context, we examine if college access influences the use of contraceptives but do not find any effect on the knowledge and use of modern contraceptives (Columns 5 and 6, respectively).

6 Conclusion

This research sheds light on the consequences of access to local colleges for women in a developing country setting. While a body of work featuring US college expansion focuses on understanding returns to education in the labor market, there is not much evidence from developing countries. Our paper provides novel evidence that access to college improves educational attainment for girls. Consistently, we find evidence of better labor market outcomes. College access in such settings can have other welfare-enhancing consequences. We shed light on some of these outcomes and find that access to local colleges reduces the prevalence of child marriage and fertility. In so far as higher age at marriage of women impacts their children positively, we also document positive developmental spillovers of investing in higher education. Thus, from a policy perspective, our findings negate the long-standing donor community view that investment in higher education has lower social returns.

While men have freedom to travel for education or work, women are more confined in their choices and more so in rural areas. Our findings indicate that rural women benefit from access to local colleges. They are more likely to attend and complete college and be employed. Men benefit too, but the benefit to women are much larger reducing the gender disparity in access to higher education. Thus, investment in higher education in the form of increased access to local colleges has the potential to empower women, reduce gender disparities, and have social benefits in developing countries.

This paper also furthers our understanding of labor market trends for women in India. A salient finding of our paper is that the increased years of schooling and college enrollment and completion translate into higher employment opportunities for women. India has witnessed a secular decline in women's labor market participation despite a continual increase in girls' schooling over time (Afridi et al., 2018; Fletcher et al., 2017). This puzzling pattern has been subjected to many speculative explanations. Amid rising concerns for women's safety, lack of local educational opportunities might be affecting women's human capital, thus affecting their labor market participation. According to a recent global poll conducted by Thompson Reuters in 2018, India has been voted the unsafest place for women in the world. Borker (2017) shows that safety concerns affect the college quality decisions of females in India. Our findings imply that improving access to college education might be effective at increasing employment opportunities for girls. The proximity of the higher education institutions may matter for not only reducing the cost of access but also the cost of safety.

Finally, there might be complementarities between increased local college access and other demand-side interventions such as increasing local recruiting efforts (Jensen, 2012). Investigating how these levers might align is an important avenue for future research.

Notes

[1]GER in higher education is defined as the ratio of population enrolled in higher education institutes to the total population in the age group 18-23.

[2]Field and Ambrus (2008) finds that each additional year that marriage is delayed increases the dowry cost to parents by approximately 11 percent.

[3]Jensen (2012) and Heath and Mobarak (2015) find that improved labor market opportunities increased girls' education in India and Bangladesh, respectively.

[4]See, for example, Ashraf et al. (2020) and Bau (2019).

[5]For instance, Krakauer (2018) documents that most schools remain unused after a school construction program in Afghanistan and Pakistan.

[6]This approach allows for selection on observables and is not predicated on an identifying assumption of parallel pre-trends. Burgess and Pande (2005) use this type of approach to identify the causal impact of rural banks on poverty and Greenstone and Hanna (2014) use it to isolate the impact of environmental regulation on child mortality in India.

[7]Education investments can have returns in both labor and marriage markets. See for example, Chiappori et al. (2009).

[8]Fertility choice consequences of a German college expansion program is studied by Kamh¨ofer and Westphal (2017).

[9]However, research has also documented that this is less relevant for academically capable students (Hoxby, 1997) and for more recent cohorts of students who are less space-bound due to technological innovations (Allen and Seaman, 2013).

[10]Doyle and Skinner (2016) examine the effects on wages for women and find that women's labor market outcomes are better.

[11]A handful of studies have highlighted the market gains from education for women. Goldin (2006) argues that greater levels of human capital investment by women increased labor market returns for women relative

to men post 1970. Similarly, Chiappori et al. (2017) find evidence for marital premium for college educated women.

[12]https://indianexpress.com/article/education/indias-gross-enrolment-ratio-in-higher-education-up-by-0-7-5012579/

[13]This program was termed as *'Scheme for providing financial assistance to New Model Colleges in Educationally Backward Districts'.*

[14]All state institutions were eligible for these grants regardless of the GER of the districts.

[15]We only use the number of colleges in a district and their year of establishment from this survey. We do not have access to the cross-walks that would enable us to use any other data collected in this survey.

[16]According to the AISHE annual report for 2018, 4.3 percent of the registered colleges did not fill the survey but they were not concentrated in any specific area.

[17]These approximations yield GER values very close to the ones used by the MHRD.

[18]For a detailed discussion of the night light data, see Henderson et al. (2012).

[19]College expansion continues to happen in all districts for a variety of other reasons as well.

[20]Between 2001 and 2009, some of the eligible districts were split into 2. The parent district is considered as treated, and the 27 offshoot districts are in the control group. The results are not sensitive to including or excluding these districts.

[21]These are the year-by-year differences-in-differences estimates.

[22]Here, γ_1 captures the linear trend in eligible districts until 2008 and γ_2 captures the difference in trend before and after the policy for the eligible districts. The sum of these two coefficients gives us the overall trend from 2009 onward.

[23]The time period of our different samples are specified in the results section.

[24]Our approach is different from Burgess and Pande (2005) in two ways. One, they use the initial financial development of states - a continuous variable - interacted with trends to estimate a trend-break model.

We do use an analog of that specification when we utilize the 2001 district GER instead of the treatment indicator to interact with the overall and post-trend (Appendix Table A.2). Two, they use a cumulative stock of the bank branches as an outcome to establish the first stage. Thus, the comparison of estimates across two consecutive years sheds light on the marginal effects. A close analog of the model we estimate (using indicators for treatment) is applied in Greenstone and Hanna (2014).

[25]Each supremum test statistic is the maximum value of the test statistic that is obtained from a series of Wald or LR (likelihood-ratio) tests over a range of possible break dates in the sample.

[26]These results are robust to inclusion of different control variables.

[27]If individuals born before 1991 are affected by this policy, say because of grade repetition, they become college eligible at a later age. In that case, we would get an attenuated result and effects would be larger post-policy change.

[28]We add Panel B in order to allay the concern that some of our controls are contemporaneous and might be endogenous. The results are very similar across the two specifications. But throughout the paper we use the estimations in Panel A to interpret the coefficients and to plot the figures.

[29]The Wald test rejects the null of no structural break in 1990 at the 1 percent significance level.

[30] One possible explanation for the larger magnitude of our estimates could be that distance constraints are more likely to bind in developing countries, and this is especially the case for females who are deemed more at risk than males.

[31]Prior research has demonstrated large effects of K-12 school infrastructure on girls' educational attainment. For example, improvements in girl-friendly school infrastructure in India led to enrollment gains of 6-7 percentage points for girls in upper primary schools (Meller and Litschig, 2016), and of 15.5 percentage points for girls in primary schools in Burkina Faso (Kazianga et al., 2019). Our estimates suggest a 16.8 percentage points increase in college enrollment over a period of eight years.

[32]Some states also had elections that year, and there was a switch in the largest party in the coalition at the federal level. Hence, there can be a lot of variation in the alignment with the federal government.

[33]In results available on request, we also find that the results are consistent in a sample of only unmarried non-migrating women.

[34]Note that the question in the survey pertains to the household migration. Unlike the US, young unmarried girls of college-going age do not typically move out of their residence without their family.

[35]We thank an anonymous referee for this suggestion

[36]The Chow test suggests that this difference across the urban and rural areas with respect to college enrollment and college completion is statistically significant at the 10 percent level of significant.

[37]We do acknowledge that wealth is measured contemporaneously and might be endogenous to the treatment. Nevertheless, the results have important policy implications.

[38]We would like to thank an anonymous referee for this suggestion.

[39]For the sake of brevity, we include the results of the specification without controls in the appendix (Table A.12).

[40]The Wald test rejects the null of no structural break in 1990 at the 1 percent significance level.

[41]In our sample, 63 percent women in rural areas and 70 percent in urban areas are not employed. Among those who are employed, main sectors of work are agriculture and allied work, services, clerical work, sales, professional and managerial positions.

[42]For the sake of brevity, we include the results of the specification without controls in the Appendix (Table A.13).

[43]Since the age at marriage for women in treated areas increases, we acknowledge the possibility that the women in treated areas after the policy have not completed their fertility cycle yet.

References

Adukia, A. (2017). Sanitation and education. *American Economic Journal: Applied Economics*, 9(2):23–59.

Afridi, F., Dinkelman, T., and Mahajan, K. (2018). Why are fewer married women joining the work force in rural India? A decomposition analysis over two decades. *Journal of Population Economics*, 31(3):783–818.

Alderman, H., Orazem, P. F., and Paterno, E. M. (2001). School quality, school cost, and the public/private school choices of low-income households in pakistan. *Journal of Human resources*, pages 304–326.

Allen, I. E. and Seaman, J. (2013). *Changing course: Ten years of tracking online education in the United States.* ERIC.

Asher, S. and Novosad, P. (2017). Politics and local economic growth: Evidence from India. *American Economic Journal: Applied Economics*, 9(1):229–73.

Ashraf, N., Bau, N., Nunn, N., and Voena, A. (2020). Bride price and female education. *Journal of Political Economy*, 128(2):591–641.

Barrera-Osorio, F., Blakeslee, D. S., Hoover, M., Linden, L. L., and Raju, D. (2011). Expanding educational opportunities in remote parts of the world: Evidence from a RCT of a public-private partnership in Pakistan. In *Third Institute for the Study of Labor (IZA) Workshop,"Child Labor in Developing Countries," Mexico City*. Citeseer.

Bau, N. (2019). Can policy change culture? Government pension plans and traditional kinship practices.

Borker, G. (2017). Safety first: Perceived risk of street harassment and educational choices of women. *Job Market Paper, Department of Economics, Brown University*.

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

Burde, D. and Linden, L. L. (2013). Bringing education to afghan girls: A randomized controlled trial of village-based schools. *American Economic Journal: Applied Economics*, 5(3):27–40.

Burgess, R. and Pande, R. (2005). Do rural banks matter? Evidence from the Indian social banking experiment. *American Economic Review*, 95(3):780–795.

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, 17(2):372–404.

Card, D. (1993). Using geographic variation in college proximity to estimate the return to schooling. Technical report, National Bureau of Economic Research.

Carneiro, P. and Heckman, J. J. (2002). The evidence on credit constraints in post-secondary schooling. *The Economic Journal*, 112(482):705–734.

Carneiro, P., Heckman, J. J., and Vytlacil, E. J. (2011). Estimating marginal returns to education. *American Economic Review*, 101(6):2754–81.

Cheema, A., Khwaja, A. I., Naseer, F., and Shapiro, J. N. (2018). Glass walls: Experimental evidence on access constraints faced by women. Technical report, Technical report.

Chiappori, P.-A., Iyigun, M., and Weiss, Y. (2009). Investment in schooling and the marriage market. *American Economic Review*, 99(5):1689–1713.

Chiappori, P.-A., Salani´e, B., and Weiss, Y. (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review*, 107(8):2109–67.

Currie, J. and Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*, 118(4):1495–1532.

Dean, J. T. and Jayachandran, S. (2019). Changing family attitudes to promote female employment. In *AEA Papers and Proceedings*, volume 109, pages 138–42.

Doyle, W. R. and Skinner, B. T. (2016). Estimating the education-earnings equation using geographic variation. *Economics of Education Review*, 53:254–267.

Field, E. and Ambrus, A. (2008). Early marriage, age of menarche, and female schooling attainment in Bangladesh. *Journal of Political Economy*, 116(5):881–930.

Fletcher, E., Pande, R., and Moore, C. M. T. (2017). Women and work in India: Descriptive evidence and a review of potential policies.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *American Economic Review*, 96(2):1–21.

Goldin, C. and Katz, L. F. (2002). The power of the pill: Oral contraceptives and women's career and marriage decisions. *Journal of Political Economy*, 110(4):730–770.

Greenstone, M. and Hanna, R. (2014). Environmental regulations, air and water pollution, and infant mortality in India. *American Economic Review*, 104(10):3038–72.

Heath, R. and Mobarak, A. M. (2015). Manufacturing growth and the lives of Bangladeshi women. *Journal of Development Economics*, 115:1–15.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102(2):994–1028.

Hoxby, C. M. (1997). How the changing market structure of us higher education explains college tuition. Technical report, National Bureau of Economic Research.

Ilie, S. and Rose, P. (2016). Is equal access to higher education in South Asia and SubSaharan Africa achievable by 2030? *Higher Education*, 72(4):435–455.

Jacoby, H. G. and Mansuri, G. (2011). *Crossing boundaries: Gender, caste and schooling in rural Pakistan*. The World Bank.

Jagnani, M. and Khanna, G. (2020). The effects of elite public colleges on primary and secondary schooling markets in India. *Journal of Development Economics*, 146:102512.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *Economics*, 7(1):63–88.

Jayachandran, S. (2020). Social norms as a barrier to women's employment in developing countries. Technical report, National Bureau of Economic Research.

Jensen, R. (2012). Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India. *The Quarterly Journal of Economics*, 127(2):753–792.

Jensen, R. and Thornton, R. (2003). Early female marriage in the developing world. *Gender & Development*, 11(2):9–19.

Jones, A. M. (2008). Afghanistan on the educational road to access and equity. *Asia Pacific Journal of Education*, 28(3):277–290.

Kamho¨fer, D. A. and Westphal, M. (2017). Fertility effects of college education: Evidence from the German educational expansion. Technical report, Ruhr Economic Papers.

Kane, T. J. and Rouse, C. E. (1995). Labor-market returns to two-and four-year college. *The American Economic Review*, 85(3):600–614.

Kazianga, H., Levy, D., Linden, L. L., and Sloan, M. (2013). The effects of "girl-friendly" schools: Evidence from the BRIGHT school construction program in Burkina Faso. *American Economic Journal: Applied Economics*, 5(3):41–62.

Kazianga, H., Linden, L. L., Protik, A., and Sloan, M. (2019). The medium-term impacts of girl-friendly schools: Seven-year evidence from school construction in Burkina Faso. Technical report, National Bureau of Economic Research.

Kerr, C. (1991). *The great transformation in higher education, 1960-1980*. SUNY Press.

Kim, J., Alderman, H., and Orazem, P. F. (1999). Can private school subsidies increase enrollment for the poor? The Quetta Urban Fellowship Program. *The World Bank Economic Review*, 13(3):443–465.

Krakauer, J. (2018). *Three cups of deceit*. Anchor.

Meller, M. and Litschig, S. (2016). Adapting the supply of education to the needs of girls: Evidence from a policy experiment in rural India. *Journal of Human Resources*, 51(3):760–802.

Mohanty, S. K., Govil, D., Chauhan, R. K., Kim, R., and Subramanian, S. (2016). Estimates of poverty and inequality in the districts of India, 2011–2012. *Journal of Development Policy and Practice*, 1(2):142–202.

Munshi, K. and Rosenzweig, M. (2006). Traditional institutions meet the modern world: Caste, gender, and schooling choice in a globalizing economy. *American Economic Review*, 96(4):1225–1252.

Muralidharan, K. and Prakash, N. (2017). Cycling to school: Increasing secondary school enrollment for girls in India. *American Economic Journal: Applied Economics*, 9(3):321– 50.

Psacharopoulos, G. and Patrinos, H. A. (2004). Returns to investment in education: A further update. *Education Economics*, 12(2):111–134.

Sekhri, S. and Debnath, S. (2014). Intergenerational consequences of early age marriages of girls: Effect on children's human capital. *The Journal of Development Studies*, 50(12):1670–1686.

Seki, M., Yamada, E., et al. (2020). Heterogeneous effects of urban public transportation on employment by gender: Evidence from the Delhi Metro. Technical report.

Sutton, M. (1998). Girls educational access and attainment. *Women in the Third World: And Encyclopedia of contemporary issues*, pages 381–396.

World Bank, U. (2000). *Higher education in developing countries: Peril and promise*. Number 440. World Bank.

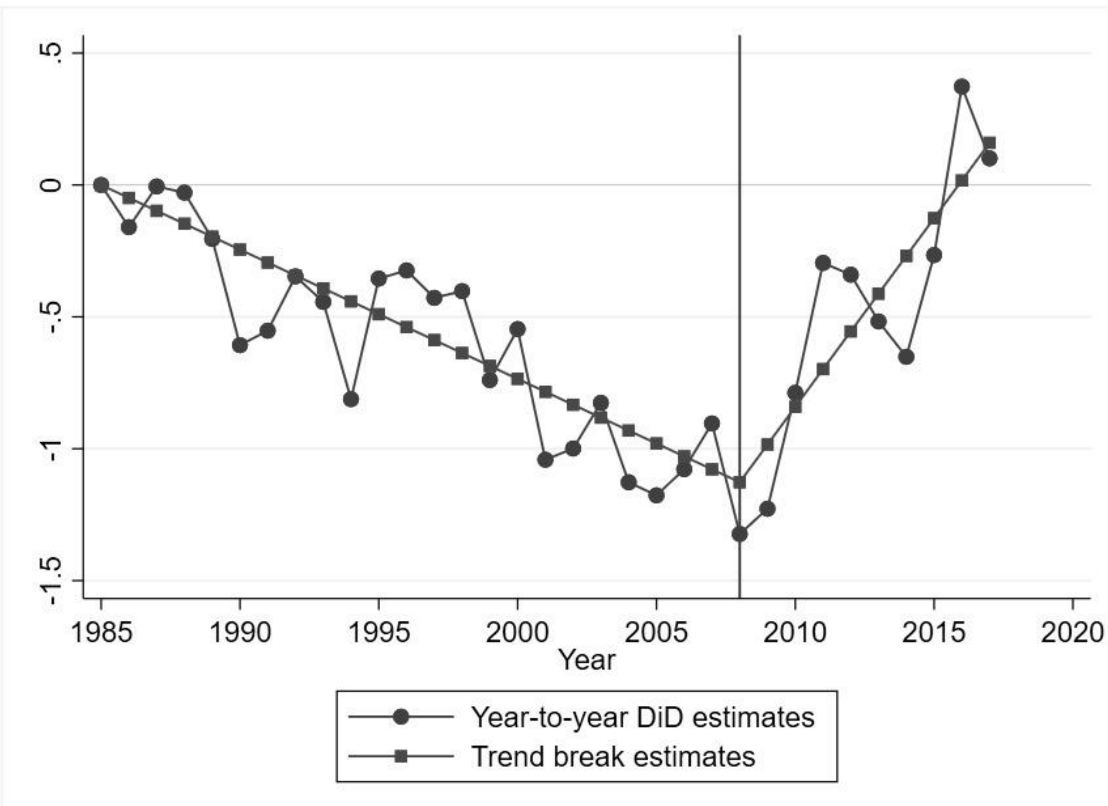
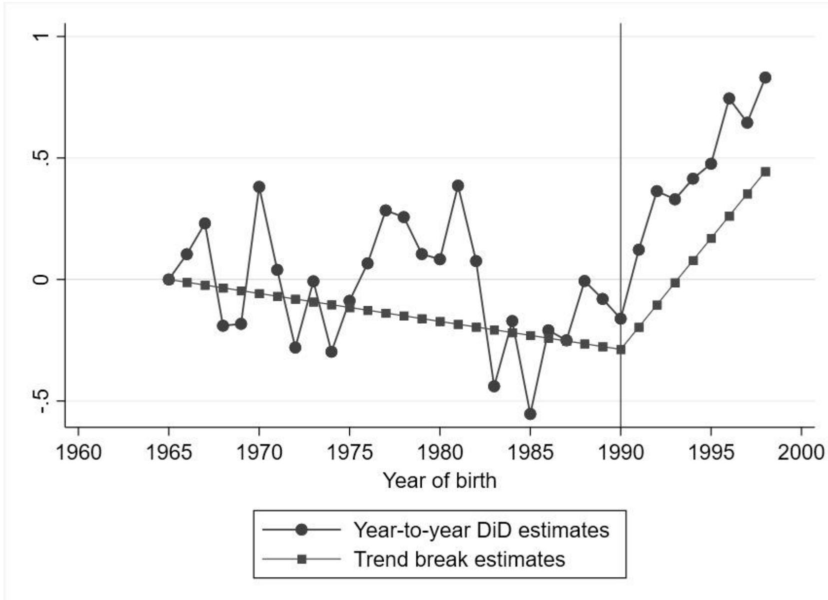
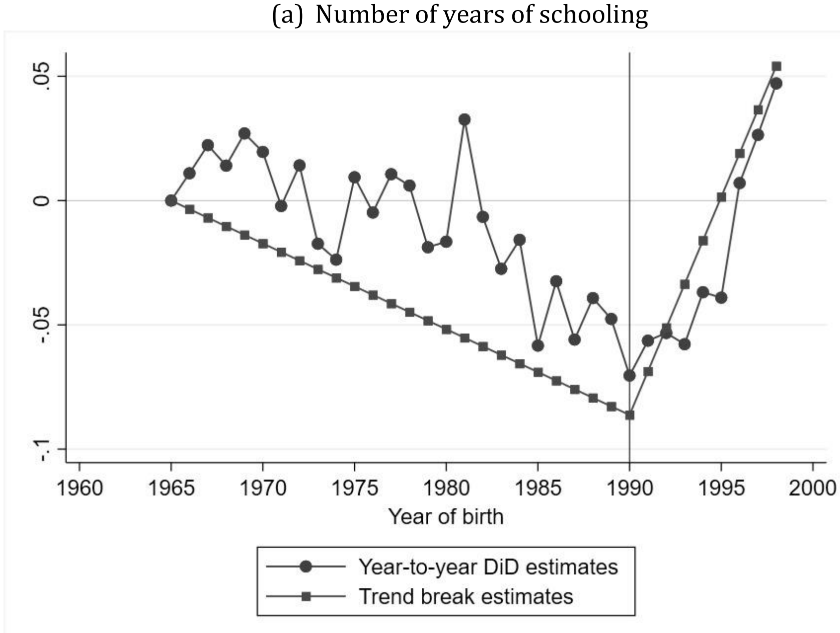


Figure 1: College Expansion by Treatment Status

Notes: The series "Year-to-year DID estimates" graphs the annual coefficients on treatment from a Difference-in-Difference regression of the form described in Equation 1. The series "Trend break estimates" graphs the annual coefficient implied by the trend-break model in Column 1 of Table 2. In both cases, the dependent variable is the number of colleges constructed in a district.



(a) Number of years of schooling



(b) College enrollment

Figure 2: Educational Attainment by Treatment Status

Notes: The series "Year-to-year DID estimates" graphs the annual coefficients on treatment from a Difference-in-Difference regression analogous to the form described in Equation 1. The series "Trend break estimates" graphs the annual coefficient implied by the trend-break model in Table 3. The dependent variable is the number of years of schooling and college enrollment in panel (a) and (b), respectively.

Table 1: Summary Statistics for Our Sample

VARIABLE	Obs	Mean	Std. Dev.	Min	Max
Panel A: Women's profile					
Years of schooling	141,641	8.58	4.54	0	20
Highest level of education is primary	141,641	0.09	0.29	0	1
Highest level of education is secondary	141,641	0.62	0.49	0	1
Enrolls into college	141,641	0.15	0.36	0	1
Finishes college	141,641	0.096	0.29	0	1
Age at marriage	33,577	18.66	4.79	0	48
Marriage before 18 years of age	141,641	0.10	0.29	0	1
Member of workforce	24,378	0.27	0.44	0	1
Panel B: Married women's husbands' profile					
Years of schooling	6,158	7.01	5.06	0	20
Highest level of education is primary	6,158	0.15	0.36	0	1
Highest level of education is secondary	6,158	0.50	0.50	0	1
Enrolls into college	6,158	0.11	0.32	0	1
Member of workforce	6,075	0.94	0.24	0	1
Panel C: Men's profile					
Years of schooling	73,347	8.25	4.64	0	20
Highest level of education is primary	73,347	0.13	0.33	0	1
Highest level of education is secondary	73,347	0.59	0.49	0	1
Enrolls into college	73,347	0.14	0.35	0	1
Finishes college	73,347	0.10	0.31	0	1
Member of workforce	73,255	0.77	0.42	0	1

Notes: Data for Panel A and Panel B are from the NFHS women's survey, while data in Panel C is from NFHS men's survey. The sample is restricted to individual that have always stayed at the same place. Only a subset of the surveyed women was asked about their employment and, if married, their husband's education and employment.

Table 2: Effect of the Grants Policy on Number of Colleges Constructed

	Number of colleges constructed					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Grant Policy x (T-1985) trend	-0.0497***	-0.0322*	-0.0434**	-0.0459**	-0.0394**	-0.0733***
	(0.0185)	(0.0169)	(0.0183)	(0.0185)	(0.0179)	(0.0180)
Grant Policy x (T-2009) trend	0.192***	0.149***	0.187***	0.194***	0.176***	0.192***
	(0.0513)	(0.0504)	(0.0510)	(0.0516)	(0.0499)	(0.0514)
Grant Policy x Post	0.0600	0.00696	-0.0862	-0.0698	-0.108	0.0600
	(0.201)	(0.188)	(0.206)	(0.207)	(0.205)	(0.201)
Control	None	Night lights	Poverty gap	Squared poverty gap	Head count ratio	State specific trend
Observations	20,955	20,955	20,394	20,394	20,394	20,955
R-squared	0.444	0.449	0.457	0.455	0.460	0.487
F-stat 1	13.37	9.063	13.60	14.08	12.57	8.765
Sig. level	0.000277	0.00271	0.000246	0.000806	0.00177	0.00319
F-stat 2	7.049	4.559	6.945	7.206	6.404	8.908
Sig. level	0.000938	0.0108	0.00104	0.000192	0.000421	0.000153

Notes: Data is from 2018 AISHE. The unit of analysis is a district-year. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for years 2009 and after and 0 otherwise. All regressions include district and year fixed effects. F-stat 1 reports the F statistic from a test that the sum of coefficients γ_1 and γ_2 from Equation 2 is 0. F-stat 2 reports the F statistic from a test of the joint significance of γ_1 and γ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table 3: Effect of the Grants Policy on Educational Attainment of Girls

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	No of years of schooling	Highest level of education is primary	Highest level of education is Secondary	Enrolls into college	Finishes college	Number of years in college
Panel A: With socio-economic controls						
Grant Policy x Overall trend	-0.0115	-0.000228	0.00463***	-0.00345***	-0.00294***	-0.0128***
	(0.0126)	(0.000661)	(0.00107)	(0.000921)	(0.000910)	(0.00377)
Grant Policy x Post Policy trend	0.103***	-0.00145	-0.0183***	0.0210***	0.0170***	0.0664***
	(0.0274)	(0.00117)	(0.00303)	(0.00319)	(0.00314)	(0.0121)
Grant Policy x Post	0.251*	0.00616	0.0725***	-0.0464***	-0.00476	-0.0305
	(0.135)	(0.00776)	(0.0169)	(0.0144)	(0.0138)	(0.0528)
Observations	110,467	110,467	110,467	110,467	110,467	110,467
F-stat 1	18.69	2.940	28.49	43.12	30.30	31.10
Sig. level	1.79e-05	0.0869	1.31e-07	1.07e-10	3.50e-07	2.45e-07
F-stat 2	9.388	1.525	18.88	21.97	15.22	15.59
Sig. level	9.59e-05	0.218	1.08e-08	5.91e-10	5.37e-08	3.63e-08
Panel B: Without socio-economic controls						
Grant Policy x Overall trend	-0.0134	-0.000175	0.00467***	-0.00360***	-0.00304***	-0.0133***
	(0.0147)	(0.000658)	(0.00110)	(0.00105)	(0.00102)	(0.00424)
Grant Policy x Post Policy trend	0.0988***	-0.00139	-0.0183***	0.0207***	0.0169***	0.0656***
	(0.0342)	(0.00123)	(0.00307)	(0.00356)	(0.00352)	(0.0138)
Grant Policy x Post	0.321*	0.00465	0.0726***	-0.0421***	-0.00173	-0.0167
	(0.166)	(0.00789)	(0.0172)	(0.0160)	(0.0152)	(0.0589)
Observations	110,467	110,467	110,467	110,467	110,467	110,467
F-stat 1	10.61	2.331	28.15	34.19	23.38	22.98
Sig. level	0.00118	0.299	1.55e-07	7.99e-09	1.66e-06	2.03e-06
F-stat 2	5.335	1.210	18.32	17.33	11.77	11.58
Sig. level	0.00504	0.127	1.84e-08	4.70e-08	9.52e-06	1.15e-05

Notes: Data is from the NFHS woman's survey. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. The variable *number of years in college* is a categorical variable which takes the value of zero if an individual does not attend college, and values between one and five for individuals attending college corresponding to the year they are in college (1, 2, and 3 for the first, second, and third year of college, respectively; 4 for year one in Masters and 5 for year two in Masters or beyond). All regressions include district and year fixed effects. Socio-economic controls include household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table 4: Effect of the Grants Policy on Educational Attainment of Girls (Full Sample)

VARIABLES	(1) No of years of schooling	(2) Highest level of education is primary	(3) Highest level of education is Secondary	(4) Enrolls into college	(5) Finishes college	(6) Number of years in college
Panel A: With socio-economic controls						
Grant Policy x Overall trend	-0.0158***	0.00183***	0.00174***	-0.00222***	-0.00205***	-0.00788***
	(0.00543)	(0.000359)	(0.000532)	(0.000361)	(0.000336)	(0.00136)
Grant Policy x Post Policy trend	0.134***	-0.00441***	-0.00586***	0.0142***	0.0144***	0.0523***
	(0.0170)	(0.000841)	(0.00171)	(0.00174)	(0.00170)	(0.00636)
Grant Policy x Post	-0.0345	0.00262	0.0560***	-0.0430***	-0.0239***	-0.101***
	(0.0659)	(0.00392)	(0.00866)	(0.00739)	(0.00637)	(0.0236)
Observations	650,773	650,773	650,773	650,773	650,773	650,773
F-stat 1	63.32	11.67	7.081	61.32	73.50	70.52
Sig. level	0	1.54e-08	0.000451	0	0	0
F-stat 2	32.26	18.51	7.798	33.45	36.79	35.41
Sig. level	0	0.000675	0.00799	0	0	0
Panel B: Without socio-economic controls						
Grant Policy x Overall trend	-0.0177***	0.00184***	0.00154***	-0.00222***	-0.00204***	-0.00783***
	(0.00580)	(0.000361)	(0.000541)	(0.000369)	(0.000344)	(0.00139)
Grant Policy x Post Policy trend	0.146***	-0.00455***	-0.00537***	0.0146***	0.0148***	0.0539***
	(0.0182)	(0.000855)	(0.00171)	(0.00177)	(0.00176)	(0.00661)
Grant Policy x Post	-0.0211	0.00285	0.0578***	-0.0433***	-0.0240***	-0.102***
	(0.0771)	(0.00397)	(0.00884)	(0.00769)	(0.00676)	(0.0251)
Observations	650,773	650,773	650,773	650,773	650,773	650,773
F-stat 1	69.28	12.45	6.284	64.69	72.60	69.53
Sig. level	0	0.000447	0.0124	0	0	0
F-stat 2	34.68	18.87	6.133	34.14	36.40	35.06
Sig. level	0	1.09e-08	0.00230	0	0	0

Notes: Data is from the NFHS woman's survey. The unit of analysis is at the individual level. **The analysis is for full sample of women (both who always lived in the same place and who did not).** Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. The variable *number of years in college* is a categorical variable which takes the value of zero if an individual does not attend college, and values between one and five for individuals attending college corresponding to the year they are in college (1, 2, and 3 for the first, second, and third year of college, respectively; 4 for year one in Masters and 5 for year two in Masters or beyond). All regressions include district and year fixed effects. Socio-economic controls include household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table 5: Effect of the Grants Policy on Men's Educational Attainment

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	No of years of schooling	Highest level of education is primary	Highest level of education is Secondary	Enrolls into college	Finishes college	Number of years in college
Panel A: With socio-economic controls						
Treatment x Overall trend	0.0212***	-0.000381	0.00284***	6.78e-05	0.000198	0.000699
	(0.00645)	(0.000466)	(0.000734)	(0.000516)	(0.000474)	(0.00188)
Treatment x Post Policy trend	0.0583**	-0.000108	-0.00586*	0.00634**	0.00247	0.0127
	(0.0259)	(0.00194)	(0.00309)	(0.00268)	(0.00231)	(0.00839)
Treatment x Post	-0.189	-0.00401	0.0229	-0.0235	0.00941	-0.000642
	(0.150)	(0.0115)	(0.0184)	(0.0152)	(0.0138)	(0.0489)
Observations	67,450	67,450	67,450	67,450	67,450	67,450
F-stat 1	9.975	0.0662	1.063	6.454	1.547	3.052
Sig. level	0.00166	0.699	0.303	0.0113	0.351	0.148
F-stat 2	10.22	0.358	7.696	3.471	1.048	1.916
Sig. level	4.27e-05	0.797	0.000498	0.0317	0.214	0.0811
Panel B: Without socio-economic controls						
Treatment x Overall trend	0.0148**	-0.000222	0.00279***	-0.000293	-0.000104	-0.000599
	(0.00705)	(0.000473)	(0.000747)	(0.000541)	(0.000496)	(0.00198)
Treatment x Post Policy trend	0.0770***	-0.000592	-0.00609*	0.00764***	0.00354	0.0170**
	(0.0263)	(0.00195)	(0.00312)	(0.00268)	(0.00232)	(0.00841)
Treatment x Post	-0.126	-0.00534	0.0248	-0.0209	0.0118	0.0112
	(0.165)	(0.0117)	(0.0184)	(0.0156)	(0.0142)	(0.0507)
Observations	67,751	67,751	67,751	67,751	67,751	67,751
F-stat 1	13.05	0.183	1.262	8.578	2.545	4.547
Sig. level	0.000328	0.669	0.262	0.00352	0.277	0.0333
F-stat 2	8.685	0.197	7.187	4.292	1.285	2.308
Sig. level	0.000190	0.821	0.000819	0.0141	0.111	0.100

Notes: Data is from the NFHS man's survey. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. The variable *number of years in college* is a categorical variable which takes the value of zero if an individual does not attend college, and values between one and five for individuals attending college corresponding to the year they are in college (1, 2, and 3 for the first, second, and third year of college, respectively; 4 for year one in Masters and 5 for year two in Masters or beyond). All regressions include district and year fixed effects. Socio-economic controls include household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table 6: Effect of the Grants Policy on Educational Attainment of Girls by Region

	Years of schooling		College enrollment		College completion	
	Urban	Rural	Urban	Rural	Urban	Rural
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: With socio-economic controls						
Grant policy x Overall trend	0.0131	-0.0262*	-0.000812	-0.00299***	-0.000216	-0.00252***
	(0.0170)	(0.0141)	(0.00147)	(0.000873)	(0.00144)	(0.000846)
Grant policy x Post Policy trend	0.0366	0.125***	0.0113**	0.0211***	0.00676	0.0156***
	(0.0348)	(0.0342)	(0.00488)	(0.00374)	(0.00440)	(0.00337)
Grant policy x Post	0.0917	0.120	-0.0116	-0.0681***	0.0133	-0.0207
	(0.201)	(0.173)	(0.0243)	(0.0177)	(0.0240)	(0.0149)
Observations	31,612	78,855	31,612	78,855	31,612	78,855
F-stat 1	2.920	13.41	5.936	31.19	3.080	21.13
Sig. level	0.0880	0.000271	0.0151	3.50e-08	0.0797	5.20e-06
F-stat 2	1.872	6.947	2.969	16.02	1.632	10.81
Sig. level	0.155	0.00104	0.0521	1.64e-07	0.196	2.43e-05
Panel B: Without socio-economic controls						
Grant policy x Overall trend	0.0175	-0.0285*	-0.000524	-0.00314***	-1.57e-05	-0.00262***
	(0.0194)	(0.0157)	(0.00161)	(0.000971)	(0.00155)	(0.000921)
Grant policy x Post Policy trend	0.00130	0.135***	0.00884*	0.0215***	0.00496	0.0159***
	(0.0386)	(0.0403)	(0.00492)	(0.00406)	(0.00456)	(0.00363)
Grant policy x Post	0.144	0.148	-0.00655	-0.0660***	0.0174	-0.0193
	(0.235)	(0.203)	(0.0258)	(0.0191)	(0.0254)	(0.0160)
Observations	31,612	78,855	31,612	78,855	31,612	78,855
F-stat 1	0.359	11.18	3.792	27.69	1.684	18.87
Sig. level	0.549	0.000876	0.149	1.96e-07	0.195	7.53e-05
F-stat 2	0.650	5.798	1.909	14.15	0.940	9.641
Sig. level	0.522	0.00320	0.0520	9.76e-07	0.391	1.63e-05

Notes: Data is from the NHFS woman's survey. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. Socio-economic control includes household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table 7: Effect of the Grants Policy on College Enrollment of Girls by Wealth

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Poorest	Poorer	Middle	Richer	Richest
Panel A: With socio-economic controls					
Grant policy x Overall trend	-0.000776*	-0.000995	-0.00152	-0.00319**	-0.000285
	(0.000467)	(0.000792)	(0.00117)	(0.00153)	(0.00183)
Grant policy x Post Policy trend	0.00699	0.00974**	0.0127***	0.0216***	0.0119**
	(0.00468)	(0.00443)	(0.00411)	(0.00476)	(0.00503)
Grant policy x Post	-0.0263	-0.0337	-0.0578**	-0.0529*	-0.0221
	(0.0267)	(0.0252)	(0.0258)	(0.0289)	(0.0257)
Observations	18,451	24,012	25,289	22,090	20,625
F-stat 1	1.861	4.571	8.701	19.32	6.742
Sig. level	0.173	0.0887	0.00889	1.30e-05	0.0344
F-stat 2	1.902	2.432	4.759	10.36	3.388
Sig. level	0.150	0.0329	0.00330	3.73e-05	0.00964
Panel B: Without socio-economic controls					
Grant policy x Overall trend	-0.000731	-0.00103	-0.00152	-0.00329**	-0.000228
	(0.000465)	(0.000788)	(0.00116)	(0.00153)	(0.00186)
Grant policy x Post Policy trend	0.00695	0.00981**	0.0129***	0.0217***	0.0113**
	(0.00468)	(0.00443)	(0.00411)	(0.00481)	(0.00501)
Grant policy x Post	-0.0268	-0.0332	-0.0591**	-0.0515*	-0.0216
	(0.0267)	(0.0253)	(0.0259)	(0.0291)	(0.0260)
Observations	18,451	24,012	25,289	22,090	20,625
F-stat 1	1.866	4.607	8.995	18.77	6.310
Sig. level	0.172	0.0322	0.00767	4.60e-05	0.0123
F-stat 2	1.785	2.483	4.908	10.15	3.172
Sig. level	0.169	0.0844	0.00281	1.71e-05	0.0426

Notes: Data is from the NHFS woman's survey. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. Socio-economic control includes household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table 8: Spillover Analysis: Effect of the Grants Policy on Educational Attainment of Girls

	(1)	(2)	(3)	(4)	(5)
VARIABLES	No of years of schooling	Highest level of education is primary	Highest level of education is Secondary	Enrolls into college	Finishes college
Treatment x Overall trend	0.0369**	0.000926	0.00441**	-0.000549	-0.000314
	(0.0187)	(0.000924)	(0.00192)	(0.00180)	(0.00183)
Treatment x Post Policy trend	0.0866**	-0.00189	-0.0176***	0.0189***	0.0121*
	(0.0416)	(0.00188)	(0.00604)	(0.00574)	(0.00638)
Treatment x Post	0.389	0.00543	0.128***	-0.0647**	-0.0203
	(0.256)	(0.0103)	(0.0337)	(0.0272)	(0.0251)
Neighboring controls x Overall trend	0.0620***	0.00147	-0.000190	0.00367*	0.00329*
	(0.0202)	(0.000923)	(0.00200)	(0.00188)	(0.00190)
Neighboring controls x Post Policy trend	-0.0256	-0.000666	0.000729	-0.00275	-0.00605
	(0.0435)	(0.00185)	(0.00633)	(0.00608)	(0.00662)
Neighboring controls x Post	0.170	-0.000878	0.0678**	-0.0226	-0.0198
	(0.267)	(0.00991)	(0.0343)	(0.0284)	(0.0262)
Observations	110,467	110,467	110,467	110,467	110,467
R-squared	0.443	0.069	0.203	0.288	0.259
F-stat 1	11.13	0.350	6.453	14.52	5.120
Sig. level	0.000898	0.614	0.993	0.0907	0.0240
F-stat 2	7.567	0.674	4.511	8.021	3.387
Sig. level	0.000565	0.510	0.920	0.857	0.195
F-stat 3	0.883	0.254	0.0100	0.0324	0.262
Sig. level	0.348	0.245	0.0113	0.000152	0.609
F-stat 4	5.150	1.408	0.00717	2.409	1.638
Sig. level	0.00604	0.554	0.0113	0.000363	0.0344

Notes: Data is from the NFHS woman's survey. The unit of analysis is at the individual level. Treatment is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Neighboring control is a dummy variable that takes a value of 1 if the control district is a neighbor to any treatment district. The omitted category is non-neighboring control districts. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. All regressions control for household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in the augmented version of Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. F-stat 3 and F-stat 4 are the corresponding test statistics (analogous to F-stat 1 and F-stat 2, respectively) for the coefficients on the interactions of the neighboring controls dummy with the overall and the post policy trends. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table 9: Effect of the Grants Policy on Female Labor Force Participation

	(1)	(2)	(3)
VARIABLES	All	Urban	Rural
Panel A: With socio-economic controls			
Grant policy x Overall trend	0.000360	0.00429	-0.00136
	(0.00224)	(0.00506)	(0.00252)
Grant policy x Post Policy trend	0.0297**	-0.00819	0.0459***
	(0.0124)	(0.0270)	(0.0152)
Grant policy x Post	-0.0233	-0.0308	-0.0385
	(0.0611)	(0.134)	(0.0720)
Observations	6,088	1,538	4,550
F-stat 1	6.169	0.0210	8.940
Sig. level	0.0133	0.885	0.00292
F-stat 2	3.133	0.381	4.545
Sig. level	0.0444	0.683	0.0110
Panel B: Without socio-economic controls			
Grant policy x Overall trend	0.000432	0.00419	-0.00147
	(0.00224)	(0.00507)	(0.00253)
Grant policy x Post Policy trend	0.0315**	-0.00742	0.0475***
	(0.0123)	(0.0267)	(0.0147)
Grant policy x Post	-0.0336	-0.0419	-0.0464
	(0.0606)	(0.132)	(0.0703)
Observations	6,088	1,538	4,550
F-stat 1	7.145	0.0146	10.25
Sig. level	0.00773	0.904	0.00145
F-stat 2	3.621	0.361	5.224
Sig. level	0.0274	0.697	0.00567

Notes: Data is from the NFHS woman's survey. Information on the husband's education and employment was collected for a subset of all surveyed women. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. Socio-economic control includes household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table 10: Effect of the Grants Policy on Husband's Education and Employment

VARIABLES	(1) No of years of schooling	(2) Enrolled into college	(3) Member of workforce	(4) Participate in agricultural work	(5) Participate in non-agricultural work
Panel A: All women					
Treatment x Overall trend	-0.0165	-0.00318**	0.000793	0.000969	-0.000176
	(0.0215)	(0.00143)	(0.000977)	(0.00193)	(0.00189)
Treatment x Post Policy trend	0.0874	0.0130	0.0183*	-0.00457	0.0228
	(0.145)	(0.00978)	(0.00992)	(0.0143)	(0.0144)
Treatment x Post	0.417	0.0347	-0.0691	0.0338	-0.103
	(0.703)	(0.0491)	(0.0424)	(0.0666)	(0.0704)
Observations	6,074	6,074	6,075	6,075	6,075
R-squared	0.432	0.261	0.191	0.354	0.333
F stat	0.238	1.034	3.661	0.0655	2.602
Sig. level	0.626	0.310	0.0562	0.798	0.107
Panel B: Women with at least a secondary education					
Treatment x Overall trend	-0.0244	-0.00542*	-0.000601	0.000782	-0.00138
	(0.0315)	(0.00279)	(0.00147)	(0.00287)	(0.00291)
Treatment x Post Policy trend	0.230	0.0283*	0.00896	-0.0112	0.0202
	(0.174)	(0.0159)	(0.0129)	(0.0177)	(0.0184)
Treatment x Post	0.0506	0.0392	-0.0274	-0.000652	-0.0267
	(0.734)	(0.0742)	(0.0558)	(0.0810)	(0.0867)
Observations	3,066	3,066	3,089	3,089	3,089
R-squared	0.378	0.315	0.301	0.387	0.385
F stat	1.445	2.118	0.419	0.352	1.065
Sig. level	0.230	0.146	0.517	0.553	0.303

Notes: Data is from the NFHS woman's survey. Information on the husband's education and employment was collected for a subset of all surveyed women. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. No socio-economic controls are included. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table 11: Effect of the Grants Policy on Child Marriage

	All		Urban		Rural	
	Age at marriage	Married before 18 years of age	Age at marriage	Married before 18 years of age	Age at marriage	Married before 18 years of age
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: With socio-economic controls						
Grant policy x Overall trend	-0.0286*	0.00152	-0.0351*	0.000824	-0.0237	0.00137
	(0.0165)	(0.00119)	(0.0196)	(0.00129)	(0.0195)	(0.00142)
Grant policy x Post Policy trend	0.0388	-0.00422**	-0.0488	-0.00153	0.0563	-0.00481**
	(0.0484)	(0.00165)	(0.0998)	(0.00156)	(0.0523)	(0.00199)
Grant policy x Post	0.240	-0.0257**	1.000**	-0.0239**	-0.0504	-0.0132
	(0.227)	(0.01000)	(0.506)	(0.0114)	(0.268)	(0.0128)
Observations	33,577	141,641	8,224	38,786	25,353	102,855
F-stat 1	0.0570	10.31	0.782	0.770	0.476	11.62
Sig. level	0.811	0.00566	0.377	0.572	0.396	0.000694
F-stat 2	1.505	5.216	2.300	0.560	0.928	5.818
Sig. level	0.223	0.00139	0.101	0.380	0.490	0.00314
Panel B: Without socio-economic controls						
Grant policy x Overall trend	-0.0263	0.00154	-0.0321	0.000756	-0.0232	0.00140
	(0.0170)	(0.00121)	(0.0204)	(0.00131)	(0.0199)	(0.00143)
Grant policy x Post Policy trend	0.0296	-0.00428**	-0.0188	-0.00115	0.0392	-0.00500**
	(0.0503)	(0.00171)	(0.101)	(0.00160)	(0.0538)	(0.00203)
Grant policy x Post	0.271	-0.0262**	0.843*	-0.0244**	0.0228	-0.0131
	(0.233)	(0.0102)	(0.510)	(0.0116)	(0.275)	(0.0129)
Observations	33,577	141,641	8,224	38,786	25,353	102,855
F-stat 1	0.00556	9.663	0.282	0.233	0.110	12.12
Sig. level	0.941	0.00196	0.216	0.630	0.494	0.00243
F-stat 2	1.228	4.869	1.537	0.263	0.706	6.079
Sig. level	0.294	0.00797	0.595	0.769	0.740	0.000535

Notes: Data is from the NHFS woman's survey. The sample includes only ever-married women. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. Socio-economic controls include household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Supplementary Appendix

A.1 Effect Size: Increase in College Enrollment vs Increase in Capacity

In this section, we compute the effect size of college enrollment in levels to assess if this is plausible given the expansion in college capacity.[44]

1. *Capacity Count Based on Official UGC Statistics: Based on our estimates, and using an average approximate number of seats of 1,417 per college, the approximate increase in enrollment capacity per district would be approximately 9,659 over a period of eight years (2009-2016)*

   The AISHE data does not provide cross-walks so we cannot estimate the exact change in capacity or number of seats in districts over years. Hence, we approximate capacity and acknowledge that it is not precise. Our capacity count is based on a report published by the University Grants Commission (UGC) of India, the nodal body regulating higher education. In their 2008 publication titled "Higher Education in India: Issues Related to Expansion, Inclusiveness, Quality and Finance", UGC summarized average enrollment per college by states based on a sample of 1400 colleges across India.[45] Based on these state-wise average survey based enrollment per college (assuming in the survey colleges, available seats must be atleast as large as enrollment), we computed the average capacity of Indian colleges by taking the weighted average of the average enrollment capacity per college of each state, where the weight is based on the population aged between 20-24 in these states. This count is 1,417.

   Based on our estimates in Table II, an additional 0.192 colleges are added each year. So, in year 1, 0.192 additional colleges are operating. This will generate an additional approximate enrollment capacity of 0.192×1,417 in the first year (where we use the fact that average enrollment per college in India is 1,417 students). In year 2, an additional 0.192 colleges are operational, which implies a total of 0.192×2 (=0.384) more colleges relative to the baseline year. This in turn implies an additional capacity of 0.384×1,417 in the second year relative to the baseline. Continuing in this manner, in year 8, there are 0.192×8×1,417 additional seats. Summing the capacity increases from years 1 through 8 yields an estimated 9,659 more seats created over the 8-year period. The calculations for the estimated capacity increases are as follows:

Capacity increase = (0.192×1×1,417) [Year 1] + (0.192×2×1,417) [Year 2] + (0.192×3×1,417)

[Year 3] +(0.192×4×1,417) [Year 4] + (0.192×5×1,417) [Year 5] + (0.192×6×1,417) [Year 6] + (0.192×7×1,417) [Year 7] +(0.192×8×1,417) [Year 8] = 9,659

Note that intake in colleges happens only in year 1 (the first year of degree programs, though very limited transfer students are accepted in years 2 or 3 of the Bachelor's program). So, seats or capacity imply year 1 capacity, not the aggregate for all years that the student is in college.

2. *Enrollment: Based on our estimates, and using an average baseline male, and female population of 86,241 and 80,932 aged 20-24 per treated district, respectively, the increase in college enrollment per district is approximately 13,327 students over a period of eight years (2009-2016).*

As mentioned in our paper, we restrict our sample to women who stayed in the same place all their lives (we do this to ensure that we know the women's region of birth and do not include migrants in our sample). Such women comprise 17 percent of 18+ aged women (as per the survey data we used). Thus, we impose this restriction on the age-relevant female population from the Census of India of 2011. According to the Census of India, 2011, there were around 80,932 females aged 20-24 per treated districts. That implies a total of (80,932*0.17) =13,758 women per district who have likely stayed in the same place. Combined with the 2.1 annual pp increase in enrollment, roughly an additional 289 (=13,758*0.021) females should have enrolled in college each year which would give us a total of 2,311 (=289*8) females enrolled in college during the eight years after the intervention among the non-movers.

We next estimate the effect for those females who moved (in-migrants). we find an annual increase of 1.1 pp in college enrollment (Appendix Table A.16). This implies an additional 739 = (0.011*80932*0.83) females should have enrolled in college each year in this sample. For 8 years, this would add up to 5,912 (=739*8) females enrolled in college. Then, we sum the enrollment from both of these samples (i.e., non-movers and movers) and resulting enrollment of 8,223 = (2,311+5,912) females during the eight years after the intervention is plausible given the approximated capacity.

The average relevant college-going male population per treated district is 86,241. For men who always stayed at the same place (65% of the sample), Table V, column (4) shows an effect on the likelihood of college enrollment of about 0.6 pp per year. This implies an additional 2,690 = (0.006*86241*0.65*8) males enrolled in college during the eight years after the intervention. For men who always stayed at the same place (35% of the sample), we find an increase of 0.1 pp per year in the likelihood of college enrollment. This implies an additional 2,414=(0.01*86241*0.35*8) males enrolled in college during the eight years after the intervention. Adding enrollment from both of these samples, we find an additional 5,104 = (2,690+2,414) males would have enrolled in college during the eight years after the intervention.

Adding enrollment for both males and females, we find an additional 13,327 = (8,223+5,104) males and females would have enrolled in college during the eight years after the intervention. We summarize this discussion in Appendix Table A.5.

3. We found that the increase in college enrollment per district is approximately 13,327 students over a period of eight years (2009-2016). For the same duration, the approximate increase in capacity per district is 9,659. The enrollment figure we obtain exceeds the approximate capacity figure. We do not have data to examine the precise increase in capacity. On scrutinizing various reports and news articles, we think there are two reasons why this approximation of capacity might be understating the available seats. Below we summarize these reasons:

(a) Some of the colleges have evening as well as morning sessions as per the AISHE reports. For example, in Tamil Nadu, during our sample period, state colleges had 2 sessions offering the same courses: morning college sessions offered classes in the morning shift between 8.45 am to 1.15 pm and the evening shift started at 1.30 pm and ended at 6.05 pm (Hindustan Times, 2021). There are 218 state colleges in Tamil Nadu in our sample. We do not have data on what is the fraction of colleges offering morning and evening sessions overall for India. So, our capacity or available seats is plausibly underestimated due to this.

(b) Colleges that were established prior to the policy sometimes have vacant seats (Indian Express, 2014) especially in technical education which is a field chosen predominantly by boys. These vacant seats vary yearly and across regions. For example, 42 percent of the seats in technical colleges in Maharashtra were vacant in 2018 (Hindustan Times, 2019). In our sample, 17% of colleges are technical or engineering. As Table A.4, reveals, the policy leads to an increase in the number of students completing high school (37 percent increase in attainment is accruing from K-12). Some of these marginal students, especially the boys who are propelled by the policy to complete high school, might be enrolling in these other institutions. We do not have data to scrutinize this possibility but a share of increase in men's enrollment could be on this account.

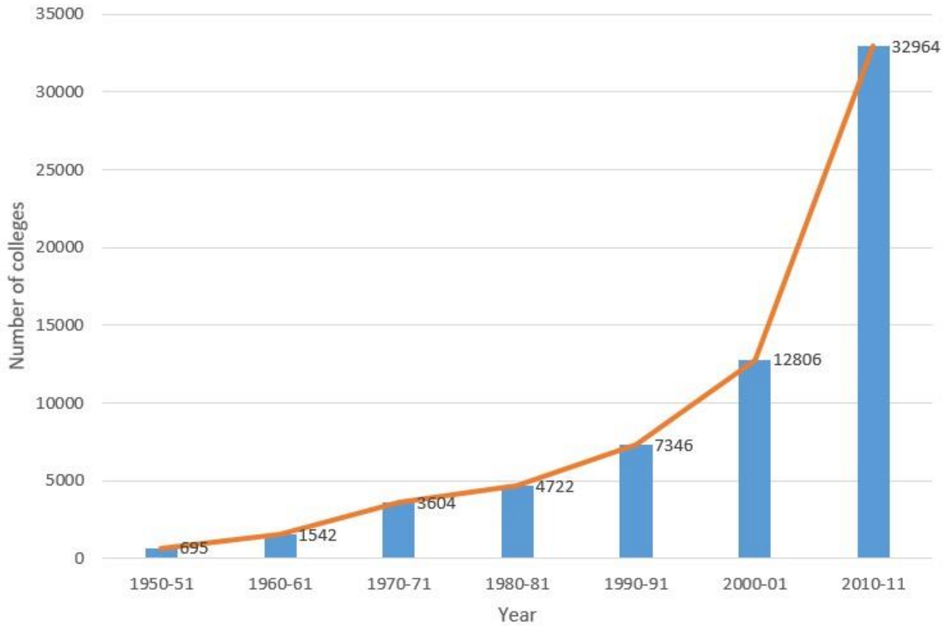
Another important point to note based on this discussion is that since our average capacity computation utilizes average enrollment in colleges in a survey conducted by the UGC (see discussion above *Capacity Count based on Official UGC Statistics*), our capacity or number of seats available is plausibly an underestimate. Our data do not allow us to specifically ascertain how much of the enrollment that exceeds the approximate capacity of the new colleges, as shown above, is on account of these two reasons.

A.2 Effect Size: Years of Schooling

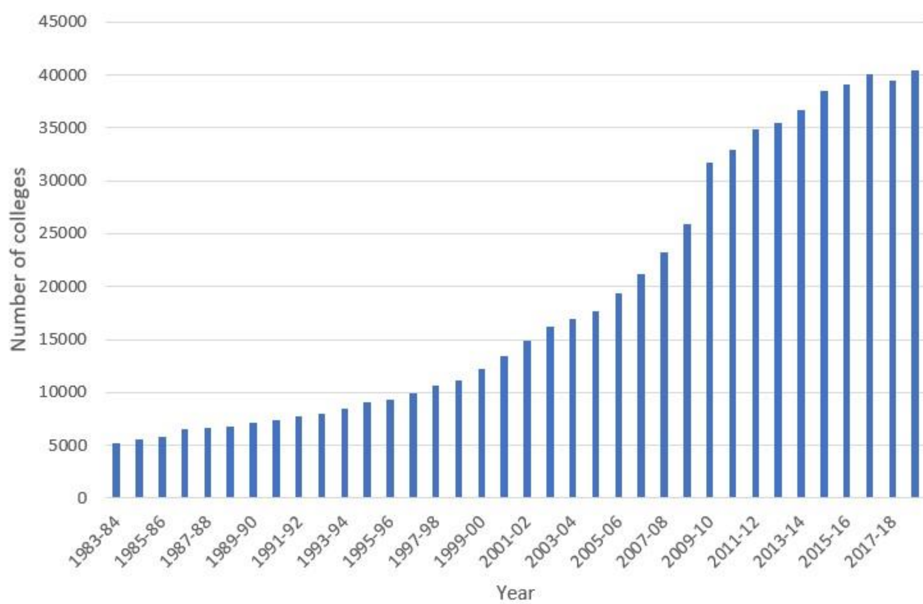
Table 3 shows that the number of years of schooling increased by 0.10 each year for the treated districts post-policy period. This translates to an increase of 0.82 years of schooling over the eight years of post-policy period. In this subsection, we try do disentangle how much of this increase is due to increase in college years and how much is due to an increase in the years of education during middle/high school years (K-12). In Appendix Table A.4, we show the effect of grant policy on number of years of schooling (Column 1), number of years of schooling for individuals who have not enrolled into college (Column 2), and number of years in college (Column 3). The variable *number of years in college* is a categorical variable which takes the value of zero if an individual does not attend college, and values between one and five for individuals attending college corresponding to the year they are in college (1, 2, and 3 for the first, second, and third year of college, respectively; 4 for year one in Masters and 5 for year two in Masters or beyond). Column 3 shows the effect of college construction on the number of college years completed. The coefficient is 0.0664 which indicates an increase of 0.53 years of college education (8*0.0664) over the eight-year period. This is comparable to the US (Doyle and Skinner, 2016).

Column 2 indicates that the increase in the total years of schooling includes an increase in the years of education during middle/high school years (K-12). We find 0.30 (=0.038*8) years increase in schooling happens for school students in the post-policy period between 2009 and 2016. The sum of the effect: 0.53 + 0.30 is what we get from an aggregate estimate in column 1. There is an incentive effect of college construction. In the literature, other studies have documented that college construction affects the incentives to go to school even if individuals do not enrol in college or complete a college degree (Jagnani and Khanna, 2020).

Appendix Figures and Tables



(a) Number of colleges in different decades



(b) Number of colleges in different years

Figure A.1: Total Number of Colleges over Time

Notes: Data for all years except 1999-2000 and 2000-01 are from different rounds of University Grant Commission (UGC) annual reports. For those two years, the numbers were extrapolated using the growth rate of colleges between 1998-99 and 2001-02.

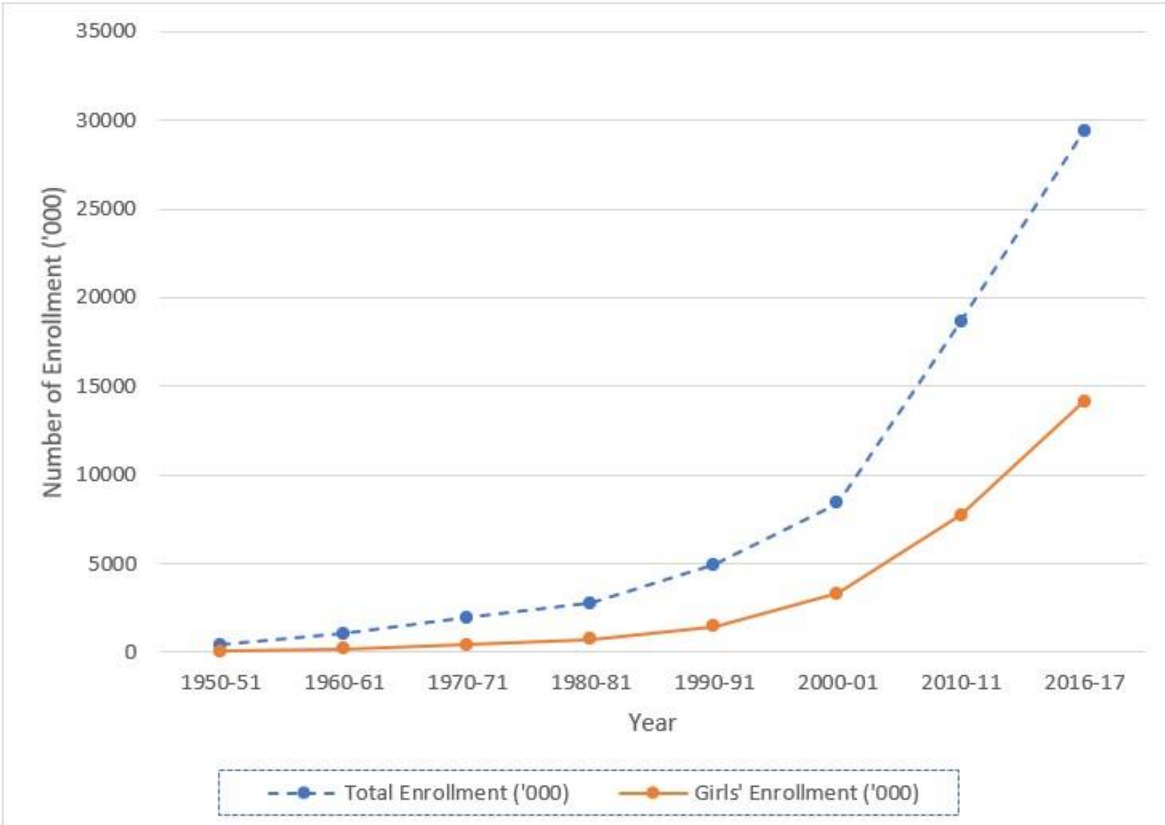


Figure A.2: College Enrollment in Different Years

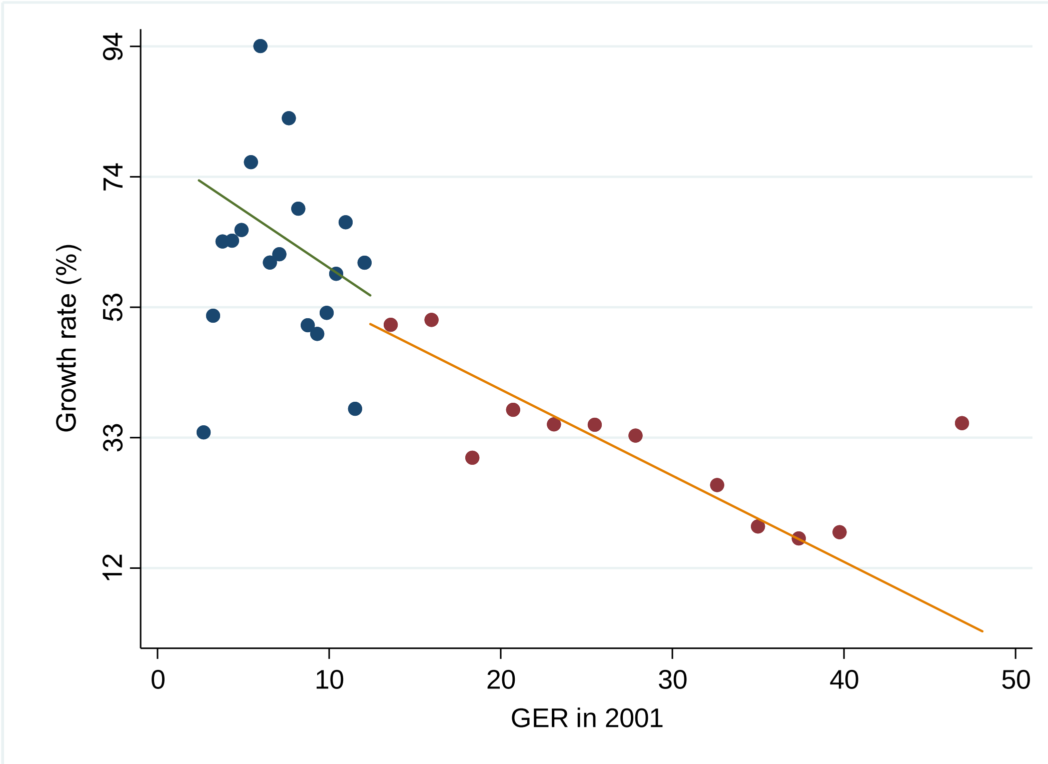


Figure A.3: 2001 GER and Growth in Colleges

Notes: The figure shows a linear fit from a regression of growth of colleges on gross enrollment ratio (GER) at the district level in 2001 separately estimated on the left and right of the GER eligibility threshold of 12.4. The dependent variable is the growth rate of total colleges between 2008 and 2018. Source: Authors' calculation.

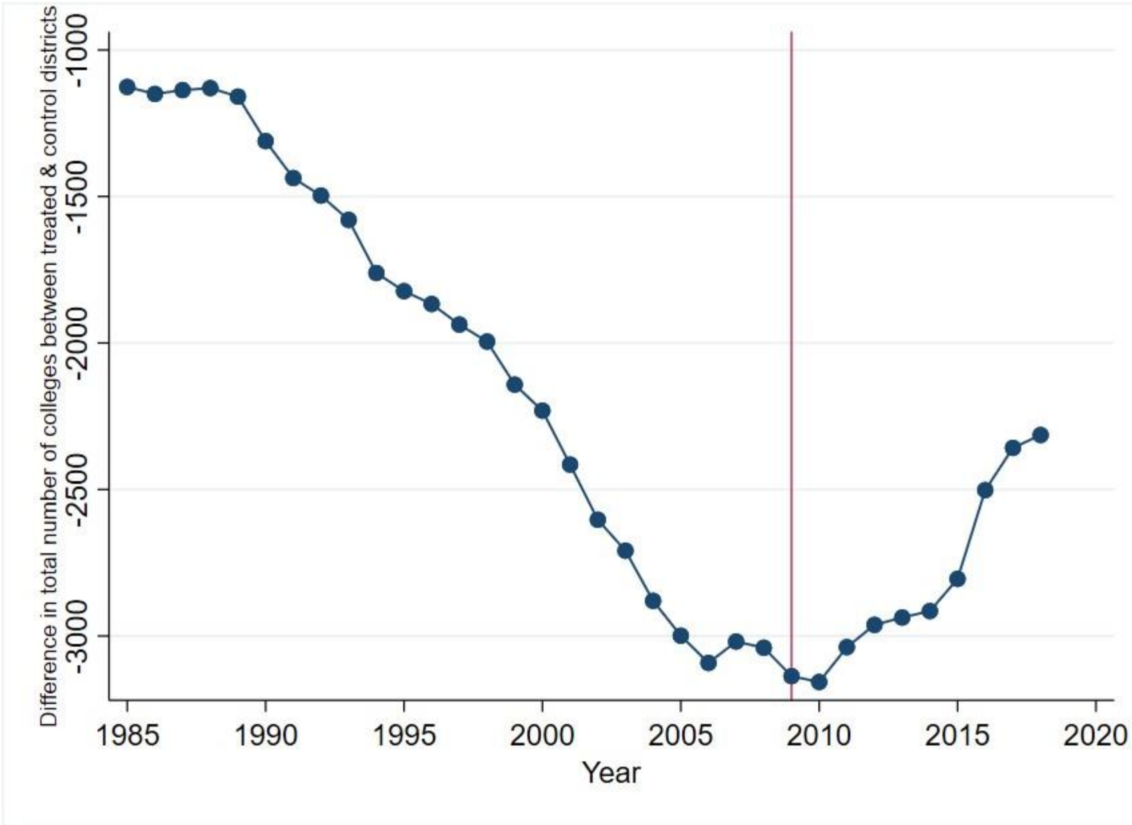


Figure A.4: Difference in the Total Stock of Colleges between the Treated and Control Districts

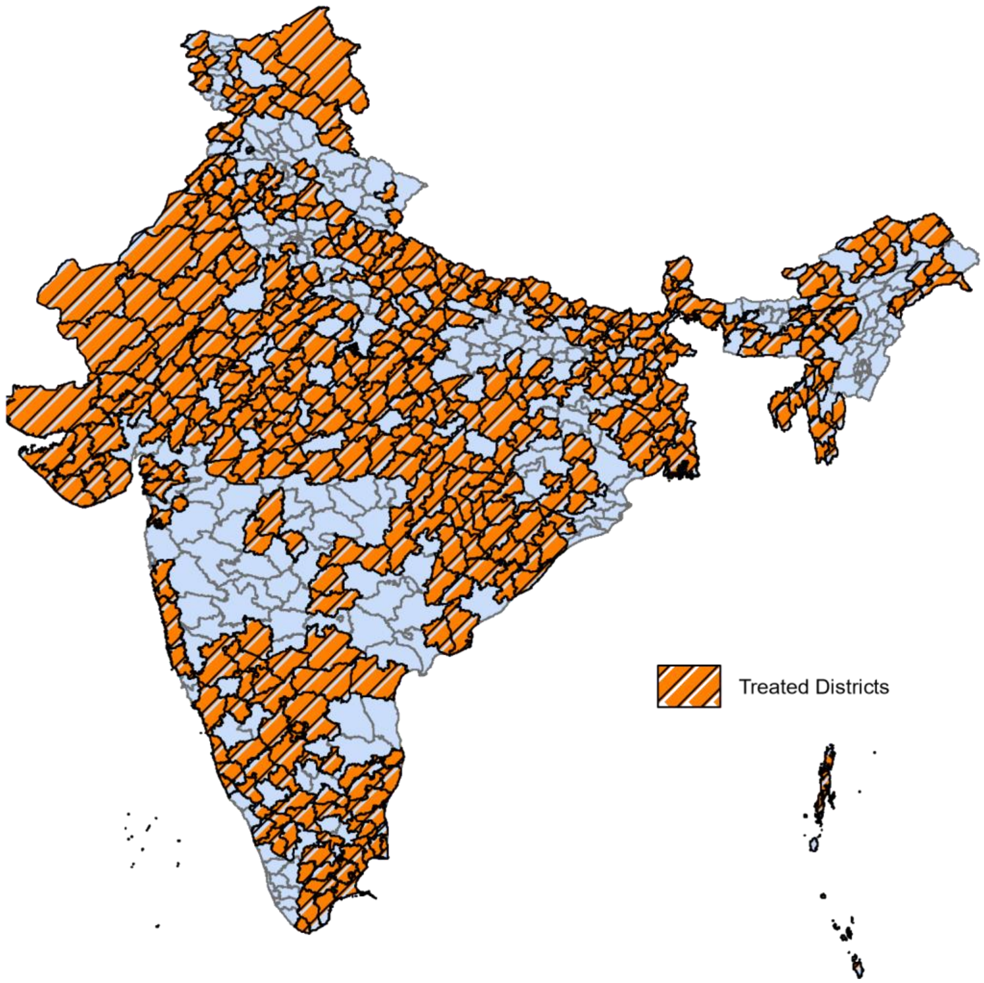
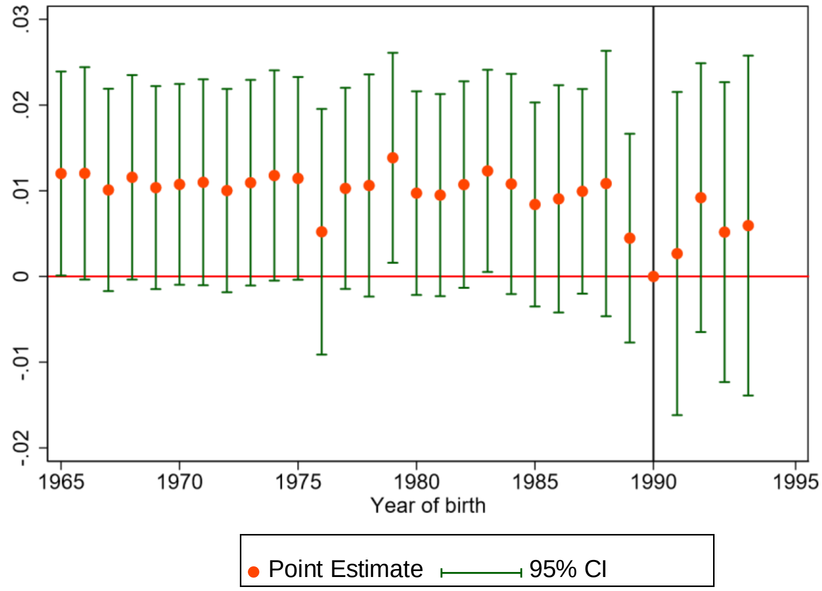
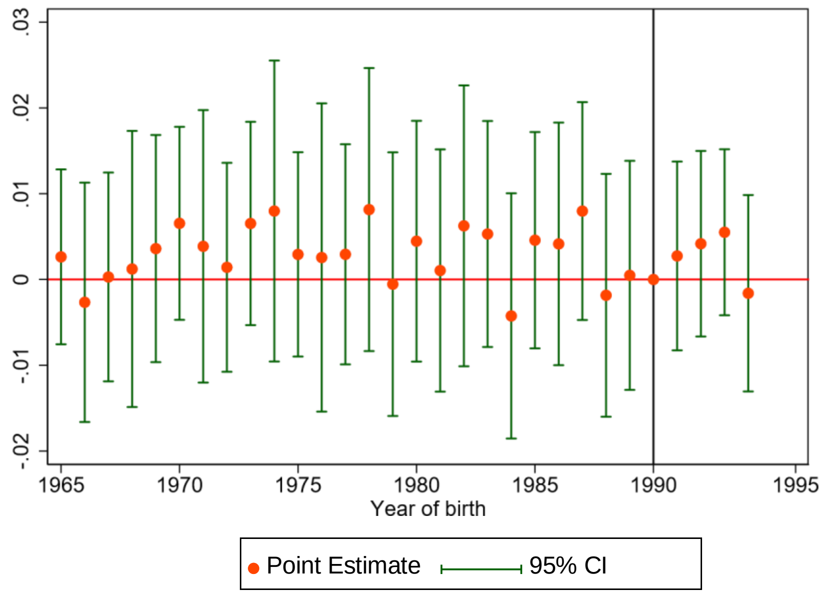


Figure A.5: Treated Districts

Note: Treated districts are colored orange while the control districts are colored grey.



(a) Migration for education



(b) Migration for work

Figure A.6: Event Study of the Effect of Grant Policy on Out-Migration

Notes: The figure plots the annual coefficients on treatment from a Difference-in-Difference regression analogous to the form described in Equation 1. The dependent variable in Panel (a) and Panel (b) is the likelihood of migrating into another district for education and work, respectively.

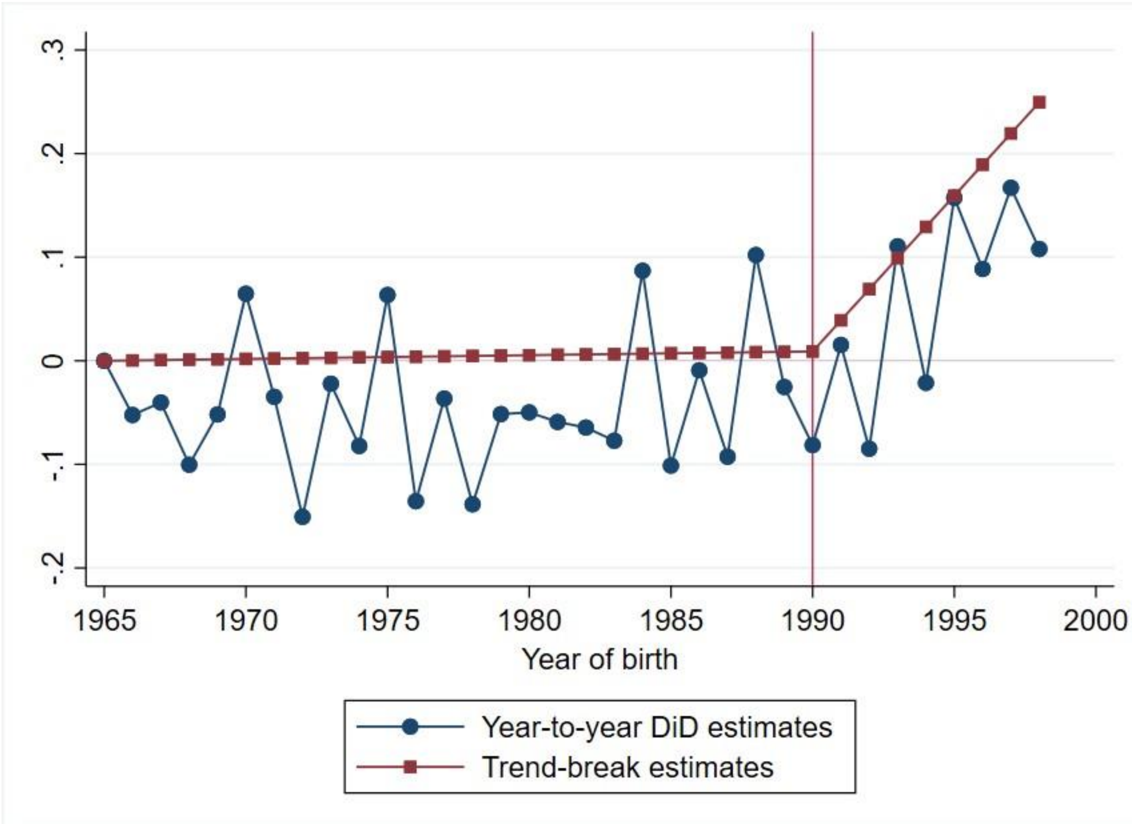


Figure A.7: Female Labor Market Participation by Treatment Status

Notes: The series "Year-to-year DID estimates" graphs the annual coefficients on treatment from a Difference-in-Difference regression analogous to the form described in Equation 1. The series "Trend break estimates" graphs the annual coefficient implied by the trend break model in Column 1 of Table 9. In both cases, the dependent variable is the likelihood that a married woman participates in the labor market.

Table A.1: Covariate Balance between Treatment and Control Districts

VARIABLES	Control mean	Difference between Treatment and Control	Standard error of difference	Observations
Panel A: Household Data				
Respondents' profile				
Years of schooling	7.47	-1.386***	0.235	40,466
Highest level of education is primary	0.12	0.039***	0.010	40,466
Highest level of education is secondary	0.41	-0.024	0.019	40,466
Enrolls into college	0.20	-0.078***	0.011	40,466
Finishes college	0.18	-0.071***	0.011	40,466
Age at marriage	19.27	-0.377*	0.225	27,147
Marriage before 18 years of age	0.37	0.037*	0.020	27,147
Member of workforce	0.40	-0.007	0.024	6,978
Participate in agricultural work	0.17	0.013	0.020	6,978
Participate in non-agricultural work	0.23	-0.021	0.017	6,978
Respondents' husbands' profile				
Years of schooling	7.33	-0.884***	0.299	5,017
Highest level of education is primary	0.14	0.026	0.017	5,017
Highest level of education is secondary	0.51	-0.037*	0.022	5,017
Enrolls into college	0.13	-0.036**	0.015	5,017
Member of workforce	0.96	-0.020**	0.009	4,951
Participate in agricultural work	0.39	0.053*	0.029	4,951
Participate in non-agricultural work	0.57	-0.073**	0.030	4,951
Panel B: District-level Information				
(From 2001 Census)				
Male-female ratio	0.93	0.005	0.005	578
Percentage of SC/ST population	0.15	0.006	0.007	578
Percentage literate	0.59	-.09***	0.009	578
Gross enrollment ratio in college	18.08	-9.93***	0.412	578
Labor force participation rate	0.39	.036***	0.005	578

Notes: Data for Panel A is from the NFHS survey. The analysis is only for the pre-policy years, i.e., cohorts born in or before 1990. Only a subset of the surveyed women was asked about their employment and (if married) their husbands' education and employment. Data for Panel B comes from the 2001 Census of India.

Table A.2: College Expansion as a Function of 2001 GER

	Number of colleges constructed				
VARIABLES	(1)	(2)	(3)	(4)	(5)
GER in 2001 x (T-1985) trend	0.00476***	0.00347**	0.00338**	0.00357**	0.00324**
	(0.00155)	(0.00135)	(0.00149)	(0.00151)	(0.00145)
GER in 2001 x (T-2009) trend	-0.0167***	-0.0136***	-0.0142***	-0.0148***	-0.0135***
	(0.00446)	(0.00429)	(0.00440)	(0.00446)	(0.00428)
GER in 2001 x Post	-0.00788	-0.00398	0.00583	0.00488	0.00551
	(0.0140)	(0.0118)	(0.0139)	(0.0140)	(0.0140)
Control	None	Night lights	Poverty gap	Squared poverty gap	Head count ratio
Observations	20,955	20,955	20,394	20,394	20,394
R-squared	0.445	0.450	0.457	0.455	0.460
F-stat 1	13.48	9.924	11.07	11.54	10.50
Sig. level	0.000261	0.00671	0.00415	0.000726	0.00545
F-stat 2	7.055	5.044	5.535	5.773	5.256
Sig. level	0.000933	0.00171	0.000931	0.00328	0.00126

Notes: Data is from the 2018 AISHE. The unit of analysis is a district-year. GER in 2001 is the gross enrollment ratio based on the 2001 Census of India. Post is an indicator variable that takes the value 1 for years 2009 and after and 0 otherwise. All regressions include district and year fixed effects. F-stat 1 reports the F statistic from a test that the sum of coefficients γ_1 and γ_2 in Equation 2 is 0. F-stat 2 reports the F statistic from a test of the joint significance of γ_1 and γ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.3: RDD Estimates of the Effect of the Grants Policy on Cumulative Colleges per 100,000 Population

	Number of total colleges per 100,000 population							
VARIABLES	2008	2009	2010	2011	2012	2013	2014	2015
Panel A: Polynomial of order 1								
Treatment	-0.206	-0.237	-0.259	-0.277	-0.289	-0.302	-0.278	-0.319
	(0.342)	(0.355)	(0.362)	(0.371)	(0.389)	(0.399)	(0.404)	(0.430)
Observations	213	222	231	231	231	231	234	243
R-squared	0.007	0.007	0.007	0.006	0.006	0.007	0.009	0.013
Bandwidth	2.758	2.911	2.996	3.044	3	3.027	3.130	3.189
Panel A: Polynomial of order 2								
Treatment	-0.026	-0.030	-0.058	-0.107	-0.089	-0.084	-0.117	-0.158
	(0.269)	(0.290)	(0.302)	(0.308)	(0.322)	(0.326)	(0.330)	(0.351)
Observations	334	342	334	331	334	342	350	354
R-squared	0.020	0.022	0.018	0.013	0.015	0.019	0.020	0.023
Bandwidth	4.532	4.645	4.532	4.462	4.537	4.658	4.755	4.825
Panel A: Polynomial of order 3								
treatment	-0.073	-0.108	-0.130	-0.107	-0.150	-0.140	-0.130	-0.187
	(0.337)	(0.368)	(0.380)	(0.394)	(0.414)	(0.425)	(0.431)	(0.465)
Observations	288	289	286	283	284	282	283	283
R-squared	0.017	0.014	0.011	0.011	0.009	0.010	0.011	0.011
Control Mean	2.679	2.813	2.901	2.959	3.017	3.067	3.149	3.203
Bandwidth	3.819	3.823	3.811	3.722	3.759	3.706	3.747	3.736
Control mean	2.679	2.813	2.901	2.959	3.017	3.067	3.149	3.203

Notes: The college data is from the 2018 AISHE. The unit of analysis is a district-year. In each regression we use optimal bandwidth as proposed by Calonico et al. (2017). Treatment is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.4: Effect of the Grants Policy on Number of Years of Schooling

VARIABLES	(1) Number of years of schooling	(2) Number of years of schooling for individuals who have not attended college	(3) Number of years in college
Grant policy x Overall trend	-0.0115	-0.00217	-0.0128***
	(0.0126)	(0.00908)	(0.00377)
Grant policy x Post Policy trend	0.103***	0.0383*	0.0664***
	(0.0274)	(0.0225)	(0.0121)
Grant policy x Post	0.251*	0.134	-0.0305
	(0.135)	(0.150)	(0.0528)
Observations	110,467	88,690	110,467
R-squared	0.443	0.415	0.298
F-stat 1	18.69	3.394	31.10
Sig. level	1.79e-05	0.184	2.45e-07
F-stat 2	9.388	1.697	15.59
Sig. level	9.59e-05	0.0659	3.63e-08

Notes: Data is from the NFHS woman's survey. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. The variable *number of years in college* is a categorical variable which takes the value of zero if an individual does not attend college, and values between one and five for individuals attending college corresponding to the year they are in college (1, 2, and 3 for the first, second, and third year of college, respectively; 4 for year one in Masters and 5 for year two in Masters or beyond). All regressions control for household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.5: Effect Size of Grants Policy on College Construction and College Enrollment

	Effect Size	8-year effect	Levels over 8 years
Increase in College Capacity	0.192	1.54 colleges	9,659 seats
Increase in enrollment for Women (Full Unrestricted sample)	0.0142	An additional 1,195 women enroll each year	9,556 students
Enrollment (weighted average for non-movers and in-migrants)			
(a) Increase in Female College Enrollment (Non-movers: Women who always lived in the same place)	0.021	An additional 289 women enroll every year	2,311 students
(b) Increase in Female College Enrollment (In-migrants-Women who did not always live in the same place)	0.011	An additional 739 women enroll every year	5,912 students
(c) Increase in Male College Enrollment (Non-movers: Men who always lived in the same place)	0.006	An additional 336 men enroll every year	2,690 students
(d) Increase in Male College Enrollment (In-migrants-men who did not always live in the same place)	0.012	An additional 302 men enroll every year	2,414 students
Total [(a) +(b) +(c) + (d)]			13,327 students

Notes: Estimates are based on authors' calculation.

Table A.6: Effect of the Grants Policy on Men's Educational Attainment (Full Sample)

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	No of years of schooling	Highest level of education is primary	Highest level of education is Secondary	Enrolls into college	Finishes college	Number of years in college
Panel A: With socio-economic Controls						
Treatment x Overall trend	0.0219***	1.86e-05	0.00277***	4.10e-05	0.000189	0.000627
	(0.00524)	(0.000379)	(0.000574)	(0.000429)	(0.000394)	(0.00160)
Treatment x Post Policy trend	0.0677***	-0.00114	-0.00654**	0.00850***	0.00553**	0.0216***
	(0.0219)	(0.00156)	(0.00264)	(0.00238)	(0.00217)	(0.00780)
Treatment x Post	-0.195	-0.00974	0.0229	-0.0286**	-0.00367	-0.0274
	(0.124)	(0.00914)	(0.0158)	(0.0134)	(0.0123)	(0.0438)
Observations	104,036	104,036	104,036	104,036	104,036	104,036
F-stat 1	17.88	0.546	2.264	14.59	8.105	9.703
Sig. level	2.70e-05	0.460	7.23e-06	0.000147	0.00456	0.00192
F-stat 2	17.93	0.274	12.06	7.921	5.141	5.910
Sig. level	2.65e-08	0.760	0.133	0.000400	0.00610	0.00286
Panel B: Without socio-economic controls						
Treatment x Overall trend	0.0207***	1.11e-05	0.00276***	-7.10e-06	0.000142	0.000427
	(0.00592)	(0.000386)	(0.000581)	(0.000461)	(0.000423)	(0.00171)
Treatment x Post Policy trend	0.0699***	-0.00106	-0.00685***	0.00876***	0.00573***	0.0224***
	(0.0228)	(0.00158)	(0.00265)	(0.00240)	(0.00220)	(0.00790)
Treatment x Post	-0.163	-0.0107	0.0243	-0.0275**	-0.00245	-0.0219
	(0.136)	(0.00916)	(0.0156)	(0.0138)	(0.0127)	(0.0454)
Observations	104,471	104,471	104,471	104,471	104,471	104,471
F-stat 1	17.37	0.472	2.676	15.25	8.380	10.07
Sig. level	3.51e-05	0.790	9.57e-06	0.000104	0.00392	0.00158
F-stat 2	15.50	0.236	11.77	8.145	5.109	5.941
Sig. level	2.67e-07	0.492	0.102	0.000322	0.00629	0.00278

Notes: Data is from the NFHS man's survey. The unit of analysis is at the individual level. **The analysis is for full sample of men (both who always lived in the same place and who did not).** Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. The variable *number of years in college* is a categorical variable which takes the value of zero if an individual does not attend college, and values between one and five for individuals attending college corresponding to the year they are in college (1, 2, and 3 for the first, second, and third year of college, respectively; 4 for year one in Masters and 5 for year two in Masters or beyond). All regressions include district and year fixed effects. Socio-economic controls include household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.7: Effect of the Grants Policy on Educational Attainments with Different Controls

VARIABLES	(1)	(2)	(3)	(4)	(5)
Panel A: Number of years of schooling					
Grant policy x Overall trend	-0.0134	-0.0104	-0.00930	-0.00899	-0.00907
	(0.0147)	(0.0150)	(0.0149)	(0.0150)	(0.0150)
Grant policy x Post Policy trend	0.0988***	0.0727**	0.103***	0.103***	0.103***
	(0.0342)	(0.0343)	(0.0353)	(0.0353)	(0.0353)
Grant policy x Post	0.321*	0.354**	0.308*	0.301*	0.299*
	(0.166)	(0.174)	(0.171)	(0.171)	(0.171)
Observations	110,467	109,837	107,024	107,024	107,024
R-squared	0.288	0.289	0.290	0.290	0.290
F-stat 1	10.61	5.540	12.09	12.21	12.21
Sig. level	0.00118	0.0189	0.000542	0.000509	0.000509
F-stat 2	10.61	5.540	12.09	12.21	12.21
Sig. level	0.00118	0.0189	0.000542	0.000509	0.000509
Panel B: College enrollment					
Grant policy x Overall trend	-0.00360***	-0.00281***	-0.00363***	-0.00358***	-0.00358***
	(0.00105)	(0.00106)	(0.00109)	(0.00111)	(0.00111)
Grant policy x Post Policy trend	0.0207***	0.0164***	0.0217***	0.0218***	0.0218***
	(0.00356)	(0.00359)	(0.00367)	(0.00368)	(0.00369)
Grant policy x Post	-0.0421***	-0.0245	-0.0449***	-0.0456***	-0.0457***
	(0.0160)	(0.0165)	(0.0166)	(0.0166)	(0.0166)
Observations	110,467	109,837	107,024	107,024	107,024
R-squared	0.194	0.196	0.198	0.198	0.197
F-stat 1	34.19	20.58	35.75	36.34	36.37
Sig. level	7.99e-09	6.82e-06	3.78e-09	2.85e-09	2.80e-09
F-stat 2	34.19	20.58	35.75	36.34	36.37
Sig. level	7.99e-09	6.82e-06	3.78e-09	2.85e-09	2.80e-09
Control	**None**	**Night light**	**Poverty gap**	**Squared poverty gap**	**Head count ratio**

Notes: Data is from the NFHS woman's survey. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. No socio-economic controls are included. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.8: Effect of the Grants Policy: Sensitivity to Different Sample Periods

	Number of colleges constructed			
VARIABLES	(1)	(2)	(3)	(4)
Grant Policy x Overall trend	-0.0396***	-0.0497***	-0.0450*	-0.0725**
	(0.0140)	(0.0185)	(0.0242)	(0.0356)
Grant Policy x Post-Policy trend	0.182***	0.192***	0.187***	0.215***
	(0.0479)	(0.0513)	(0.0558)	(0.0659)
Grant Policy x Post	-0.0246	0.0600	0.0299	0.163
	(0.191)	(0.201)	(0.215)	(0.245)
Observations	24,130	20,955	17,780	14,605
R-squared	0.420	0.444	0.480	0.510
F-stat 1	13.42	13.37	13.29	13.18
Sig. level	0.000269	0.000277	0.000289	0.000305
F-stat 2	7.208	7.049	6.649	6.593
Sig. level	0.000803	0.000938	0.00139	0.000305
Sample Period	1981-2017	1985-2017	1990-2017	1995-2017

Notes: Data is from 2018 AISHE. The unit of analysis is a district-year. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for years 2009 and after and 0 otherwise. All regressions include district and year fixed effects. F-stat 1 reports the F statistic from a test that the sum of coefficients γ_1 and γ_2 in Equation 2 is 0. F-stat 2 reports the F statistic from a test of the joint significance of γ_1 and γ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.9: Changes in Political Affiliation by Treatment Status

VARIABLES	Affiliation with ruling alliance (1)	(2)
Grant Policy x Overall trend	0.00787	0.0189
	(0.0154)	(0.0148)
Grant Policy x Post Policy trend	-0.00611	-0.0176
	(0.0185)	(0.0161)
Grant Policy x Post	-0.0159	-0.0156
	(0.0228)	(0.0237)
Control	None	State-specific trend
Observations	6,686	6,686
R-squared	0.370	0.560
F-stat 1	0.0680	0.0634
Sig. level	0.794	0.801
F-stat 2	0.222	0.883
Sig. level	0.801	0.414

Notes: Data is from the Election Commission of India. The sample period is 2005-2017. The unit of analysis is a district-year. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for years 2009 and after and 0 otherwise. All regressions include district and year fixed effects. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.10: Changes in Political Affiliation and Migration by Treatment Status

	Affiliation with ruling alliance		Migrated in last 5 years	
VARIABLES	(1)	(2)	(3)	(4)
Grant Policy x Post	-0.00743	0.00564	0.00710	0.00709
	(0.0308)	(0.0281)	(0.00680)	(0.00481)
Post	-0.0445*	-0.0517**	-0.0257***	-0.0257***
	(0.0252)	(0.0222)	(0.00606)	(0.00429)
Constant	0.387***	0.386***	0.0318***	0.0318***
	(0.00581)	(0.0133)	(0.00156)	(0.00110)
Control	District-specific trend	State-specific trend	HH FEs	District FEs
Observations	6,686	6,686	77,092	77,092
R-squared	0.725	0.558	0.532	0.045

Notes: Political affiliation data is from the Election Commission of India. The sample period is 2005-2017. The unit of analysis is a district by year. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for years 2009 and after and 0 otherwise. Columns 1 and 2 present estimated coefficients from the difference-in-difference estimates of the intervention (college construction grants) on the fraction of total constituencies in a district that are affiliated with the ruling alliance. Columns 1 and 2 control for district fixed effects. Migration data is from two rounds of IHDS survey. Columns 3 and 4 present estimated coefficients from the difference-in-difference estimates of the intervention (college construction grants) on the probability of in-migration in last five years. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.11: Effect of the Grants Policy on Out-Migration

	(1)	(2)
VARIABLES	Migration for employment	Migration for education
Grant policy x Overall trend	-3.32e-05	-0.000205**
	(0.000126)	(9.94e-05)
Grant policy x Post Policy trend	-0.000721	0.000863
	(0.00139)	(0.00319)
Grant policy x Post	0.00223	-0.00329
	(0.00430)	(0.00940)
Observations	54,504	54,504
R-squared	0.015	0.032
F-stat 1	0.303	0.0425
Sig. level	0.582	0.120
F-stat 2	0.198	2.136
Sig. level	0.821	0.837

Notes: Data is from 2012 IHDS tracking sheet data. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1994 (18 years or older in 2012–the year of the survey). The post trend goes from cohort birth years 1991 to 1994. All regressions include district and year fixed effects. No socio-economic controls are included. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.12: Spillover Analysis: Effect of the Grants Policy on Educational Attainment of Girls

	(1)	(2)	(3)	(4)	(5)
VARIABLES	No of years of schooling	Highest level of education is primary	Highest level of education is Secondary	Enrolls into college	Finishes college
Treatment x Overall trend	0.0370**	0.000933	0.00440**	-0.000543	-0.000310
	(0.0186)	(0.000924)	(0.00193)	(0.00180)	(0.00184)
Treatment x Post Policy trend	0.0882**	-0.00196	-0.0176***	0.0190***	0.0121*
	(0.0419)	(0.00190)	(0.00606)	(0.00579)	(0.00643)
Treatment x Post	0.388	0.00565	0.128***	-0.0651**	-0.0207
	(0.260)	(0.0103)	(0.0336)	(0.0273)	(0.0252)
Neighboring controls x Overall trend	0.0621***	0.00147	-0.000214	0.00369*	0.00330*
	(0.0200)	(0.000923)	(0.00202)	(0.00188)	(0.00191)
Neighboring controls x Post Policy trend	-0.0232	-0.000724	0.000748	-0.00259	-0.00595
	(0.0438)	(0.00187)	(0.00634)	(0.00612)	(0.00667)
Neighboring controls x Post	0.164	-0.000507	0.0684**	-0.0234	-0.0205
	(0.271)	(0.00998)	(0.0342)	(0.0285)	(0.0263)
Observations	110,467	110,467	110,467	110,467	110,467
R-squared	0.439	0.069	0.203	0.285	0.257
F-stat 1	11.26	0.388	6.439	14.54	5.126
Sig. level	0.000840	0.645	0.921	0.000339	0.192
F-stat 2	7.693	0.697	4.464	8.090	3.433
Sig. level	0.000500	0.498	0.0114	0.000151	0.0329
F-stat 3	0.997	0.212	0.00980	0.0452	0.239
Sig. level	0.318	0.252	0.0119	0.832	0.0239
F-stat 4	5.304	1.382	0.00796	2.482	1.656
Sig. level	0.00519	0.533	0.992	0.0844	0.625

Notes: Data is from the NFHS woman's survey. The unit of analysis is at the individual level. Treatment is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Neighboring control is a dummy variable that takes a value of 1 if the control district is a neighbor to any treatment district. The omitted category is non-neighboring control districts. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. No socio-economic controls are included. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in the augmented version of Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. F-stat 3 and F-stat 4 are the corresponding test statistics (analogous to F-stat 1 and F-stat 2, respectively) for the coefficients on the interactions of the neighboring controls dummy with the overall and the post policy trends. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.13: Effect of the Grants Policy on Husband's Education and Employment

	(1)	(2)	(3)	(4)	(5)
VARIABLES	No of years of schooling	Enrolled into college	Member of workforce	Participate in agricultural work	Participate in non-agricultural work
Panel A: All women					
Grant policy x Overall trend	-0.0255	-0.00360**	0.000872	0.00126	-0.000393
	(0.0250)	(0.00153)	(0.000975)	(0.00199)	(0.00193)
Grant policy x Post Policy trend	-0.0605	0.00695	0.0181*	-4.05e-05	0.0182
	(0.149)	(0.0105)	(0.00990)	(0.0142)	(0.0144)
Grant policy x Post	1.404*	0.0757	-0.0709*	0.00441	-0.0753
	(0.770)	(0.0508)	(0.0424)	(0.0667)	(0.0715)
Observations	6,074	6,074	6,075	6,075	6,075
R-squared	0.250	0.151	0.188	0.301	0.289
F-stat 1	0.326	0.105	3.645	0.00769	1.600
Sig. level	0.568	0.746	0.0567	0.930	0.448
F-stat 2	0.615	2.784	2.069	0.209	0.804
Sig. level	0.541	0.0626	0.127	0.812	0.206
Panel B: Women with at least a secondary education					
Grant policy x Overall trend	-0.0230	-0.00505*	-0.000426	0.00104	-0.00146
	(0.0316)	(0.00286)	(0.00147)	(0.00304)	(0.00298)
Grant policy x Post Policy trend	0.213	0.0264*	0.00881	-0.00912	0.0179
	(0.174)	(0.0159)	(0.0128)	(0.0170)	(0.0181)
Grant policy x Post	0.0546	0.0391	-0.0310	-0.0141	-0.0169
	(0.736)	(0.0747)	(0.0563)	(0.0819)	(0.0898)
Observations	3,066	3,066	3,089	3,089	3,089
R-squared	0.375	0.310	0.296	0.334	0.343
F-stat 1	1.234	1.851	0.429	0.230	0.849
Sig. level	0.267	0.0843	0.513	0.836	0.357
F-stat 2	0.881	2.486	0.264	0.180	0.548
Sig. level	0.415	0.174	0.768	0.632	0.578

Notes: Data is from the NFHS woman's survey. Information on the husband's education and employment was collected for a subset of all surveyed women. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. No socio-economic controls are included. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.14: Effect of the Grants Policy on Male Labor Force Participation

	(1)	(2)	(3)
VARIABLES	All	Urban	Rural
Panel A: With socio-economic Controls			
Grant Policy x Overall trend	0.000489	0.000652	0.000261
	(0.000385)	(0.000827)	(0.000405)
Grant Policy x Post Policy trend	-7.06e-05	-0.00282	0.00163
	(0.00229)	(0.00426)	(0.00263)
Grant Policy x Post	0.0352**	0.0348	0.0303*
	(0.0152)	(0.0324)	(0.0176)
Observations	72,934	18,125	54,809
R-squared	0.371	0.405	0.372
F-stat 1	0.0359	0.270	0.553
Sig. level	0.850	0.643	0.575
F-stat 2	0.878	0.442	0.554
Sig. level	0.416	0.603	0.457
Panel B: Without socio-economic controls			
Grant policy x Overall trend	0.000624	0.000902	0.000288
	(0.000381)	(0.000818)	(0.000405)
Grant policy x Post Policy trend	8.33e-05	-0.00283	0.00166
	(0.00231)	(0.00434)	(0.00264)
Grant policy x Post	0.0318**	0.0289	0.0300*
	(0.0152)	(0.0328)	(0.0176)
Observations	73,255	18,207	55,048
R-squared	0.359	0.392	0.364
F-stat 1	0.101	0.206	0.579
Sig. level	0.750	0.650	0.447
F-stat 2	1.500	0.700	0.628
Sig. level	0.224	0.497	0.534

Notes: Data is from the NFHS man's survey. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. Socio-economic controls include household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.15: Effect of the Grants Policy on Fertility and Child Preference

VARIABLES	(1) Total children ever born	(2) Number of living children	(3) Ideal number of children	(4) Son preference	(5) Knows about modern contraceptive method	(6) Uses modern contraceptive method
Panel A: With socio-economic controls						
Grant Policy x Overall trend	0.00331	0.00420	-0.00254	-0.000183	-7.53e-05	0.000376
	(0.00517)	(0.00482)	(0.00250)	(0.000731)	(0.000410)	(0.00112)
Grant Policy x Post Policy trend	-0.00514	-0.00133	0.00228	0.00282**	0.000778	0.00490
	(0.0153)	(0.0141)	(0.00403)	(0.00142)	(0.000849)	(0.00563)
Grant Policy x Post	-0.144**	-0.141**	-0.0334	-0.0177**	-0.00418	0.00480
	(0.0701)	(0.0674)	(0.0271)	(0.00864)	(0.00465)	(0.0264)
Observations	34,482	34,482	108,855	108,673	110,467	34,457
R-squared	0.415	0.415	0.325	0.069	0.101	0.231
F-stat 1	0.0170	0.0489	0.00637	4.841	0.853	0.864
Sig. level	0.896	0.662	0.596	0.0281	0.356	0.636
F-stat 2	0.208	0.413	0.518	2.441	0.451	0.453
Sig. level	0.812	0.825	0.936	0.0879	0.637	0.353
Panel B: Without socio-economic controls						
Grant policy x Overall trend	0.00130	0.00235	-0.00249	-0.000181	-8.67e-05	0.000358
	(0.00560)	(0.00526)	(0.00254)	(0.000736)	(0.000413)	(0.00111)
Grant policy x Post Policy trend	-0.00478	-0.00173	0.00264	0.00287**	0.000763	0.00493
	(0.0156)	(0.0142)	(0.00420)	(0.00145)	(0.000848)	(0.00562)
Grant policy x Post	-0.174**	-0.169**	-0.0373	-0.0183**	-0.00374	0.00340
	(0.0712)	(0.0688)	(0.0277)	(0.00875)	(0.00468)	(0.0266)
Observations	34,482	34,482	108,855	108,673	110,467	34,457
R-squared	0.327	0.322	0.312	0.064	0.095	0.229
F-stat 1	0.0622	0.00235	0.00202	4.873	0.788	0.864
Sig. level	0.803	0.902	0.618	0.0276	0.375	0.353
F-stat 2	0.0522	0.103	0.481	2.451	0.426	0.448
Sig. level	0.949	0.961	0.964	0.0871	0.653	0.639

Notes: Data is from the NHFS woman's survey. Columns (1), (2), and (6) are based on the ever-married women sample, whereas Columns (3), (4) and (5) are based on the whole sample. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. Socio-economic controls include household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.16: Effect of the Grants Policy on Educational Attainment of Girls

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	No of years of schooling	Highest level of education is primary	Highest level of education is Secondary	Enrolls into college	Finishes college	Number of years in college
Panel A: With socio-economic controls						
Treatment x Overall trend	-0.0157***	0.00204***	0.00137**	-0.00200***	-0.00187***	-0.00703***
	(0.00543)	(0.000364)	(0.000536)	(0.000353)	(0.000324)	(0.00130)
Treatment x Post Policy trend	0.130***	-0.00490***	-0.00239	0.0111***	0.0128***	0.0448***
	(0.0192)	(0.000967)	(0.00168)	(0.00157)	(0.00179)	(0.00647)
Treatment x Post	-0.00573	0.00198	0.0379***	-0.0261***	-0.0170***	-0.0676***
	(0.0707)	(0.00447)	(0.00864)	(0.00714)	(0.00655)	(0.0244)
Observations	540,306	540,306	540,306	540,306	540,306	540,306
F-stat 1	43.36	10.05	0.445	44.59	49.35	47.36
Sig. level	9.49e-11	1.54e-09	0.0368	0	0	0
F-stat 2	22.82	20.95	3.320	25.94	25.49	24.04
Sig. level	2.68e-10	0.00159	0.505	5.28e-11	0	8.61e-11
Panel B: Without socio-economic controls						
Treatment x Overall trend	-0.0176***	0.00204***	0.00114**	-0.00198***	-0.00186***	-0.00694***
	(0.00583)	(0.000366)	(0.000547)	(0.000361)	(0.000331)	(0.00133)
Treatment x Post Policy trend	0.143***	-0.00504***	-0.00186	0.0116***	0.0132***	0.0467***
	(0.0201)	(0.000977)	(0.00170)	(0.00158)	(0.00181)	(0.00657)
Treatment x Post	0.00918	0.00227	0.0398***	-0.0264***	-0.0172**	-0.0687***
	(0.0791)	(0.00451)	(0.00876)	(0.00743)	(0.00683)	(0.0254)
Observations	540,306	540,306	540,306	540,306	540,306	540,306
F-stat 1	50.17	10.99	0.220	49.44	52.71	50.89
Sig. level	0	0.000970	0.112	0	0	0
F-stat 2	25.64	21.02	2.193	27.12	26.65	25.48
Sig. level	0	1.44e-09	0.639	0	0	0

Notes: Data is from the NFHS woman's survey. The unit of analysis is at the individual level. **The sample consists of women who did not always stay in the same place.** Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. The variable *number of years in college* is a categorical variable which takes the value of zero if an individual does not attend college, and values between one and five for individuals attending college corresponding to the year they are in college (1, 2, and 3 for the first, second, and third year of college, respectively; 4 for year one in Masters and 5 for year two in Masters or beyond). All regressions include district and year fixed effects. Socio-economic controls include household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.