DISCUSSION PAPER SERIES

DP16321

Centrality-Based Spillover Effects

Asadul Islam, Michael Vlassopoulos, Yves Zenou and Xin Zhang

DEVELOPMENT ECONOMICS

LABOUR ECONOMICS

ISSN 0265-8003

Centrality-Based Spillover Effects

Asadul Islam, Michael Vlassopoulos, Yves Zenou and Xin Zhang

Discussion Paper DP16321
Published 02 July 2021
Submitted 30 June 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Centrality-Based Spillover Effects

Abstract

We study the role of social network structure in peer-to-peer educational spillovers by leveraging a two-year field experiment in primary schools in rural Bangladesh. We implement a randomized educational intervention-the provision of free after-school tutoring-offered to a random subsample of students in treatment schools. We exploit the experimentally induced across-classroom variation in the centrality of treated students to provide the first causal evidence of centrality-based spillover effects. We find that a one standard deviation (SD) increase in the average centrality of treated students within a classroom leads to improvements in the test scores of their untreated classmates of 0.57 SD in English and 0.62 SD in math. Further evidence indicates that more central students have higher academic ability, better social skills, and interact more with classmates on matters related to learning, which suggests that they can be more influential for their peers. In addition, we offer the private tutoring intervention to the most central students on a separate group of schools. We find that this targeted approach provides larger educational benefits both for treated and untreated students than the policy that treats a random subset of students. We conclude that targeting the most central students in a network to offer an intervention can be a cost-effective way to improve the educational outcomes of all students in a classroom.

JEL Classification: I21, O15, Z13

Keywords: Network centrality, Spillovers, Education

Asadul Islam - asadul.islam@monash.edu
Centre for Development Economics and Sustainability (CDES), Department of Economics, Monash University, Australia

Michael Vlassopoulos - m.vlassopoulos@soton.ac.uk
University of Southampton

Yves Zenou - yves.zenou@monash.edu
Monash University and CEPR

Xin Zhang - xin.zhang@monash.edu
Monash University

Acknowledgements
We thank Arthur Campbell and Claudio Labanca for their helpful comments. We also thank BRAC in Bangladesh for conducting fieldwork, training, and surveys. We acknowledge funding support from the Australian Development Research Award Scheme (ADRAS) from AusAID (now DFAT) and the Faculty of Business and Economics at Monash University. We also thank the Global Development and Research initiative (GDRI), BRAC Education program, and the Department of Primary Education (DPE) of the Government of Bangladesh for administrative support and implementation of the intervention. We thank Firoz Ahmed, Jenarul

Islam, Tanvir Mojumder, and Sukanto Roy for excellent research assistance.

Centrality-Based Spillover Effects

Asad Islam[*] Michael Vlassopoulos[†] Yves Zenou[‡] Xin Zhang[§]

July 1, 2021

Abstract

We study the role of social network structure in peer-to-peer educational spillovers by leveraging a two-year field experiment in primary schools in rural Bangladesh. We implement a randomized educational intervention–the provision of free after-school tutoring–offered to a random subsample of students in treatment schools. We exploit the experimentally induced across-classroom variation in the centrality of treated students to provide the first causal evidence of centrality-based spillover effects. We find that a one standard deviation (SD) increase in the average centrality of treated students within a classroom leads to improvements in the test scores of their untreated classmates of 0.57 SD in English and 0.62 SD in math. Further evidence indicates that more central students have higher academic ability, better social skills, and interact more with classmates on matters related to learning, which suggests that they can be more influential for their peers. In addition, we offer the private tutoring intervention to the most central students on a separate group of schools. We find that this targeted approach provides larger educational benefits both for treated and untreated students than the policy that treats a random subset of students. We conclude that targeting the most central students in a network to offer an intervention can be a cost-effective way to improve the educational outcomes of all students in a classroom.

Keywords: Network centrality, spillovers in education, private tutoring, Bangladesh, field experiment.

JEL classification codes: I21, O15, Z13.

[*]Monash University, Australia. E-mail: asadul.islam@monash.edu.

[†]University of Southampton, UK, and IZA. Email: m.vlassopoulos@soton.ac.uk.

[‡]Monash University, Australia, CEPR, and IZA. Email: yves.zenou@monash.edu.

[§]Monash University, Australia. E-mail: xin.zhang@monash.edu.

[1]We thank Arthur Campbell and Claudio Labanca for their helpful comments. We also thank BRAC in Bangladesh for conducting fieldwork, training, and surveys. We acknowledge funding support from the Australian Development Research Award Scheme (ADRAS) from AusAID (now DFAT) and the Faculty of Business and Economics at Monash University. We also thank the Global Development and Research initiative (GDRI), BRAC Education program, and the Department of Primary Education (DPE) of the Government of Bangladesh for administrative support and implementation of the intervention. We thank Firoz Ahmed, Jenarul Islam, Tanvir Mojumder, and Sukanto Roy for excellent research assistance.

1 Introduction

Educational spillovers across students are recognized as key inputs to the education production function, alongside the important contribution of teachers and parents. Indeed, a large body of literature in economics and, more broadly, in the social sciences has demonstrated that various dimensions of heterogeneity in classroom composition, such as, gender, race, and ability, influence the educational outcomes of students (Sacerdote, 2011, 2014; Epple & Romano, 2011). Overlaying classroom makeup, students are embedded in networks of social relationships with their classmates with various degrees of connectedness, which would appear to be an important channel for the operation of spillover effects across students. For example, one might expect that between-classmate spillovers would be weaker in a classroom where students are socially isolated than in one in which many strong ties exist between classmates that facilitate the transmission of knowledge. Surprisingly, the role that the network structure in which students are embedded in plays in the strength of educational spillovers has been largely neglected in the literature examining peer effects in education. This is an important omission because understanding how network structure impacts knowledge spillovers in educational settings and beyond can inform the design of interventions targeting influential actors to achieve a more cost-effective diffusion of knowledge.

In this paper, we investigate whether students who are more central in the friendship network of their classroom generate stronger spillover effects on their classmates than those who are less central. To address this question, we carry out a large field experiment with primary school students in rural Bangladesh, specifically designed to overcome the well-known empirical challenges associated with identifying spillover effects across peers (Manski, 1993). In particular, in our experimental design schools are first randomized into treatment and control schools, and then treatment schools are randomized into one of three treatment arms. In the first treatment arm (T1), students are further randomized into receiving free after-school private tutoring in groups three times per week during the course of two academic years following a partial population design (Moffitt, 2001). The rest of the children in the treatment schools do not receive our intervention. For schools in the second treatment arm (T2), the intervention also includes additional one-to-one tutoring sessions at home once a week. In other words, within schools, in these two treatment arms, a random subset of students receives our intervention, with the share of treated students varying across treatment schools.[1]

We also collect information on the friendship network of the students before the start of the intervention and are thus able to attach a measure of network centrality to each student. We use this information to also deliver the intervention in a targeted way to schools in the third treatment arm (T3). In particular, in these schools, we offer private tutoring to the most central students in the class (as assessed at baseline). That is, the third treatment arm resembles the first, except that treated students are not randomly picked but are instead selected on the basis of their eigenvector

[1] Other supplementary educational interventions have been considered in developing countries, such as providing remedial teaching support to children (Banerjee et al., 2007), training volunteers to hold after school reading camps (Banerjee et al., 2010), or offering computer-aided after school customized instruction (Muralidharan et al., 2019).

centrality. This allows us to evaluate how spillover effects of the random intervention treatment compare to this alternative approach of targeting the most connected students.

Our outcomes include measures of students' learning in English and math assessed on multiple occasions: before the intervention and continuing into the middle and end of the two-year program. A second group of outcomes that we consider are measures of noncognitive skills of the students measured at the endline. These skills encompass three key domains: self-control, social skills, and motivation.

Our sample consists of over 14,000 primary school students (in year 3 and 4) across 254 schools in rural Bangladesh. This is a suitable context to carry out our study for two main reasons. First, despite considerable progress in educational attainment, especially in primary schools, raising educational standards remains an important challenge in developing countries such as Bangladesh (Glewwe & Muralidharan, 2016; World Bank, 2017). Second, many developing countries operate under tight public education spending budgets, making it difficult to support educational interventions on a large scale; thus, the phasing-in of an educational program aimed at raising student learning with partial coverage of the student population, such as our intervention, can be a relevant policy option.

We find that the intervention generates significant improvements in the educational outcomes of treated students in schools where private tutoring is provided relative to students in control schools who receive no intervention. The improvements are present both when the treatment is offered to a random subgroup of students in each class and in the targeted case where the treatment is offered to a subgroup of the most central students. Comparison of schools in which tutoring is offered randomly (T1) to those where it is offered in a targeted way (T3) shows that the latter generates larger benefits in the learning of treated children: 0.75 SD as compared to 0.38 SD in English, and 0.78 SD as compared to 0.48 SD in math.

More importantly for our purposes, our design also allows us to estimate spillover effects, which is the main objective of this paper. First, we examine the overall spillover effect of the intervention by comparing the test scores of untreated students in treated schools to students in control schools where the intervention did not take place. The overall spillover effects after two years are large and also much stronger–more than double–in the targeted intervention than in the random one: 0.59 SD as compared to 0.25 SD in English, and 0.68 SD as compared to 0.30 SD in math.

We then exploit the fact that, in schools where the intervention was offered randomly, there is exogenous variation across classrooms in the degree of network centrality of treated students for the causal identification of what we refer to as *centrality-based* spillover effects of the intervention. We document that centrality spillovers are present and are large: we estimate that a one SD increase in the average centrality of treated students within a class leads to improvements in the test scores of their untreated classmates by 0.57 SD in English and 0.62 SD in math. Further analysis indicates that variation in the average academic ability of treated students (measured by their test scores at baseline) does not fully account for the centrality spillover effects, which suggests that centrality spillovers do not simply capture variation in the ability of treated students.

To the best of our knowledge, this is the first study to provide causal estimates of *centrality-based* spillover effects within the context of an educational setting, which constitutes our main contribution. In addition, we explore why centrality spillovers are important in this context. To address this issue, we investigate the characteristics of the central students. We find that the centrality of a student at baseline is positively associated with test scores in both English and math. We also find that more central students have better social skills and tend to interact more with classmates on matters related to schoolwork. We argue that the fact that more central students are good students, have strong social skills, and engage more intensely with their classmates in their learning is a possible explanation for why they are able to generate more knowledge spillovers to their classmates.

After establishing the presence of centrality-based spillover effects, we turn our attention to an evaluation of whether a targeted intervention is preferable to one carried out randomly in terms of the educational achievements of both treated and untreated students. For example, one might be concerned that with our targeted intervention weaker students might be neglected, and there might be an increase in inequality between stronger and weaker students.

To aid interpretation and understanding of the trade-off associated with the two types of interventions, we develop a theoretical model that builds on the framework of Ballester et al. (2006) but allows the possibility that some agents in the network are treated while others are not, as is the case in our field experiment. First, the model shows that an individual's action (in our context, educational outcome) is proportional to measures of network centrality of that individual, which is consistent with our evidence. Second, the model highlights a possible trade-off between the direct and spillover effects generated by the targeted intervention. That is, moving from random treatment to targeted treatment could lead to some losses in total direct benefits (if randomly treated students benefit more from the intervention than the most central ones), but the spillover effects are larger under the targeted treatment because the students receiving the treatment are more connected and thus generate larger externalities to their classmates. Consequently, in aggregate, the benefits are likely to be larger under the targeted approach. In our experiment, we find that the targeted intervention improves average outcomes for both treated and untreated students relative to the random intervention while showing no indication of more dispersion or weaker students falling behind. Moreover, we show that the targeted policy has a stronger *contagion effect* achieving a more widespread *diffusion* of the benefits of the intervention among untreated students. In other words, by providing treatment to central students, we show that untreated students benefit more because there are *more students* who improve their test scores and they attain *higher test scores* than the untreated students in classrooms where treatment was given to students at random. This leads us to conclude that targeting interventions to central players of a social network can be an effective and relatively cheap way to capitalize on the knowledge spillovers they generate on their classmates.

Related literature Our paper contributes to several strands of the literature.

Network centrality Centrality is a core concept in the study of networks (Wasserman & Faust, 1994; Jackson, 2008, 2019).[2] Measures of centrality help characterize myriad features of network models, ranging from which individuals we should target to facilitate the spread of innovations (Banerjee et al., 2013), which agents have the largest influence over others' beliefs (Golub & Jackson, 2010), and which players choose higher equilibrium actions in various network games (Ballester et al., 2006; Galeotti et al., 2010; Bochet et al., 2020).

Some studies have addressed the question of which measure of centrality is appropriate to predict which behavior. For example, in cases in which there are strong complementarities in behaviors, such as in crime, education, or R&D collaborations, Katz-Bonacich centrality has proven useful in describing the activity of each agent (Lindquist & Zenou, 2014; Calvó-Armengol et al., 2009; Battaglini & Patacchini, 2018; Lee et al., 2021). In contrast, when studying the diffusion of information, Katz-Bonacich centrality is not always a strong predictor of which people are the most influential seeds for the process, and other centrality measures outperform it. Indeed, in investigating microfinance diffusion in 43 different villages in India, Banerjee et al. (2013) find that the eigenvector centrality and diffusion centrality of the first contacted individuals (i.e., the set of original injection points in a village) are the only significant predictors of the eventual diffusion.

An earlier study that is closely related to ours is that of Calvó-Armengol et al. (2009), who, using data from a survey of high school students in the United States (AddHealth), show that the network position of students (measured by their Katz-Bonacich centrality) has a *direct* positive impact on their academic achievement.[3] We complement this previous study by providing *causal* evidence of the role of a student's network centrality on the strength of the *spillover effects* that they can exert on their classmates.

A few other recent studies examine the role of network centrality of agents in various applications. Breza & Chandrasekhar (2019) carry out a savings field experiment in rural India in which they randomly assign savers to monitors and find that more central monitors lead to larger increases in savings. Chandrasekhar et al. (2018) find that, in the absence of contract enforcement, individuals tend to behave more cooperatively when interacting with more central partners. In a related study, Breza et al. (2015) investigate whether efficiency in a trust game can improve under the presence of a monitoring third party and whether the effectiveness of the monitor increases with her network centrality (measured by eigenvector centrality). Beaman et al. (2021) find that targeting central farmers leads to faster adoption of a new agricultural technology in Malawi. Finally, Mohnen (2021) studies the role of network position in knowledge spillovers among co-authoring scientists, exploiting the sudden death of a co-author as a shock to the network. We contribute to this line of research by investigating the role of network centrality in spillovers in an educational setting.

[2]For overviews of the network literature, see Jackson (2008), Ioannides (2013), and Jackson et al. (2017).

[3]See, also, Jain & Langer (2019), who examine how the students' network size, distance, prestige, and connections to influential individuals impact their academic performance. They find that increasing closeness centrality within the network negatively affects student performance measured by grade point average.

Spillover effects in education

Our paper contributes to research measuring spillover effects in educational settings, often by taking advantage of randomized interventions. List et al. (2020) document the neighborhood spillover effects of an early childhood intervention that aimed at improving the educational outcomes of a sample of disadvantaged children in the United States. Bennett & Bergman (2021) find that a parent information intervention on student absences has spillover effects across the students' social network. Avvisati et al. (2014) find that a program of parent–school meetings in France aimed at enhancing parental involvement in their children's education not only improved the behavior of pupils of participating families but also had a spillover effect on children whose parents did not participate in the program. Abramitzky et al. (2021) find spillover effects of a reform in Israel that affected families that lived in kibbutzim communities and led to improved academic performance of affected students compared to their high-school peers whose families were not affected by the reform.[4]

These papers document spillovers in education, but none have investigated how the network position of a student affects the strength of the spillovers they can exert on their peers, which is the main contribution of this paper.

The remainder of this paper is organized as follows. Section 2 provides the context of our study and our experimental design. Section 3 describes the data and discusses sample construction. Section 4 presents results on the direct effect of the intervention, while section 5 presents results on the spillover effects and explores the underlying mechanisms. Section 6 offers a comparison between the random provision of the intervention and the one that targets more central students. The last section offers some concluding remarks.

2 Context and experimental design

2.1 Education in Bangladesh

The education system in Bangladesh is broadly divided into three major levels: primary, secondary, and tertiary education. Primary education, defined as basic education, is compulsory for children aged six to ten years (grades 1–5). Article 17 of the Bangladesh Constitution states that all children between the ages of six and ten are to be provided with free basic education. The Primary Education (Compulsory) Act was introduced in 1990 with the objective of realizing universal primary enrolment.

Bangladesh, like many other countries in South Asia, has traditionally been characterized by low school enrollment. In 1990, the net enrollment in primary schools was 69% for girls and 75% for boys.[5] However, the country has made tremendous progress in recent decades in improving educational attainment. Major achievements of the last decade in primary education include gender parity

[4] There is also a very large literature on peer effects in education (see the overviews by Epple & Romano (2011); Sacerdote (2011, 2014)), however, to the best of our knowledge, none of these papers investigate the causal effect of individual *centrality* on the education outcomes of peers.

[5] These statistics were retrieved from https://data.worldbank.org/indicator/SE.PRM.NENR?locations=BD.

in enrollment, improvement in gross and net enrollment, and a reduction in dropout and improvement in the completion of the cycle.[6] National education budgets have been increased substantially, particularly in the area of primary education. At the primary level, the target teacher–student ratio is 40 students per teacher, and in 2018 about 55% of schools met that target. The target of offering free textbooks in the first month of the school year to all students has also been achieved.

However, in terms of classroom learning, the perception is that the level is low, particularly in rural areas (Asadullah & Chaudhury, 2013). This is due to the low level of effective teaching hours, high levels of student and teacher absenteeism, and low teaching input. Children from poor families are particularly disadvantaged, as they are less likely to have educated parents who could help them in their studies at home or afford study materials.

2.2 The intervention

Our study was carried out in collaboration with a local non-governmental organization (NGO), named Global Development Research Initiative (GDRI). The study took place in two districts, Satkhira and Khulna, which are fairly typical of many parts in rural Bangladesh. Figure A1 in the Online Appendix A.1 presents the location of these two districts in Bangladesh.

Private tutoring refers to the supplementary teaching provided by the private sector, in addition to the provision of education from the mainstream schooling system. In Bangladesh, families are increasingly employing private tutors to help their children with school work, making it a good context to introduce private tutoring intervention.

Our main intervention involved offering free private tutoring outside school hours to treated students over the course of two academic years in order to improve their educational outcomes. Three days a week, private tutors provided two hours of supplementary teaching to grade 3 and 4 students, teaching mathematics and English. Grade 3 and 4 students were chosen because it is the best period to provide extra education to these students as a foundation for their further studies. English and math are the subjects that have the highest failure rates and were considered to be important subjects in which students need extra help after school. Indeed, English and math are the subjects that are a strong predictor of whether a student drops out of school or continues on to the next grade.

Private tutors were hired by the NGO and were attached to a school to familiarize themselves with the school's activities and receive feedback from school teachers about the academic progress of students in school. Before the tutors started to deliver tutoring, they were trained for one week. Throughout the two-year experiment, the private tutors were also trained for three days every two months during the school year to make sure that all of them were following the school curriculum while identifying and sharing the students' concerns and problems and solving them with the help of educators. In contrast to school teachers, private tutors were locals or came from nearby villages and were recent college graduates with a bachelor degree, but they were also currently unemployed.

[6] In 2018, the net enrollment rate, attendance rate, and average dropout rate at the primary level stood at 97.8%, 88.6%, and 18.6%, respectively (Government of Bangladesh, 2019).

Hence, they had an incentive to teach well to improve their employment prospects. Because they were not professionally qualified, they were paid at a fraction of the wage of a regular school teacher.

In terms of the private tutoring, it was conducted on a group basis, so treated students who received private tutoring studied together in a group of about 10–12 students. These private tutoring sessions took place *after school*. Each week, private tutors provided three sessions, each lasting two hours.

In addition, in some schools students also received a *home visit* by the tutor, which involved a one-day tutoring in the students' home. The private tutors met the students' parents to keep them informed about their children's progress in school. The tutor also provided parents with information about how to take care of children properly, such as preparing meals and creating a good study environment. The home visit lasted for one hour. The purpose of the home visit for tutoring was to make sure that there were more interactions with parents about the child's progress.

To summarize, two types of interventions were conducted among treated schools: (i) Private tutoring by a tutor in a group of about 10–12 students for three sessions of two hours per week after school; (ii) Private tutoring plus a home visit one day a week for one hour to individually tutor students and parents at home. Note that both interventions lasted for *two years*—the duration of the field experiment.

2.3 Experimental design

Treatments Our experimental design involves randomly allocating participating schools into one of the following 4 groups:

1. Treatment 1 (T1): Private tutoring is provided to a random sample of students.
2. Treatment 2 (T2): Private tutoring and home visits are provided to a random sample of students.
3. Treatment 3 (T3): Private tutoring is provided to a sample of students with the highest eigenvector-centrality.
4. Control: No intervention is provided.

The main difference between T1, T2, and T3 that we focus on in this paper is that in T1 and T2 treatment was assigned randomly, whereas in T3 treatment was provided to the most central students. Our main interest in this paper is evaluating whether the spillover effects from these two different ways of delivering an intervention differ. Along the way, we will also report the direct effect of the various treatments; however, this aspect is not our main focus in this paper.

Randomization We use a two-stage randomization design, where we first assign each school to one of the four experimental groups described above and then assign students within treatment schools (T1, T2, and T3) to receive treatment. In total, 254 primary (public) schools participated in our study, randomly selected from two rural districts in Bangladesh, Satkhira, and Khulna. These 254 primary schools were randomly allocated into treatment and control groups, resulting in 100 schools allocated to the control group and the other 154 schools allocated to one of the treatment groups: 58 schools were randomly assigned to Treatment 1 (T1: private tutoring), 58 schools were randomly assigned to Treatment 2 (T2: private tutoring and home visits), and 38 schools were randomly assigned to Treatment 3 (T3: centrality-based private tutoring). Figure A2 in the Online Appendix A.10 illustrates the allocation of schools into control and treatment schools.

Within T1 and T2 schools, a subgroup of students were randomly chosen to receive the treatment. In the 38 T3 schools, treated students were *not* chosen at random; rather, selection was based on their eigenvector centrality. In fact, in each school, only students with the highest centrality were treated, and these students only received private tutoring, i.e., the same intervention as treated students in T1. Table 1 provides the sample sizes of each group for the schools in our sample. Overall, in T1 and T2, about 80% of students are treated, and there is variation in the share of treated students across schools. This is because, in each treated school, we aimed to provide the treatment to about 20–24 students per grade, split into two tutoring groups. Therefore, variation in the classroom size of the school creates variation in the share of treated students. In T3, the share of treated students is smaller at 49% because we wanted to ensure that we have a sufficiently large number of untreated students in this treatment arm to be able to estimate the spillover effect on them.

Table 1: Overview of Experimental Design

	T1	T2	T3	Control	Total
Treated students	2,647	2,358	1,307	0	6,312
Untreated students	650	613	1,370	5,271	7,904
Total	3,297	2,971	2,677	5,217	14,216
Treatment Saturation	80.3%	79.4%	48.8%		
No. of schools	58	58	38	100	254

Timeline Figure A3 in the Online Appendix A.1 displays the timeline of our experiment, which took place between December 2013 ($t = 0$) and December 2015 ($t = 2$). Within this two-year time frame, we surveyed participating students on two occasions ($t = 0, 2$) and administered exams (test scores) *three times* ($t = 0, 1, 2$). First, in December 2013 ($t = 0$), before our intervention, a baseline survey on household characteristics as well as on students' *network* was conducted among all grade 2 and 3 students in participating schools. The baseline survey contains information on students as well as their parents' pre-experiment characteristics and network relationships. Also, at baseline, a pre-experiment test was implemented to students in grades 2 and 3 (students in grade 2 are subject to treatment in 2014 when they reach grade 3). Each student was tested in English and math.

Our interventions started in March 2014 (students are busy with extra-curricular activities in January and February), so students were now in grades 3 and 4. As stated above, our experiment lasted for two years from January 2014 until December 2015. The same set of students were observed and tested over these two years, resulting in a two-year panel data set. One year after the intervention, in December 2014 ($t = 1$), we conducted a mid-line test of the participants. It included a second English and math test. Finally, two years after the start of the intervention, in December 2015 ($t = 2$), the experiment finished, and an endline survey was implemented. The endline survey contained a survey on household characteristics as well as on students' *network*. A third exam in English and math was also implemented to collect the students' test scores two years after the intervention started. Moreover, in December 2015 ($t = 2$), we collected information on the non-cognitive skills of students.

Let us explain how the network information was collected at the baseline (December 2013) and the endline (December 2015). All students in a grade were asked to nominate and rank ten of their closest friends among their classmates. To be more precise, we asked each student to pick their ten best friends and rank them from 1 (very close friends) to 10 (least close friends). It is possible to map out the students' entire social network because it is fairly typical in Bangladesh for an entire grade to be placed in one classroom.

3 Data

3.1 Outcome variables

Test scores Our first set of outcome variables measures students' learning through the standardized tests in *English* and *math* that we administered. In the analysis, we normalize these test scores for each student by subtracting the average test score of the control group in the respective year and dividing by the SD. The tests were based on materials drawn from the relevant textbooks. Separate tests were conducted for students in each grade in each year. Program staff administered the tests in the classrooms at the schools. The test items consist of both multiple-choice and short questions. The test is intended to assess problem-solving capacities in mathematics (e.g., geometric skills and complexly worded problems) and English comprehension requiring students verbal, language, and reasoning skills. Local school teachers and educators were consulted to ensure that the tests are appropriate for the grade level.

Noncognitive skills Our second group of outcome variables consist of measures of the noncognitive skills of the students. To obtain these measures, we designed several questions to explore three main domains of noncognitive skills of each student: *self-control*, *social skills*, and *motivation*. Teachers and not the private tutors provided an evaluation of each student in these dimensions. In Table A11 in Appendix A.7, we provide the underlying questions used to construct each measure. For each skill, we aggregate the corresponding questions into one variable by calculating the average across the responses (all responses are numerical), so all measures take values in the range $[0, 1]$. We then normalize the measures by substracting the mean and then dividing by the SD of the control group. We also construct an *Index*, which is the average of these three measures. Figure A4 presents the distribution of our four measures of noncognitive skills (before normalization).

3.2 Descriptive statistics and balance checks

As stated above, the baseline survey was conducted in December of 2013 ($t = 0$). According to Table A1 in the Online Appendix A.2, the average age of a household head is around 39, and roughly 60% of them can read and write. Moreover, the average monthly income of a household is roughly 6,000 takas, and 50% of the students are male.

We perform a series of balance checks reported in Appendix A.2. Table A1 compares students in treatment schools (T1, T2, and T3) with students in the control schools along their individual and household characteristics and finds that, indeed, there are no statistical differences in any of them. In Table A2, we report a second balance check in which we compare within treatment schools (T1, T2, and T3) students who were assigned to receive treatment to those who were not. Again, we do not find significant differences in terms of their observable characteristics. Finally, in Table A3 we assess the balance between treatment and control schools in terms of head teacher characteristics, such as their education, age, and gender, and other school-level characteristics, such as class size, number of teachers, and facilities. There is no indication of systematic differences across the treatment arms in any of these characteristics.

3.3 Missing test scores

As is common in studies that track student outcomes over time, there are cases of students with missing follow-up test scores. The main reason for a missing test score in our sample is that a student may have been absent from school on the day that we administered the test.

To start with, 17,102 students took the baseline test in English and math that we administered at $t = 0$. Of those, 3,723 (21.8%) have missing test scores at $t = 1$, and 2,886 (16.9%) have missing test scores at $t = 2$. In our baseline sample and analysis, we include the 14,216 students that have test scores at both baseline and endline. The fraction of students with missing endline test scores is slightly lower among control schools (15.8%) than treatment schools (18.6%). However, among the students with missing endline scores, we find no significant differences in baseline test scores in English and math across control and treatment schools, suggesting that the composition of students who miss endline test scores (in terms of baseline educational achievement) is similar across treatment and control schools. Tables A4 and A5 in Appendix A.3 report these results, which suggest that differential missingness of endline test scores by treatment status is not a concern in our analysis. Nevertheless, in Appendix A.6, we provide bounds for the impact that missing test scores have on our main results (overall treatment effect in Table A9 and spillover effect in Table A10) by applying the procedure suggested in Lee (2009). We find them to be robust.

3.4 Take-up of Treatment

Treatment take-up in our study was high (97%). Table A6 in Appendix A.4 reports the fraction of students who enrolled in the private tutoring sessions out of those who were randomly assigned to receive a treatment (or were selected on the basis of their centrality in the case of T3) for each treatment arm separately. Take-up was slightly higher in T2 than in the other treatment groups. We focus on presenting results on the basis of treatment assignment (intention to treat effects), though, given the overall high rates of treatment compliance, our estimates can essentially be interpreted as average treatment effects.

4 Direct treatment effects

In this section, we evaluate the direct effect of our interventions. This is defined as the impact of the intervention on treated students within treatment schools.

Our main estimating equation is as follows:

$$y_{i,g,s,t} = \alpha_0 + \beta_1 T1 + \beta_2 T2 + \beta_3 T3 + \gamma\, y_{i,g,s,0} + X'_{i,g,s}\delta + Grade_{i,g,s,0} + \epsilon_{i,g,s,t} \tag{1}$$

where $y_{i,g,s,t}$ is the outcome of interest (test score or measure of noncognitive skill) of student i, in grade g, school s, and time t; $T1, T2$, and $T3$ are indicator variables for a school receiving the private tutoring treatment, the private tutoring plus home visit treatment, and the centrality-based private tutoring treatment, respectively; $y_{i,g,s,0}$ is the baseline outcome measured at $t = 0$ (December 2013) when available (test scores); $X_{i,g,s}$ denotes student and household characteristics, including gender, father's age, literacy and occupation, household income, and the number of children in the household.[7] $Grade$ is an indicator of being in the senior grade we examine, and $\epsilon_{i,g,s,t}$ is an error term. In all our regressions, standard errors are clustered at the school level, as treatment is determined at that level.

Equation (1) determines the *direct* impact of treatment (private tutoring for $T1-$schools, private tutoring and home visits for $T2-$schools, private tutoring provided to the most central students for $T3-$schools) on test scores and measures of noncognitive skills of treated students. We estimate (1) on a sample that consists of treated students in treatment schools and all students in control schools. We expect β_1, β_2, and β_3 to be positive. Further, we expect β_2 to be larger in magnitude than β_1, as students in T2 receive additional support (home visits) relative to those in T1. We might also expect β_3 to be larger than β_1 because the composition of treated students in T3 is different from that in T2, that is, treated students in T3 are more central than those in T1 and have better baseline test scores. We cannot ex ante compare β_2 and β_3 since the treatment as well as the students receiving treatment were different in terms of their network characteristics.

[7]For observations where data on the control variables are missing, we imputed missing values and added an indicator variable denoting this in the specifications that we estimate.

4.1 Test scores

We first present results of the estimation of equation (1) for test scores. Table 2 displays these results.[8] Several interesting findings emerge.

Table 2: Direct effect on test scores

	English			Math		
	2013	2014	2015	2013	2014	2015
	(1)	(2)	(3)	(4)	(5)	(6)
T_1	-0.039	0.335***	0.382***	-0.027	0.496***	0.481***
	(0.065)	(0.077)	(0.082)	(0.064)	(0.082)	(0.068)
T_2	0.00162	0.506***	0.901***	-0.0337	0.722***	0.965***
	(0.063)	(0.071)	(0.088)	(0.069)	(0.073)	(0.070)
T_3	0.0490	0.456***	0.706***	0.016	0.710***	0.783***
	(0.072)	(0.082)	(0.094)	(0.081)	(0.085)	(0.077)
$y_{i,g,s,0}$		0.271***	0.211***		0.242***	0.218***
		(0.015)	(0.017)		(0.018)	(0.017)
Observations	11,583	10,986	11,583	11,583	10,986	11,583
No. of schools	254	251	254	254	251	254
H_0: $T_3 = T_1$	0.274	0.083	0.000	0.631	0.013	0.000
H_0: $T_3 = T_2$	0.550	0.423	0.030	0.594	0.875	0.015
H_0: $T_1 = T_2$	0.579	0.002	0.000	0.934	0.002	0.000
H_0: $T_1 = T_2 = T_3$	0.548	0.009	0.000	0.852	0.006	0.000

Notes: The sample includes all treated students in T1, T2, and T3 and all students in control schools. T_1 is an indicator for attending a T1 school, T_2 is an indicator for attending a T2 school, and T_3 is an indicator for attending a T3 school. All regressions control for grade level, student gender, age of father, literacy of father, occupation of father, parent income, number of children, and a dummy indicating whether a student missed any baseline household characteristics. Standard errors are clustered at school level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1.

First, we see that, in the pretreatment test scores, there are no differences across treatment and control students (columns (1) and (4)), which provides further reassurance that the experimental groups are balanced. Second, after the intervention, we find that there is a positive and statistically significant effect of treatment on both English and math standardized test scores, and this effect is slightly more pronounced in the endline, especially for English, at least for T2 and T3. The effects are sizable. Focusing on the endline test scores in 2015, they range from 0.38 SD in T1 to 0.9 SD in T2 for English and from 0.48 SD in T1 to 0.97 SD in T2 for math, with the effects in T3 lying in between (all direct treatment effects are significant with $p < 0.01$). Third, these treatment effects are statistically distinguishable from each other for both English and math (p-values reported at bottom of Table 2). In particular, we see that the direct treatment effect in T3 is substantially larger than

[8] In Table A7 of Appendix A.5 we also report the overall treatment effect of the interventions, that is, the impact of treatment (private tutoring for $T1-$schools, private tutoring and home visits for $T2-$schools, private tutoring provided to the most central students for $T3-$schools) on the test scores of all students in treatment schools.

that in T1, implying that more central students benefit more from the intervention. We return to this observation in section 6, where we carry out an overall evaluation of T1 and T3.

4.2 Noncognitive skills

We next examine the direct treatment effects on *noncognitive skills*. Table 3 displays these results. Note that we only elicited measures of noncognitive skills at the endline, and for a slightly smaller number of students (10,991) than those for whom we have test scores.[9]

Compared to the control students, treated students across all three treatment groups improve in "motivation" and "social skills" and, globally, in their noncognitive skills index. However, we do not find a treatment effect on "self-control," although this can probably be attributed to the fact that there is less variation in this variable, as most students are assigned high scores in this dimension (see figure A4). Treatment effects are slightly stronger in T2 compared to T1 and T3; however, we fail to reject the hypothesis that the effects are equal in any of the pairwise comparisons. In terms of magnitude, the treatment effects on the index range from 0.19 SD in T1 to 0.28 SD in T3, suggesting non-negligible effects, though smaller than the ones obtained on test scores.[10] This is not surprising given that the interventions were designed to mainly address students' knowledge of English and math, which the test scores measure directly.

[9]Table A12 of Appendix A.8 reports a set of balance tests of baseline characteristics across treatment arms for the subsample of 10,991 students for whom we have measures of noncognitive skills.

[10]In Table A8 of Appendix A.5, we also report the overall treatment effect of the interventions, that is, the impact of treatment (private tutoring for $T1-$schools, private tutoring and home visits for $T2-$schools, private tutoring provided to the most central students for $T3-$schools) on the noncognitive skills of all students in treatment schools.

Table 3: Direct effect on noncognitive skills

	Self Control	Motivation	Social skills	Index
	(1)	(2)	(3)	(4)
T_1	0.106	0.266**	0.197**	0.189**
	(0.096)	(0.103)	(0.087)	(0.077)
T_2	0.139	0.409***	0.287***	0.278***
	(0.111)	(0.110)	(0.096)	(0.091)
T_3	0.158	0.251**	0.207*	0.205**
	(0.106)	(0.113)	(0.113)	(0.089)
Observations	8,911	8,911	8,911	8,911
No. of schools	248	248	248	248
H_0: $T_3 = T_1$	0.602	0.907	0.934	0.866
H_0: $T_3 = T_2$	0.869	0.237	0.547	0.489
H_0: $T_1 = T_2$	0.755	0.251	0.414	0.353
H_0: $T_1 = T_2 = T_3$	0.866	0.407	0.695	0.635

Notes: The sample includes all treated students in T1, T2, and T3 and all students in control schools for whom we have measures of noncognitive skills. T_1 is an indicator for attending a T1 school, T_2 is an indicator for attending a T2 school, and T_3 is an indicator for attending a T3 school. All regressions control for grade level, student gender, age of father, literacy of father, occupation of father, parent income, number of children, and a dummy indicating whether a student missed any baseline household characteristics. Standard errors are clustered at school level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1.

5 Spillover effects

We have seen that there are clear positive effects of treatment on both the test scores and the noncognitive skills of treated students. We now turn our attention to the main objective of this paper, which is to measure spillover effects. In other words, we next examine whether the increased knowledge of treated students spills over to untreated students in the same classroom. Importantly, we assess whether the centrality of treated students matters for the strength of the spillover effects.

First, we consider "standard" spillover effects, comparing untreated students in treatment schools to students in control schools. Then, we focus on T1 and T2 schools to probe the source and nature of the spillover effects. In particular, we examine whether, among the randomly treated students, those who are *more central* in the classroom network at baseline generate larger spillovers to untreated students than less central randomly treated students.

5.1 Spillover effects on the untreated students

Comparing untreated students in treatment schools to students in control schools As a baseline measure of spillover effects of the intervention, we compare the outcomes of *untreated* students in T1, T2, and T3 to the outcomes of students in the control group. We do this by estimating equation (1) on this subsample of students consisting of untreated students from T1, T2, and T3 schools and students from control schools. Any differences in outcomes across these two types of students can be attributed to the intervention *indirectly* benefiting untreated students within the treated schools; hence, we refer to it as a spillover effect. This spillover effect could arise due to peer-to-peer spillovers or due to the parents of untreated students and their teachers adjusting their inputs and effort.

Results are presented in Table 4 for test scores and in Table 5 for noncognitive skills, respectively. As we can see, the number of students is smaller in this analysis because we are focusing only on untreated students in treatment schools and students in control schools.[11]

Starting with the results reported in columns (1) and (4) of Table 4, we see that in the pretreatment baseline, as we would expect, there are no systematic differences between untreated students in treatment schools and students in control schools. T2 untreated students appear to perform slightly worse than those in the control schools in both English and math, with the difference in the former being significant at 10%. However, after the introduction of our intervention, we begin to see large and statistically significant improvements in the test scores of untreated students in all treatments relative to control students. Also, consistent with the evidence seen above related to the direct effect, the spillover effects are more pronounced at the endline than the midline. Focusing on the endline, the spillover effects are large and significant with $p < 0.01$ (with the exception of English in T1 in which $p<0.10$). The effects are also much stronger—more than double—and statistically distinguishable in T3 (and T2) relative to T1, 0.59 SD as opposed to 0.25 SD in English, and 0.68 SD as opposed to 0.30 SD in math.

Spillover effects on noncognitive skills are reported in Table 5. In line with the evidence on the direct effect, we find that untreated students in T2 and T3 schools improve in the dimensions of "motivation" and "social skills," whereas in T1 schools the coefficients are not statistically significant. The effects are slightly larger for T2 (0.42 SD for "motivation" and 0.34 SD for "social skills") than T3 (0.26 SD for "motivation" and 0.28 SD for "social skills"), although the differences between these two treatment arms are not statistically significant.

[11] Also, the total number of schools is smaller than in Table 2 because in some schools there are no untreated students.

Table 4: Spillover effects on test scores

	English			Math		
	2013	2014	2015	2013	2014	2015
	(1)	(2)	(3)	(4)	(5)	(6)
T_1	-0.050	0.162	0.249*	0.173	0.254*	0.298***
	(0.195)	(0.120)	(0.150)	(0.149)	(0.136)	(0.114)
T_2	-0.214*	0.231**	0.680***	-0.253	0.392***	0.830***
	(0.110)	(0.106)	(0.160)	(0.194)	(0.102)	(0.165)
T_3	0.029	0.297***	0.594***	-0.102	0.505***	0.676***
	(0.090)	(0.098)	(0.103)	(0.089)	(0.100)	(0.090)
$y_{i,g,s,0}$		0.308***	0.218***		0.294***	0.233***
		(0.020)	(0.022)		(0.023)	(0.022)
Observations	7,904	7,202	7,904	7,904	7,202	7,904
No. of schools	206	183	206	206	183	206
H_0: $T_3 = T_1$	0.705	0.286	0.029	0.010	0.091	0.003
H_0: $T_3 = T_2$	0.070	0.565	0.607	0.470	0.336	0.374
H_0: $T_1 = T_2$	0.455	0.608	0.032	0.078	0.357	0.005
H_0: $T_1 = T_2 = T_3$	0.192	0.556	0.054	0.146	0.226	0.003

Notes: The sample includes all untreated students in T1, T2, and T3 and all students in control schools. T_1 is an indicator for attending a T1 school, T_2 is an indicator for attending a T2 school, and T_3 is an indicator for attending a T3 school. All regressions include controls for grade level, student gender, age of father, literacy of father, occupation of father, parent income, number of children, and a dummy indicating whether a student missed any baseline household characteristics. Standard errors are clustered at school level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table 5: Spillover effects on noncognitive skills

	Self Control	Motivation	Social skills	Index
	(1)	(2)	(3)	(4)
T_1	-0.095	0.191	0.265	0.120
	(0.143)	(0.125)	(0.183)	(0.103)
T_2	0.162	0.423*	0.341**	0.309*
	(0.161)	(0.226)	(0.148)	(0.167)
T_3	0.117	0.259**	0.277*	0.218**
	(0.122)	(0.130)	(0.147)	(0.110)
Observations	5,854	5,854	5,854	5,854
No. of schools	191	191	191	191
H_0: $T_3 = T_1$	0.183	0.672	0.955	0.466
H_0: $T_3 = T_2$	0.797	0.508	0.751	0.628
H_0: $T_1 = T_2$	0.171	0.341	0.738	0.299
H_0: $T_1 = T_2 = T_3$	0.296	0.629	0.929	0.534

Notes: The sample includes all untreated students in T1, T2, and T3 and all students in control schools for whom we have measures of noncognitive skills.T_1 is an indicator for attending a T1 school, T_2 is an indicator for attending a T2 school, and T_3 is an indicator for attending a T3 school. All regressions include controls for grade level, student gender, age of father, literacy of father, occupation of father, parent income, number of children, and a dummy indicating whether a student missed any baseline household characteristics. Standard errors are clustered at school level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1.

To summarize, the intervention seems to have generated substantial spillovers to untreated students in treatment schools in terms of both test scores and noncognitive skills. A second notable finding is that the spillovers in test scores are larger in T3 schools, where more central students are treated, than in T1 schools, where students are treated at random.

5.2 Centrality-based spillover effects

As stated above, the main aim of this paper is to examine whether spillover effects are stronger if transmitted through more central children in the social network. To test this, let us first provide a formal definition of the measure of centrality that we employ, i.e., *eigenvector centrality.*[12]

Eigenvector centrality definition Consider a network $\mathbf{g}$ in which $g_{ij} = 1$ indicates that there is a friendship link between students i and j (i.e., if either i has nominated j or j has nominated i), and $g_{ij} = 0$, otherwise. The eigenvector centrality $E_{i,g,s,t}(\mathbf{g})$ of individual (i, g, s) at time t in

[12]There are other measures of network centrality (Wasserman & Faust, 1994; Jackson, 2008; Jackson et al., 2017; Bloch et al., 2019), but eigenvector centrality is the most common measure used in network economics, especially in empirical work. Further, the theoretical foundations of eigenvector centrality are both axiomatic (Palacios-Huerta & Volij, 2004; Dequiedt & Zenou, 2017; Bloch et al., 2019) and microfounded (Golub & Jackson, 2010; Banerjee et al., 2013; Elliott & Golub, 2019; Bochet et al., 2020; Sadler, 2020).

network $\mathbf{g}$ is defined using the following recursive formula:

$$E_{i,g,s,t}(\mathbf{g}) = \frac{1}{\lambda_1(\mathbf{g})} \sum_{j=1}^{n} g_{ij} E_{j,g,s,t}(\mathbf{g}), \tag{2}$$

where $\lambda_1(\mathbf{g})$ is the largest eigenvalue of network $\mathbf{g}$. By the Perron-Frobenius theorem, using the largest eigenvalue guarantees that $E_{i,g,s,t}(\mathbf{g})$ is strictly positive. In matrix form, we have:

$$\lambda_1(\mathbf{g})\mathbf{E}_{g,s,t}(\mathbf{g}) = \mathbf{G}\mathbf{E}_{g,s,t}(\mathbf{g}), \tag{3}$$

where $\mathbf{G} = [g_{ij}]$ is the adjacency matrix of network $\mathbf{g}$, and $\mathbf{E}_{g,s,t}$ is the column vector of eigenvector centralities in grade g in school s at time t. We then normalize the eigenvector centralities in each classroom (i.e., grade g) by requiring that they sum to 1. This ensures that the average centrality stays (relatively) constant over different grades and schools since it is equal to $1/n_g$, where n_g is the number of students in classroom/grade g. It also guarantees that the range of the eigenvector centrality $E_{i,g,s,t}$ is $[0, 1]$.

Eigenvector centrality assigns relative scores to all nodes in the network based on the principle that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes.

Centrality exposure Our main explanatory variable is the *centrality exposure*, $\overline{E}^T_{g,s,0}$, defined as the average eigenvector centrality of treated students in grade g in school s at time $t = 0$. For an untreated student i in grade g in school s at time $t = 0$, her *centrality exposure* is thus given by:

$$\overline{E}^T_{g,s,0} = \frac{\sum_j E^T_{j,g,s,0}}{N^T_{i,g,s}}, \tag{4}$$

where $E^T_{j,g,s,0}$ is the eigenvector centrality of a *treated* individual j, in grade g, in school s, at time $t = 0$ defined in (2), and $N^T_{i,g,s}$ is the number of treated students, who go to the same school s and are in the same grade g as student i. As stated above, we measure the centrality exposure rate at the baseline because the intervention might affect the network centrality of the students by making them more popular and thus more central.

Econometric specification To assess centrality spillover effects, we estimate the following specification:

$$y_{i,g,s,t} = \alpha_0 + \alpha_1 \overline{E}^T_{g,s,0} + \gamma y_{i,g,s,0} + X'_{i,g,s}\beta + \delta T2 + Grade_{i,g,s,0} + \epsilon_{i,g,s,t}, \tag{5}$$

in which the main variable of interest is the *centrality exposure* $\overline{E}_{g,s,0}^{T}$ defined in (4).[13] In this specification, beyond the usual household characteristics, we add additional controls to capture the characteristics of the network $\mathbf{g}$ that the untreated student (i, g, s) belongs to; these include network size (i.e., number of students in a classroom), network density, and network diameter.[14] We also include an indicator for whether the school s is assigned to T2. As above, standard errors are clustered at the school level.

Note that, in this analysis, as we are restricting our sample to untreated students from T1 and T2 schools, we pool observations across the two treatments, but we do include a dummy for treatment (T_2 in equation (5)). This is done mainly to avoid the problem of unreliable cluster-robust standard errors due to having a small number of clusters (schools) with untreated students (37 in T1 and 34 in T2). Beyond this inference consideration, we believe there is little loss from pooling these two treatments, as evaluating whether centrality spillovers are different across the T1 and T2 interventions is not important for the main message of this paper.

Identification The main novelty of this paper is that it exploits across-classroom random variation in the centrality exposure ($\overline{E}_{g,s,0}^{T}$) of untreated students induced by our experimental design in T1 and T2 for the causal identification of the influence of central classmates on their classmates' learning. Since students were randomly selected to receive the intervention, it is reasonable to expect that $\overline{E}_{g,s,0}^{T}$ is exogenous with respect to the educational outcomes of untreated students.

Figure A6 in Appendix A.10 plots the distribution of centrality exposure. We see that it is roughly normally distributed and that there is considerable variation between classrooms that we will exploit to identify the impact of centrality. To provide further support for our identification, we perform a type of balancing test in which we examine whether variation in centrality exposure is related to the variation in predetermined student and household characteristics. Table A14 in Appendix A.10 reports coefficients from a set of bivariate regressions of centrality exposure on baseline test scores and on each of the student and household characteristics. We find no significant correlations in most cases, with the exception of two household characteristics that we control for in our analysis. This provides support for our identification strategy.

5.3 Centrality-based spillovers on test scores

Table 6 displays the results separately for English and math test scores. Note that because we have standardized $\overline{E}_{g,s,0}^{T}$, the estimated coefficients can be interpreted as the impact of a one SD in the exposure rate.

Columns (1) and (4) present a placebo test: the centrality spillover at baseline. As expected, the effects are insignificant, that is, the centrality of treated students does not affect untreated students before the intervention. Instead, in columns (2) and (3) for English and (5) and (6) for math, we can

[13]In presenting regression results, we normalize $\overline{E}_{g,s,0}^{T}$ by dividing it by its standard deviation to ease interpretation of the coefficient α_1.

[14]We present some summary statistics of network characteristics in T1 and T2 classrooms in Appendix A.9.

see that centrality-based spillover effects are positive and statistically significant (p<0.05). Taking the point estimate at endline, we find that a one SD increase in the average centrality of treated students leads to an increase in the test scores of untreated students of 0.38 SD in English and 0.32 SD in math.

Table 6: Centrality-based spillovers on test scores

	English			Math		
	2013	2014	2015	2013	2014	2015
	(1)	(2)	(3)	(4)	(5)	(6)
$\overline{E}^{T}_{g,s,0}$	0.239	0.328**	0.376**	-0.0761	0.453**	0.321**
	(0.341)	(0.145)	(0.183)	(0.305)	(0.172)	(0.153)
$y_{i,g,s,0}$		0.346***	0.242***		0.289***	0.167***
		(0.023)	(0.044)		(0.039)	(0.038)
T_2	-0.243	-0.063	0.385**	-0.415	-0.022	0.437***
	(0.192)	(0.112)	(0.155)	(0.258)	(0.115)	(0.149)
Observations	1,263	1,112	1,263	1,263	1,112	1,263
No. of schools	71	53	71	71	53	71

Note: The sample includes all untreated students in T1 and T2 schools. $\overline{E}^{T}_{g,s,0}$ is the average centrality of treated students by own gender. All regressions control for baseline household characteristics, grade level, a dummy indicating whether any baseline household characteristics are missing, and three network characteristics, including network size, network density, and network diameter. Standard errors are clustered at the school level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1.

5.4 Centrality-based spillovers on noncognitive skills

So far, we have examined the effect of centrality exposure on the test scores in English and math of untreated students. Now, we would like to study the same relationship but on the noncognitive scores of untreated students. Table 7 displays these results.

Table 7: Centrality-based spillovers on noncognitive skills

	(1) Self control	(2) Social skills	(3) Motivation	(4) Index
$\overline{E}^{T}_{g,s,0}$	0.367*	0.672***	0.717**	0.585***
	(0.194)	(0.238)	(0.285)	(0.208)
T_2	0.238	0.018	0.182	0.146
	(0.215)	(0.162)	(0.257)	(0.180)
Observations	937	937	937	937
No. of schools	59	59	59	59

Note: The sample includes all untreated students in T1 and T2 schools for whom we have measures of noncognitive skills. $\overline{E}^{T}_{g,s,0}$ is the average centrality of treated students. All results control for baseline household characteristics, a dummy indicating whether missing any baseline household characteristics, class level, and network characteristics, including network size, network density, and network diameter. Standard errors are clustered at the school level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1.

The results obtained indicate that centrality-based spillover effects on noncognitive skills are generally positive and statistically significant. We find them to operate through all three types of noncognitive skills, although the strongest impact seems to occur on "motivation." The effect sizes are also shown to be slightly larger than those shown on test scores in Table 6.

5.5 Understanding centrality-based spillover effects

Our main results in the previous subsection suggest that the more "exposed" an untreated student is to highly central treated students, the higher his or her test scores are in English and math, and there are also improvements in measures of noncognitive skills. In this section, we probe the question of how centrality-based spillovers might operate.

To this end, we attempt to understand the profile of central students to gain some insight into why they are effective in generating spillovers. We first examine the correlation between observable characteristics and the network (eigenvector) centrality of students. In particular, we look at the correlation between eigenvector centrality (at baseline) and academic ability, measured by baseline test scores, for all students in T1 and T2. Table 8 presents these results. For test scores in both English and math, we see that the correlation is positive, with and without additional controls. This indicates that students who do well academically in school tend to also be more popular.

For noncognitive skills, as we only have measures at endline, we look at their contemporaneous correlation with centrality measured at endline. Table 9 presents these results. Interestingly, among the noncognitive skills we observe a positive and statistically significant correlation between eigenvector centrality and social skills, indicating that naturally popular students have good social skills.

Table 8: The correlation between centrality and test scores (in 2013)

	English		Math	
	(1)	(2)	(3)	(4)
$E_{i,g,s,0}$	0.205***	0.258***	0.204***	0.257***
	(0.044)	(0.037)	(0.050)	(0.044)
Controls	No	Yes	No	Yes
Observations	6,268	6,268	6,268	6,268
No. of schools	116	116	116	116

Notes: The sample includes all students in T1 and T2 schools. All characteristics are measured at the baseline (2013). $E_{i,g,s,0}$ is the centrality of a student in year 2013. Columns (2) and (4) control for baseline household characteristics, a dummy indicating whether missing any baseline household characteristics, grade level, and three network characteristics, including network size, network density, and network diameter. Standard errors are clustered at the school level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table 9: The correlation between centrality and noncognitive skills (in 2015)

	Self control	Motivation	Social skills	Index
	(1)	(2)	(3)	(4)
$E_{i,g,s,2}$	-0.019	0.057	0.086**	0.041
	(0.045)	(0.048)	(0.042)	(0.037)
Controls	Yes	Yes	Yes	Yes
Observations	5,025	5,025	5,025	5,025
No. of schools	116	116	116	116

Notes: The sample includes all students in T1 and T2 schools for whom we have measures of noncognitive skills. $E_{i,g,s,2}$ is the centrality of a student measured in year 2015. All regressions include controls for baseline household characteristics, grade level, a dummy indicating whether missing any baseline household characteristics, and three network characteristics, including network size, network density, and network diameter. Standard errors are clustered at the school level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1.

In addition, we have seen in Table 4 that spillover effects on test scores of untreated students are larger in T3 than T1. To assess whether this might be driven by the fact that treated students in T3–who are the most central students in the class–are perhaps more engaged with their classmates in school work, we make use of data collected from a student survey at the baseline. Specifically, we focus on a set of questions that examine the extent to which students interact and seek support from their classmates in matters related to studying and schoolwork. We directly compare treated students in T3 to those in T1 in these dimensions. Table 10 presents these results. We see that treated students in T3 tend to report being generally more likely to seek and receive help from classmates than those in T1, suggesting that they tend to be more likely to exchange information with their classmates on matters related to learning in school.

Table 10: Intensity of interaction with classmates (T1 vs T3)

	Questions	T1	T3	T1=T3
Q1	Do you approach others when you face difficulties regarding study?	0.29	0.32	0.604
Q2	Do you often ask others when you cannot understand the lesson?	0.26	0.43	0.030
Q3	Do you often receive help from others to do your classwork?	0.24	0.31	0.284
Q4	Do you communicate with others when you are upset?	0.80	0.86	0.197
Q5	Do you talk to others when you receive poor grades?	0.37	0.51	0.038
	Observations	2,604	1,287	

Notes: The sample includes all treated students in T1 and T3 schools. Q1, Q4, and Q5 are measured by indicator variables that are equal to 1 if the answer is yes. Q2 and Q3 are answered on a scale from 1 to 5. We converted these answers to indicators taking the value of 1 if a student answers 4 or 5 and 0 otherwise. Means are reported. P-values are obtained from OLS regressions, with standard errors clustered at school level.

Overall, these results indicate that more central treated students are academically stronger, have better social skills, and are more likely to engage with their classmates in schoolwork-related matters. These qualities provide a plausible explanation for why more central treated students generate larger spillovers: they are better at transmitting their knowledge (due to treatment) to their untreated classmates.

To probe further whether the centrality-based spillover effects we estimate in Table 6 are mainly capturing the fact that more central students are academically more advanced, we also examine whether variation in the average ability of treated students, as measured by their average test scores at baseline, explains untreated students' test scores at the midline and endline. To do so, we estimate a version of equation (5) in which we replace the centrality exposure with the average test scores of treated students before treatment. We do this for test scores in English and math. The results, presented in Table 11, indicate that exposure to students with better initial test scores in English has a positive and statistically significant impact on the test scores of untreated students at the endline but not at the midline. A one SD increase in the average English test score of the treated is associated with a 0.17 SD increase in the test score of untreated students. For math, the effects are smaller and not statistically significant. These results suggest that the centrality spillover effects estimated above do not simply capture variation in the average ability of treated students. The ability of treated students seems to play a role, at least for spillovers in English test scores, but does not fully account for the centrality spillovers estimated in Table 6, suggesting that other attributes such as social skills and intensity of interaction with peers play an important role as well.

Table 11: Initial ability-based spillover effects

	English		Math	
	2014	2015	2014	2015
	(1)	(2)	(3)	(4)
$\bar{y}^T_{g,s,0}$	0.019	0.170***	0.047	-0.008
	(0.053)	(0.053)	(0.058)	(0.073)
$y_{i,g,s,0}$	0.346***	0.190***	0.265***	0.168***
	(0.029)	(0.037)	(0.041)	(0.042)
T_2	-0.031	0.450***	0.034	0.475***
	(0.115)	(0.143)	(0.122)	(0.171)
Observations	1,112	1,263	1,112	1,263
No. of schools	53	71	53	71

Notes: The sample includes all untreated students in T1 and T2 schools. $\bar{y}^T_{g,s,0}$ is the average standardized test scores of all treated students in a classroom in the baseline year (t=0). $y_{i,g,s,0}$ is the standardized baseline test score of student i, and T_2 is an indicator for attending a T2 school. All specifications include controls for baseline characteristics, including student gender, age of father, literacy of father, occupation of father, household income, number of children in the household, a dummy indicating whether any household characteristics are missing, a class-level indicator, and three network characteristics, including network size, density, and diameter. Standard errors are clustered at the school level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1

.

5.6 A theoretical framework

In order to gain more intuition about the mechanisms at work, let us develop a simple theoretical model building on the framework in Ballester et al. (2006) and Calvó-Armengol et al. (2009).

5.6.1 Benchmark model

Consider a finite set of students $\mathcal{N} = \{1, \ldots, n\}$, with each student corresponding to a node in a network $\mathbf{g}$. We keep track of social connections in network $\mathbf{g}$ through its adjacency matrix $\mathbf{G} = [g_{ij}]$, where $g_{ij} = 1$ if nodes i and j ($i \neq j$) are connected and $g_{ij} = 0$ otherwise. Thus, $\mathbf{G}$ is a zero-diagonal symmetric square matrix.

Students in network $\mathbf{g}$ decide how much study effort to exert in education. As in the empirical analysis, we denote by $y_{i,g,s,t}$ the study effort level of agent i in grade g in school s at time t. For the ease of the presentation, in the theoretical model, we omit the subscript g, s, t and thus refer to the effort $y_{i,g,s,t}$ as $y_{i,t}$. Denote by $\mathbf{y}_t = (y_{1,t}, ..., y_{n,t})'$ the vector of effort profile in network $\mathbf{g}$. Each agent i, t selects an effort $y_{i,t} \geq 0$ and obtains a payoff $U_i(\mathbf{y}_t, \mathbf{g})$ that depends on the effort profile $\mathbf{y}_t$ and on the underlying network $\mathbf{g}$ in the following way:

$$U_i(\mathbf{y}_t, \mathbf{g}) = \underbrace{(\pi_i + \alpha_1 \sum\nolimits_{j \in \mathcal{N}} g_{ij} y_{j,t}) y_{i,t}}_{\text{payoff}} - \underbrace{\tfrac{1}{2} y^2_{i,t}}_{\text{cost}}, \tag{6}$$

with $\alpha_1 \geq 0$. This utility has a standard cost-payoff structure. The payoff increases in own effort $y_{i,t}$

with the marginal payoff given by $\pi_i + \alpha_1 \sum_{j=1}^{n} g_{ij} y_{j,t}$. The parameter $\alpha_1 \geq 0$ is the *social-multiplier* or *spillover* coefficient, which captures strategic complementarity in the effort of two students who are connected through a friendship tie. That is, if a student increases her effort in studying, then the marginal utility of her friend's effort also increases. In our education setting, the complementarity could arise because of a desire to imitate a friend and/or because of the direct learning benefits stemming from the interaction with friends. The cost part of the utility function (6) is a *direct* cost of exerting effort given by $\frac{1}{2} y_i^2$.

The term π_i represents the *exogenous heterogeneity* of agent i's "productivity" in education activities and is given by

$$\pi_i = \gamma y_{i,0} + X_i' \beta + Grade_{i,0} + \epsilon_{i,t}, \tag{7}$$

where, as in the empirical analysis, $y_{i,0}$ is the baseline outcome measured at $t = 0$ (cognitive test scores); X_i' denotes student and household characteristics, including gender, father's age, literacy and occupation, household income, and the number of children in the household; $Grade_{i,0}$ is an indicator of being in the senior grade we examine, and $\epsilon_{i,t}$ is an error term. Remember that $y_{i,t}$ is the outcome of interest and denotes either a test score or corresponds to a measure of noncognitive skill of student i at time t.

5.6.2 Explaining the correlation between test score or non-cognitive skills and eigenvector centrality

In equilibrium, each agent maximizes her utility, and the best-response function is given by:

$$y_{i,t} = \alpha_1 \sum\nolimits_{j \in \mathcal{N}} g_{ij} y_{j,t} + \pi_i. \tag{8}$$

In Proposition 1 in Appendix B, we show under which condition this game has a unique Nash equilibrium given by (8). We show that, when students choose their education effort, in the unique Nash equilibrium, the education effort will be proportional to their Katz-Bonacich or eigenvector centrality. In particular, since Proposition 1 implies that $y_{j,t} = E_{j,t}(\mathbf{g})$, equation (8) can be written as:

$$y_{i,t} = \alpha_1 \sum\nolimits_{j \in \mathcal{N}} g_{ij} E_{j,t}(\mathbf{g}) + \pi_i, \tag{9}$$

which corresponds to our econometric specification (5).

In other words, more central students exert higher equilibrium effort, reaping the complementarities from their friendship ties. This is consistent with the evidence in our data, as shown in Tables 8 and 9, where we find a strong positive correlation between test scores or non-cognitive skills and individual (eigenvector) centrality in the network.

5.6.3 A model with treated and untreated students

Consider the same model, but now the population of students is exogenously divided between treated (T) and untreated (NT) students, i.e., $\mathcal{N} = \{1, \ldots, n\} = \mathcal{N}^T + \mathcal{N}^{NT}$. Since we randomly select the treated and untreated students, there is no correlation between centrality and treatment. However, as we will show below, treated students who are more central will have a stronger impact on their peers than less central treated students.

Consistent with our empirical results (direct treatment effects; Tables 2 and 3), we assume that, when a student obtains a treatment (either private tutoring (T1) or home visits (T2)), this increases her productivity π_i. Thus, for a treated student, we have:

$$\pi_i^T = \pi_i + \overline{\pi},$$

where $\overline{\pi} > 0$ and $\pi_i := \pi_i^{NT}$ (productivity of the untreated students) is defined in Equation (7).[15] This implies that treated students generate more spillover effects than untreated students. To capture this aspect, instead of (6), we now assume:

$$U_i^T(\mathbf{y}_t, \mathbf{g}) = \pi_i^T y_{i,t}^T - \frac{1}{2}\left(y_{i,t}^T\right)^2 + \alpha_1\theta \sum_{j=1}^{n} g_{ij}^{Intra} y_{i,t}^T y_{j,t}^T + \alpha_1 \sum_{j=1}^{n} g_{ij}^{Inter} y_{i,t}^T y_{j,t}^{NT}, \tag{10}$$

$$U_i^{NT}(\mathbf{y}_t, \mathbf{g}) = \pi_i^{NT} y_{i,t}^{NT} - \frac{1}{2}\left(y_{i,t}^{NT}\right)^2 + \alpha_1 \sum_{j=1}^{n} g_{ij}^{Intra} y_{i,t}^{NT} y_{j,t}^{NT} + \alpha_1\theta \sum_{j=1}^{n} g_{ij}^{Inter} y_{i,t}^{NT} y_{j,t}^{T}, \tag{11}$$

where $\theta > 1$ and where $\mathbf{G}^{Inter} = \left(g_{ij}^{Inter}\right)$ is the adjacency matrix for only the *inter-type links*, i.e., only between treated (T) and untreated (NT) students, and $\mathbf{G}^{Intra} = \left(g_{ij}^{Intra}\right)$ is the adjacency matrix for only the *intra-type links*, i.e., only between treated (T) and treated (T) students and between untreated (NT) and untreated (NT) students. Obviously, $\mathbf{G}^{Inter} + \mathbf{G}^{Intra} = \mathbf{G}$. For simplicity and without loss of generality, we order the players so that the n^T first players are of type T and the next n^{NT} ones are of type NT, with $n = n^T + n^{NT}$, so that the first n^T rows of the $\mathbf{G}$ matrix correspond to the type$-T$ students and the last n^{NT} rows corresponds to the type$-NT$ students.

As stated above, this specification implies that the spillover effects from treated students are stronger than from untreated students, that is, the intensity of the spillover effects from treated and untreated students are $\alpha_1\theta$ and α_1, respectively, with $\alpha_1\theta > \alpha_1$, since $\theta > 1$.

The best-reply functions for the treated and untreated students are respectively given by:

$$y_{i,t}^T = \pi_i^T + \alpha_1\theta \sum_{j=1}^{n} g_{ij}^{Intra} y_{j,t}^T + \alpha_1 \sum_{j=1}^{n} g_{ij}^{Inter} y_{j,t}^{NT}, \tag{12}$$

[15] We assume that the productivity benefit of treatment $\overline{\pi}$ is not a function of centrality. In principle, this need not be the case; it may well be that more central students benefit more because they are on average better students. This is ultimately an empirical issue, so to keep the exposition of the model simple, we take here the neutral view that treatment effects are homogeneous with respect to centrality.

$$y_{i,t}^{NT} = \pi_i^{NT} + \alpha_1 \sum_{j=1}^{n} g_{ij}^{Intra} y_{j,t}^{NT} + \alpha_1 \theta \sum_{j=1}^{n} g_{ij}^{Inter} y_{j,t}^{T}, \tag{13}$$

where $\alpha_1^T = \alpha_1 \theta > 0$ and $\alpha_1^{NT} = \alpha_1 > 0$.

In Proposition 2 in Appendix B, we show under which condition this game has a unique Nash equilibrium given by (12) and (13). Using the results in Proposition 1 that show that education efforts are equal to their eigenvector centrality, as in (9), we can replace the efforts in the right-hand side of these equations with their eigenvector centrality. This illustrates that more central students exert a greater externality on the education efforts of their peers.

5.6.4 Targeting randomly students (T1)

Let us illustrate our empirical results by considering the network displayed in Figure 1 with four students.

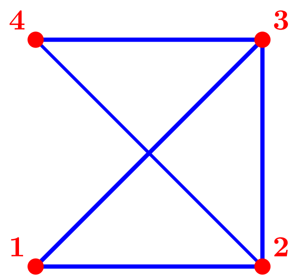


Figure 1: A network with four students

As in treatment T1 (or T2), we consider a policy that *randomly* treats students. Assume that students 1 and 2 are treated (type T) while students 3 and 4 are not (type NT). Assume that $\theta = 4$, that is, treated students generate four times more spillovers than untreated students. Define $\mathbf{A}^{Intra}$ as a diagonal matrix where, on the diagonal, the first $n^T = 2$ rows have $\theta = 4$ while the remaining $n^{NT} = 2$ rows have 1. Similarly, define $\mathbf{A}^{Inter}$ as a diagonal matrix where, on the diagonal, the first $n^T = 2$ rows have 1 while the remaining $n^{NT} = 2$ rows have $\theta = 4$. We have:

$$\mathbf{A}^{Intra} = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{A}^{Inter} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

and

$$\mathbf{G}^{Intra} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \mathbf{G}^{Inter} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

It is easily verified that $\rho(\mathbf{G}) = 2.56$, and the corresponding vector of eigenvector centralities is given by:

$$\mathbf{E}_t(\mathbf{g}) = \begin{pmatrix} E_1(\mathbf{g}) \\ E_2(\mathbf{g}) \\ E_3(\mathbf{g}) \\ E_4(\mathbf{g}) \end{pmatrix} = \begin{pmatrix} 1 \\ 1.28 \\ 1.28 \\ 1 \end{pmatrix}. \tag{14}$$

Indeed, as can be seen in Figure 1, students 2 and 3 have the highest eigenvector centrality. Using Proposition 2, we can calculate the unique Nash equilibrium of this game. To have an explicit solution, assume that $\alpha = 0.05$, and $\pi_1 = \pi_2 = \pi_3 = \pi_4 = 1$ and $\overline{\pi} = 2$. Indeed, for simplicity and to be able to compare them, we assume that all students are ex ante, that is, before treatment, the same and thus have the same productivity. After treatment, we have $\pi_1^T = \pi_2^T = 3$, $\pi_3^{NT} = \pi_4^{NT} = 1$, that is, students 1 and 2, who were treated, triple their productivity. Then, we obtain:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 3 + 0.2y_2 + 0.05y_3 \\ 3 + 0.2y_1 + 0.05\,(y_3 + y_4) \\ 1 + 0.2\,(y_1 + y_2) + 0.05y_4 \\ 1 + 0.2y_2 + 0.05y_3 \end{pmatrix}. \tag{15}$$

And the unique Nash equilibrium is given by:

$$\begin{pmatrix} y_1^T \\ y_2^T \\ y_3^{NT} \\ y_4^{NT} \end{pmatrix} = \begin{pmatrix} 3.94 \\ 4.02 \\ 2.69 \\ 1.94 \end{pmatrix}. \tag{16}$$

Not surprisingly, student 2 has the highest outcome (test score or non-cognitive skills) because she is the most central (in terms of eigenvector centrality) in the network and has been treated, that is, received private tutoring, which increases her productivity. Interestingly, student 3, who has a higher eigenvector centrality than student 1 (see (14)), ends up with a lower outcome because she received private tutoring (*direct effect*) but also because her extra link (compared to student 1) is with an untreated student (student 4), who generates weak spillover effects. Indeed, we assume that $\theta = 4$, that is, treated students generate four times more spillover effects than untreated students.

5.6.5 Targeting the most central students (T3)

Consider the same network displayed in Figure 1 but instead of having a *random* treatment (as in T1 or T2), we consider a *targeted* treatment as in T3, that is, a policy that treats the *most central* (in terms of eigenvector centrality) students in the network. Thus, assume now that students 2 and 3 are treated (see (14)). This implies that $\pi_2^T = \pi_3^T = 3$, $\pi_1^{NT} = \pi_4^{NT} = 1$. The inter-link and intra-link productivity and adjacency matrices change and are now given by

$$\mathbf{A}^{Intra} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{A}^{Inter} = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

and

$$\mathbf{G}^{Intra} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \mathbf{G}^{Inter} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

With exactly the same parameters values, it is easily verified that the best-reply functions are given by:

$$\begin{pmatrix} y_1^{NT} \\ y_2^T \\ y_3^T \\ y_4^T \end{pmatrix} = \begin{pmatrix} 1 + 0.2\,(y_2 + y_3) \\ 3 + 0.05\,(y_1 + y_4) + 0.2 y_3 \\ 3 + 0.05\,(y_1 + y_4) + 0.2 y_2 \\ 1 + 0.2\,(y_2 + y_3) \end{pmatrix}. \tag{17}$$

Compared with (15), student 4 (who is not treated in both policies) benefits from more spillovers in the targeted policy because she is only linked to students 2 and 3, who are the most central and now treated. For student 1, there is a trade off: she experiences less of a direct effect in terms of productivity (which is reduced from 3 to 1) but obtains more spillover effects since she is linked to 2 and 3, who are now treated. Student 3 is clearly better off. The unique Nash equilibrium is now given by:

$$\begin{pmatrix} y_1^T \\ y_2^T \\ y_3^{NT} \\ y_4^{NT} \end{pmatrix} = \begin{pmatrix} 2.63 \\ 4.08 \\ 4.08 \\ 2.63 \end{pmatrix}. \tag{18}$$

We see that, compared to the random targeting, nearly all students benefit from this policy. Student 1 is the only one who has a lower outcome because the spillover gain (*indirect effect*) of being a friend to student 3, who is now treated, cannot compensate for the lost in extra productivity (*direct effect*) of not being treated.

In summary, this model provides a simple mechanism that can explain our main empirical results shown in Tables 6 and 7. Students with higher centrality in a network have not only higher ability and better social skills than less central students but also benefit more from the treatment. Consequently, they are more able to transfer this knowledge to their untreated classmates.

6 Implementing the policy: Does targeting central students work?

In T3, we implemented a policy where the choice of treated students was based on their centrality, so that only the most central students (in terms of eigenvector centrality) were treated. We next want to compare the effects of centrality-based education policies to those obtained through the standard approach of offering the intervention randomly. For example, suppose that in a classroom of 50 students a certain intervention (i.e., private tutoring) can be offered to only 10 of the students due to limited resources. In the centrality-based approach, we would first determine the network centrality of each of the 50 students and then offer the treatment only to the 10 students among the 50 who have the highest eigenvector centrality in the classroom. In the randomized approach (as in T1), the treatment would be offered randomly to 10 students in the class. How would the two approaches fare in terms of the educational achievements of both treated and untreated students under the two schemes?

We can use the evidence from our study to address this question. We can compare the performance of students in T3 schools to those in T1 schools. We have seen in the analysis above that both treated and untreated students in T3 do better on average than those in T1, which is prima facie evidence in support of the superiority of T3. However, one might be concerned that T3 could lead to more inequality in student outcomes because the more central students who are better students to start with are the ones who receive the treatment.

To check this, Figures 2 and 3 display the cumulative distribution function of test scores in English and math, respectively, for T1 and T3 schools and separately for treated and untreated students. These figures indicate that the distribution of T3 first-order stochastically dominates that of T1, showing no indication of more dispersion or weaker students falling behind.[16] This is true in both English and math and for both treated and untreated students.

[16] A Kolmogorov-Smirnov test rejects equality of the corresponding cumulative distributions in all cases ($p<0.01$).

Figure 2: CDF of endline English test scores for students in T1 and T3 schools

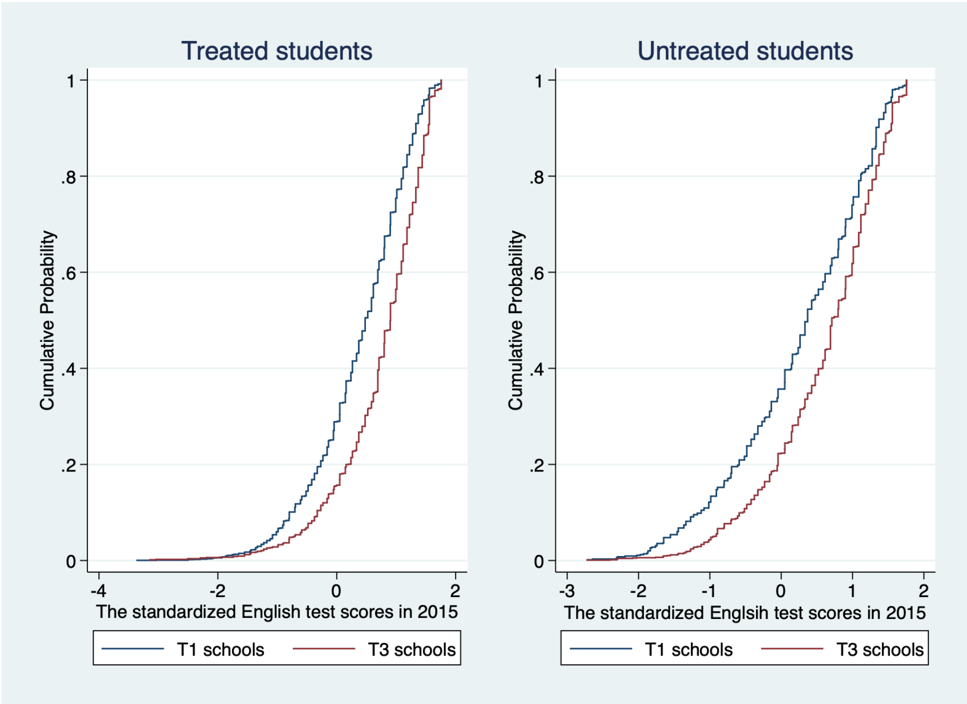
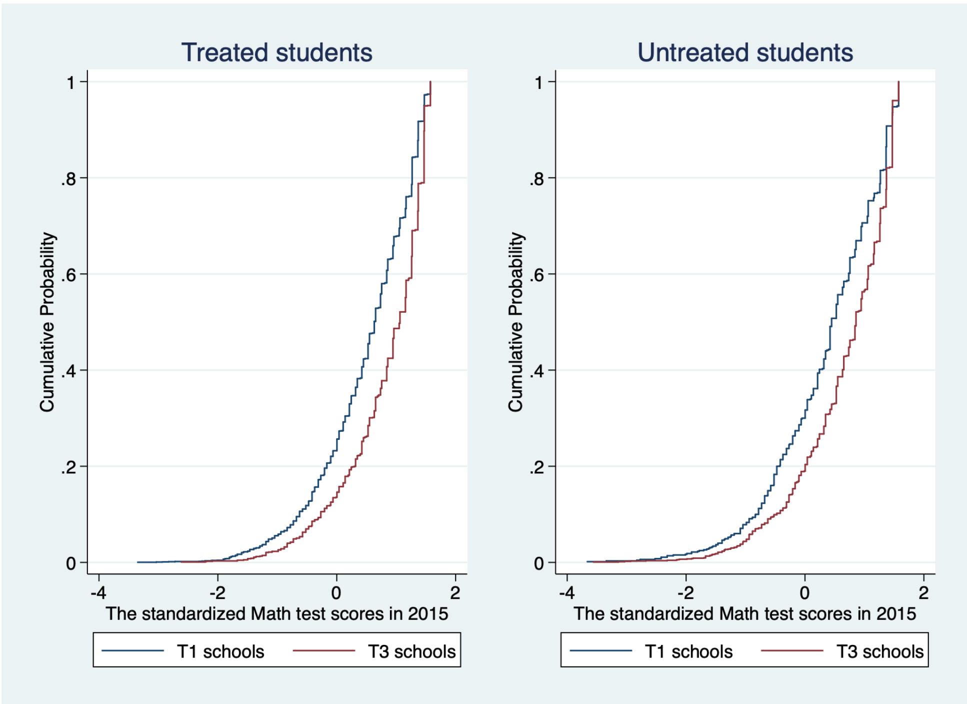


Figure 3: CDF of endline Math test scores for students in T1 and T3 schools



Beyond differences in the strength of the spillover effect of the intervention across T1 and T3, it is important to understand whether there are differences in its *spread* among untreated students. In other words, we would like to know the share of untreated students that benefit from the spillover effects of treated students. To address this question, we examine the share of untreated students that improved their test scores by comparing the baseline and the endline scores across T1 and T3 schools; as a reference, we also report the corresponding share for control schools. In English, 45% of the students in the control schools, 55% of the untreated students in T1 schools, and 64% of the untreated students in T3 schools experienced an improvement in test scores. For math, the respective shares are 44% in control schools, 50% in T1 schools, and 69% in T3 schools. The higher fraction of untreated students in T1 and T3 that saw improvements relative to students in control schools can be attributed to the spillover effects of the intervention. The higher fraction in T3 relative to T1 schools indicates that T3 has a stronger *contagion effect* by achieving a more *widespread diffusion* of the spillover benefits of the intervention among untreated students. This implies that targeting central instead of random students leads, not only, to stronger spillover effects but also to a larger spread. That is, compared to T1 schools, in T3 schools, a *larger* fraction of students increase their test scores and this increase is *higher* for all these students.

If we go back to our theoretical model of Section 5.6 and, in particular, to the network displayed in Figure 1, we showed that, when comparing the random (T1) and targeted (T3) policies, there could be a trade-off between the increased productivity (*direct effect*) of the treated and the spillover effects (*indirect effects*). That is, it may be that low-central students benefit more than high-central students from the direct treatment effects, while more central students generate more spillover effects to untreated students than less central students (Tables 6 and 7) and, by definition, affect more students since they are more connected. However, in Tables 2 and 3 we showed that, for both test scores and non-cognitive skills, the direct treatment effects are, in fact, stronger for central students than for random students. This implies that targeting central students not only has a higher direct effect, but also a larger indirect effect.

Let us calculate the total outcome for the network displayed in Figure 1 under a random and a targeted policy. Using (16), the total outcome under the random policy is equal to (superscript RA stands for *random*):

$$Y^{RA} = y_1^T + y_2^T + y_3^T + y_4^T = 12.59.$$

Using (18), the total outcome under the targeted policy is given by (superscript TA stands for *targeted*):

$$Y^{TA} = y_1^T + y_2^T + y_3^T + y_4^T = 13.42 > 12.59 = Y^{RA}.$$

Total outcome is higher under the targeted (T3) than the random (T1) treatment. If we also compare the *variance* in outcomes across the two policies, it is easily verified that the targeting one has a lower variance than the random policy (0.7 versus 1.02). In other words, not only does the targeted policy have a higher total outcome; it also has a lower variance. This is consistent with Figures 2 and 3.

In summary, providing treatment to the most central students is an effective way of increasing the test scores and noncognitive skills of all students in a classroom. In particular, because of strong and large spillover effects, the untreated students benefit from this policy since there are more students who improve their test scores and they attain higher test scores than the untreated students in classrooms where treatment was given to students at random.

7 Conclusion

One neglected aspect of the large body of literature on spillover and peer effects in education is the role of the social network structure in which students are embedded within their schools.[17] In this paper, we carry out a two-year field experiment in primary schools in rural Bangladesh to investigate this issue. Through our randomized intervention, we engineer an exogenous uplift in the learning and educational outcomes of a random subsample of students through the provision of private after-school tutoring and then causally estimate the spillover effect on the outcomes of their classmates through peer influence. In addition, for a separate group of schools, we offer the same private tutoring to the subgroup of students that are the most central in the social network of their class. Our findings indicate large gains in terms of test scores from the intervention for treated students and more importantly for their untreated classmates. We then show that the gains for untreated students are larger when their treated classmates are more central, which provides the first evidence that network structure is important for educational spillovers.

The treatment that targets the most central students to receive the educational intervention delivers better outcomes for everyone in the classroom, either as direct beneficiaries of the treatment or through the spillover. To the best of our knowledge, this has not been previously shown through a carefully designed randomized intervention in an educational setting. We believe that our evidence has an important policy implication showing that this type of targeted approach to the deployment of an educational intervention can be very effective, particularly in limited-resource contexts such as the one that we study. Of course, further evidence from different interventions and settings would be welcome to provide additional support for this approach.

It is worth noting that Carrell et al. (2013) caution that designing "optimal" peer groups to boost the academic achievement of low-ability students is challenging because of the possible endogenous response of students to the assignment, which could undermine the attainment of a desired outcome. Our approach of targeting central students overcomes this issue, as we take the composition of the peer groups as fixed (the classroom) and then leverage the existing social network to increase the strength of spillovers generated by an educational intervention.

For purposes of practical adoption and scalability of our approach in other educational settings, targeting central students requires collecting network data from classrooms. While this may cause

[17] There are papers that study the direct impact of friends and networks on education outcomes (see, e.g., Babcock (2008); Lin (2010); Bifulco et al. (2011); Fletcher et al. (2020); Norris (2020)), but they do not examine how individual network centrality affects education, especially through spillover effects.

various challenges, there are other approaches that one could consider that would proxy the procedure that we followed, such as asking teachers or school administrators or a sample of students to identify the most popular students or asking how many links have a certain trait (Breza et al., 2020). Such techniques have been shown to be effective in identifying individuals that can effectively spread information (Banerjee et al., 2019). We leave it for future research to ascertain whether similar alternative methods of selecting central students would generate equally strong educational spillover effects as the ones we found in this study.

References

Abramitzky, R., Lavy, V., & Pérez, S. (2021). The long-term spillover effects of changes in the return to schooling. *Journal of Public Economics*, *196*, 104369.

Asadullah, M. N. & Chaudhury, N. (2013). Primary schooling, student learning, and school quality in rural Bangladesh. Center for Global Development Working Paper No. 349.

Avvisati, F., Gurgand, M., Guyon, N., & Maurin, E. (2014). Getting parents involved: A field experiment in deprived schools. *Review of Economic Studies*, *81*(1), 57–83.

Babcock, P. (2008). From ties to gains? Evidence on connectedness and human capital acquisition. *Journal of Human Capital*, *2*(4), 379–409.

Ballester, C., Calvó-Armengol, A., & Zenou, Y. (2006). Who's who in networks. wanted: The key player. *Econometrica*, *74*(5), 1403–1417.

Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2013). The diffusion of microfinance. *Science*, *341*(6144).

Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2019). Using gossips to spread information: Theory and evidence from two randomized controlled trials. *The Review of Economic Studies*, *86*(6), 2453–2490.

Banerjee, A. V., Banerji, R., Duflo, E., Glennerster, R., & Khemani, S. (2010). Pitfalls of participatory programs: Evidence from a randomized evaluation in education in india. *American Economic Journal: Economic Policy*, *2*(1), 1–30.

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

Battaglini, M. & Patacchini, E. (2018). Influencing connected legislators. *Journal of Political Economy*, *126*(6), 2277–2322.

Beaman, L., BenYishay, A., Magruder, J., & Mobarak, A. M. (2021). Can network theory-based targeting increase technology adoption? *American Economic Review*, *111*(6), 1918–1943.

Bennett, M. & Bergman, P. (2021). Better together? Social networks in truancy and the targeting of treatment. *Journal of Labor Economics*, *39*(1), 1–36.

Bifulco, R., Fletcher, J., & Ross, S. (2011). The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health. *American Economic Journal: Economic Policy*, *3*, 25–53.

Bloch, F., Jackson, M. O., & Tebaldi, P. (2019). Centrality measures in networks. *Available at SSRN 2749124*.

Bochet, O., Faure, M., Long, Y., & Zenou, Y. (2020). Perceived competition in networks. CEPR Discussion Paper No. 15582.

Bonacich, P. & Lloyd, P. (2001). Eigenvector-like measures of centrality for asymmetric relations. *Social Networks*, *23*, 191–201.

Breza, E. & Chandrasekhar, A. G. (2019). Social networks, reputation, and commitment: Evidence from a savings monitors experiment. *Econometrica*, *87*(1), 175–216.

Breza, E., Chandrasekhar, A. G., & Larreguy, H. (2015). Network centrality and institutional design: Evidence from a lab experiment in the field. Unpublished manuscript, Harvard University.

Breza, E., Chandrasekhar, A. G., McCormick, T., & Pan, M. (2020). Using aggregated relational data to feasibly identify network structure without network data. *American Economic Review*, *110*(8), 2454–2484.

Calvó-Armengol, A., Patacchini, E., & Zenou, Y. (2009). Peer effects and social networks in education. *The Review of Economic Studies*, *76*(4), 1239–1267.

Carrell, S. E., Sacerdote, B. I., & West, J. E. (2013). From natural variation to optimal policy? the importance of endogenous peer group formation. *Econometrica*, *81*(3), 855–882.

Chandrasekhar, A. G., Kinnan, C., & Larreguy, H. (2018). Social networks as contract enforcement: Evidence from a lab experiment in the field. *American Economic Journal: Applied Economics*, *10*(4), 43–78.

Dequiedt, V. & Zenou, Y. (2017). Local and consistent centrality measures in parameterized networks. *Mathematical Social Sciences*, *88*, 28–36.

Elliott, M. & Golub, B. (2019). A network approach to public goods. *Journal of Political Economy*, *127*(2), 730–776.

Epple, D. & Romano, R. E. (2011). Peer effects in education: A survey of the theory and evidence. In *Handbook of Social Economics*, volume 1 (pp. 1053–1163). Elsevier.

Fletcher, J., Ross, S., & Zhang, Y. (2020). The effect of classmate characteristics on post-secondary outcomes: Evidence from the add health. *Journal of Urban Economics*, *116*, 103241.

Galeotti, A., Goyal, S., Jackson, M. O., Vega-Redondo, F., & Yariv, L. (2010). Network games. *The Review of Economic Studies*, *77*(1), 218–244.

Glewwe, P. & Muralidharan, K. (2016). Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. In *Handbook of the Economics of Education*, volume 5 (pp. 653–743). Elsevier.

Golub, B. & Jackson, M. O. (2010). Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, *2*(1), 112–49.

Government of Bangladesh (2019). Bangladesh primary education annual sector performance report 2019. Technical report, Directorate of Primary Education.

Ioannides, Y. (2013). *From Neighborhoods to Nations: The Economics of Social Interactions.* Princeton: Princeton University Press.

Jackson, M. O. (2008). *Social and Economic Networks.* Princeton: Princeton University Press.

Jackson, M. O. (2019). *The Human Network. How Your Social Position Determines Your Power, Beliefs, and Behaviors.* New York: Pantheon / Penguin Random House.

Jackson, M. O., Rogers, B. W., & Zenou, Y. (2017). The economic consequences of social-network structure. *Journal of Economic Literature*, *55*(1), 49–95.

Jain, T. & Langer, N. (2019). Does who you know matter? Unraveling the influence of peers' network attributes on academic performance. *Economic Inquiry*, *57*(1), 141–161.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, *76*(3), 1071–1102.

Lee, L.-F., Liu, X., Patacchini, E., & Zenou, Y. (2021). Who is the key player? A network analysis of juvenile delinquency. *Journal of Business & Economic Statistics*, *39*(3), 849–857.

Lin, X. (2010). Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics*, *28*(4), 825–860.

Lindquist, M. J. & Zenou, Y. (2014). Key players in co-offending networks. *IZA Discussion Paper.*

List, J. A., Momeni, F., & Zenou, Y. (2020). The social side of early human capital formation: Using a field experiment to estimate the causal impact of neighborhoods. *NBER Working Paper No. 28283.*

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

Meyer, C. (2000). *Matrix Analysis and Applied Linear Algebra.* Philadelphia: SIAM.

Moffitt, R. A. (2001). Policy interventions, low-level equilibria, and social interactions. *Social dynamics*, *4*(45-82), 6–17.

Mohnen, M. (2021). Stars and brokers: Knowledge spillovers among medical scientists. *Management Science,* forhcoming.

Muralidharan, K., Singh, A., & Ganimian, A. J. (2019). Disrupting education? Experimental evidence on technology-aided instruction in India. *American Economic Review*, *109*(4), 1426–60.

Norris, J. (2020). Peers, parents, and attitudes about school. *Journal of Human Capital*, *14*(2), 290–342.

Palacios-Huerta, I. & Volij, O. (2004). The measurement of intellectual influence. *Econometrica*, *72*, 963–977.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In *Handbook of the Economics of Education*, volume 3 (pp. 249–277). Elsevier.

Sacerdote, B. (2014). Experimental and quasi-experimental analysis of peer effects: two steps forward? *Annual Review of Economics*, *6*(1), 253–272.

Sadler, E. (2020). Ordinal centrality. *Available at SSRN 3594819*.

Wasserman, S. & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

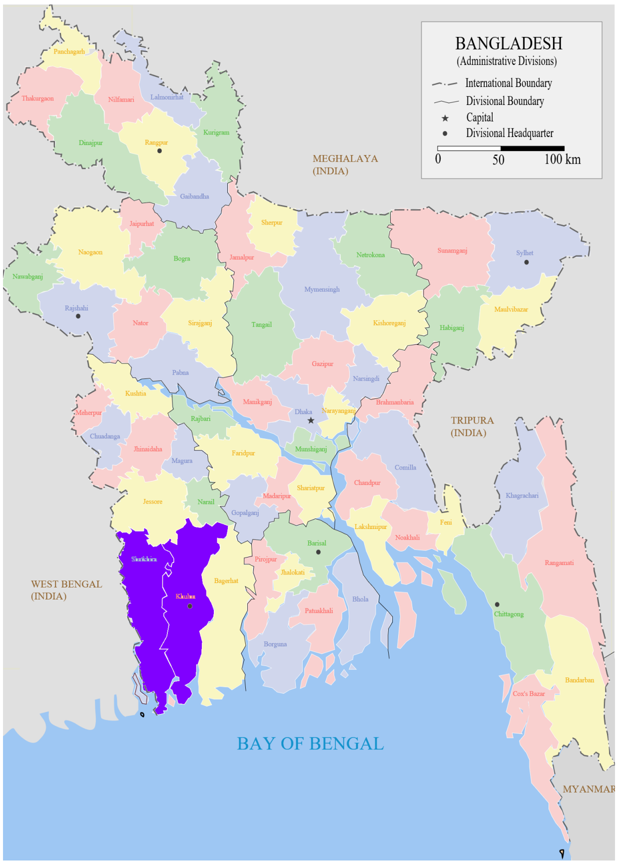
World Bank (2017). *World development report 2018: Learning to realize education's promise*. The World Bank.

Appendix

A Additional Figures and Tables

A.1 Map, timing, randomization process

Figure A1: Location of the study



Notes: The field experiment took place in Satkhira and Khulna, which are the two purple shaded areas as marked on the map.

Figure A2: Randomization process

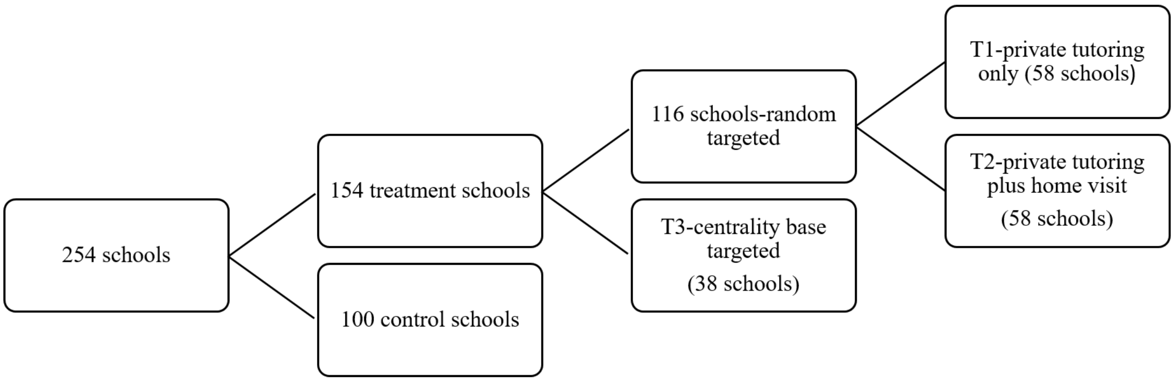
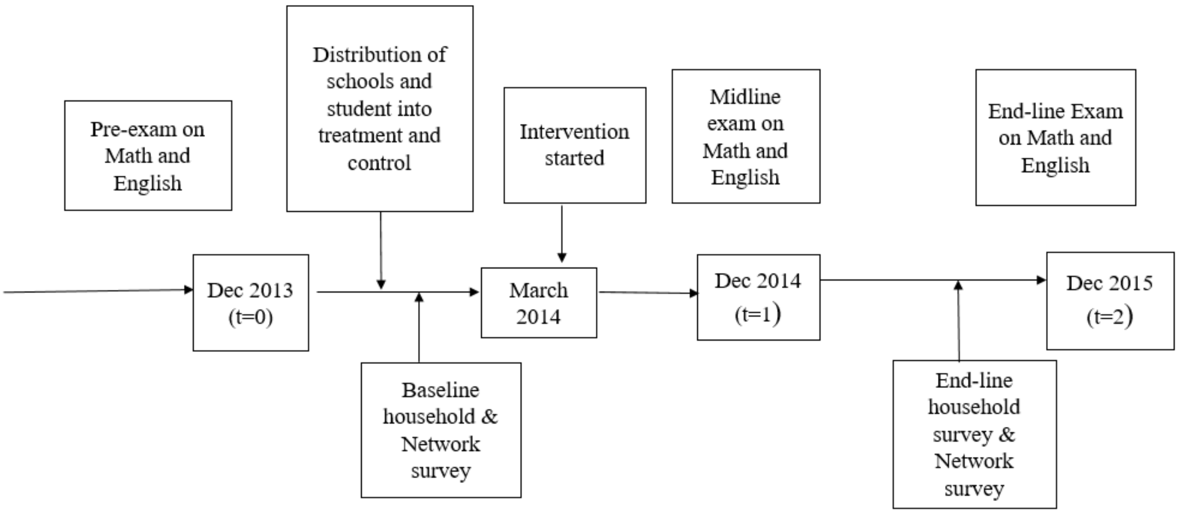


Figure A3: Timeline



A.2 Balance checks

Table A1: Balance checks across treatment and control schools

	T1	T2	T3	Control	T1=T2=T3=C
Test score-English (baseline)	0.093	0.090	0.167	0.109	0.114
Test score-math (baseline)	0.174	0.084	0.112	0.131	0.270
Student gender	0.50	0.49	0.50	0.49	0.347
Age of household head	39.03	39.12	38.90	39.03	0.651
Literacy of household head	0.63	0.63	0.62	0.63	0.648
Occupation of household head in agriculture	0.62	0.63	0.61	0.61	0.440
Household Income	6352.1	6267	6337.7	6331.7	0.658
Number of children	1.63	1.65	1.66	1.66	0.112
Observations	3297	2971	2677	5271	

Notes: Student gender is a dummy that equals to 1 for boys, and 0 for girls. Literacy is a dummy that equals to 1 if a father can both read and write, it is 0 if he cannot. Occupation of household head is a dummy that equals to 1 if it is agriculture related, it is 0 if not. Household income is the monthly household income in taka. Means are reported for each baseline characteristic. ANOVA tests are conducted for continuous baseline variables including age, household income and household size. Chi square tests are implemented if it is a dummy variable including student gender, literacy of household head, and occupation of household head.

Table A2: Balance checks across treated and untreated students within treatment schools

	Treated	Untreated	p-values
	(1)	(2)	(1)=(2)
Test score-English (baseline)	0.133	0.069	0.393
Test score-math (baseline)	0.147	0.073	0.350
Student gender	0.53	0.53	0.700
Age of household head	39.14	38.74	0.459
Literacy of household head	0.626	0.626	0.991
Occupation of household head	0.612	0.638	0.381
Household Income	6262.95	6455.07	0.504
Number of children	1.65	1.62	0.422
Observations	6,312	2,633	

Notes: Student gender is a dummy that equals to 1 for boys, and 0 for girls. Literacy is a dummy that equals to 1 if a father can both read and write, it is 0 if he cannot. Occupation of household head is a dummy that equals to 1 if it is agriculture related, it is 0 if not. Household income is the monthly household income in taka. Means are reported for each baseline characteristic. ANOVA tests are conducted for continuous baseline variables including age, household income and household size. Chi-square tests are reported for dummy variables including student gender, literacy of household head, and occupation of household head.

Table A3: Balance check of school characteristics

	T1	T2	T3	Control	T1=T2=T3=Control
Headteacher characteristics					
Bachelor's degree	0.59	0.68	0.65	0.64	0.839
Male	0.79	0.95	0.81	0.81	0.090
Age	43.1	44.7	46.1	45.1	0.105
Experience	15.8	17.1	16.4	15.8	0.839
School characteristics					
Accommodation	0.33	0.48	0.35	0.33	0.251
Toilet facility	0.94	0.92	0.89	0.92	0.795
Electricity (school)	0.17	0.19	0.16	0.13	0.723
Electricity (area)	0.44	0.38	0.38	0.38	0.879
Wall-Brick	0.79	0.77	0.76	0.73	0.860
Roof-Concrete	0.81	0.79	0.76	0.71	0.528
Floor-Concrete	0.88	0.84	0.81	0.85	0.847
Number of rooms	4.5	4.6	4.5	4.3	0.933
Number of teachers	4.6	4.5	4.4	5.1	0.109
Share of male teachers	0.50	0.53	0.51	0.53	0.815
Class size	26.4	28.4	24.5	25.4	0.624
No. of schools	57	56	37	100	

Notes: Bachelor's degree is an indicator of whether a headteacher holds a bachelor's or higher qualification. Experience is the years of teaching experience of the headteacher in this school. Accommodation is an indicator for whether the school provides sufficient accommodation space for students. Toilet facility indicates whether the school has a toilet; Electricity(school) indicates whether the school itself has electricity facilities; Electricity (area) whether the area where the school is located has electricity facilities; Wall-Brick indicates whether the school's wall is made of brick; Roof-Concrete equals to 1 if the school's roof is made of concrete; Floor-Concrete equals to 1 of the floor is made of concrete. One way Anova tests are conducted to compare the group means across T1, T2 T3 and control schools for continuous variables. Chi square tests are conducted to compare the group means across T1, T2, T3 and control schools for indicator variables. Means are reported for each variable.

A.3 Missing Test Scores

Table A4: Overview of Missing Test Scores, by Treatment

	T1	T2	T3	Control	Total
Baseline Sample	3,297	2,971	2,677	5,271	14,216
Missing Endline Test Score	678	522	478	1,208	2,886
Total	3,975	3,493	3,155	6,479	17,102
Missing %	17.1%	14.9%	15.2%	18.6%	16.9%

Table A5: Differences in baseline test scores of students with missing endline test scores

	English	Math
T1	-0.142	-0.026
	(0.121)	(0.106)
T2	-0.023	-0.094
	(0.122)	(0.091)
T3	-0.105	-0.099
	(0.124)	(0.117)
N	2,886	2,886

Note: The sample consists of students with missing endline scores. Reported coefficients are estimated from a regression of test scores in English (col 1) and Math (col 2) on treatment indicators. Pairwise tests (T1=T2; T1=T3; T2=T3) fail to reject equality of coefficients, in both regressions. Standard errors are clustered at school level and reported in parentheses.

A.4 Treatment Take-up

Table A6: Treatment take-up rate

	T1	T2	T3	Total
Invited	2,647	2,358	1,307	6312
Attended	2,553	2,335	1,199	6087
Take up rate	96.45%	99.02%	91.74%	96.44%

Notes: Invited denotes students who received the invitation to be treated. Attended denotes students who received the invitation and actually participated.

A.5 Overall Treatment Effect

Table A7: Overall treatment effect: English & math test scores

	English			Math		
	2013	2014	2015	2013	2014	2015
	(1)	(2)	(3)	(4)	(5)	(6)
T_1	-0.041	0.306***	0.359***	0.010	0.457***	0.451***
	(0.079)	(0.080)	(0.087)	(0.074)	(0.086)	(0.071)
T_2	-0.0396	0.454***	0.857***	-0.0777	0.658***	0.939***
	(0.068)	(0.073)	(0.099)	(0.085)	(0.074)	(0.089)
T_3	0.0394	0.379***	0.649***	-0.0450	0.606***	0.728***
	(0.071)	(0.088)	(0.095)	(0.076)	(0.090)	(0.080)
$y_{i,g,s,0}$		0.279***	0.216***		0.248***	0.211***
		(0.0147)	(0.0172)		(0.0158)	(0.0151)
Observations	14,216	13,379	14,216	14,216	13,379	14,216
No. of schools	254	251	254	254	251	254
H_0: $T_3 = T_1$	0.385	0.351	0.001	0.554	0.112	0.000
H_0:$T_3 = T_2$	0.322	0.284	0.043	0.747	0.527	0.020
H_0:$T_1 = T_2$	0.991	0.014	0.000	0.386	0.010	0.000

Note: The sample includes all students in T1, T2, T3 treatment schools and control schools. All columns include controls for class level, student gender, age of father, literacy of father, household income and number of children, and a dummy indicating whether any baseline household characteristics are missing. Standard errors are clustered at school level and reported in parentheses. The Table also reports p-values of pairwise tests of equality of coefficients on the treatments. ***p<0.01, **p<0.05, *p<0.1.

Table A8: Overall treatment effect on noncognitive skills

	Self Control	Motivation	Social skills	Index
	(1)	(2)	(3)	(4)
T_1	0.070	0.260***	0.214**	0.181**
	(0.101)	(0.100)	(0.092)	(0.076)
T_2	0.148	0.435***	0.316***	0.300***
	(0.108)	(0.114)	(0.095)	(0.092)
T_3	0.132	0.249**	0.234** *	0.205**
	(0.111)	(0.110)	(0.118)	(0.0920)
Observations	10,991	10,991	10,991	10,991
No. of schools	253	253	253	253
$T_3 = T_1$	0.567	0.923	0.878	0.803
$T_3 = T_2$	0.895	0.166	0.548	0.388
$T_1 = T_2$	0.466	0.164	0.371	0.218

Notes: The sample includes all students in T1, T2, T3 treatment schools and all students in control schools for whom we have measures of noncognitive skills. All columns include controls for grade level, student gender, age of father, literacy of father, household income and number of children, and a dummy indicating whether any baseline household characteristics are missing. Standard errors are clustered at school level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1.

A.6 Lee bounds

Table A9: Lee bounds analysis: Overall treatment effect

	English		Math	
	2014	2015	2014	2015
	(1)	(2)	(3)	(4)
	English		Math	
Treatment	0.390***	0.610***	0.584***	0.698***
	(0.076)	(0.080)	(0.072)	(0.067)
Lee bounds	[0.285, 0.457]	[0.574, 0.685]	[0.460,0,659]	[0.632, 0.771]
Observations	13,379	14,216	13,379	14,216
No. of schools	251	254	251	254

Notes: The sample includes all students in T1, T2, T3 treatment schools and control schools. $Treatment$ is an indicator variable for attending one of the three treatment (T1, T2, or T3) schools. All columns include controls for binary covariates including baseline test scores (above/below median), student gender, grade level, parent age (above/below median), parent occupation, parent income (above/below median) and number of children (above/below median).

Table A10: Lee bounds analysis: Spillover effects

	(1)	(2)	(3)	(4)
	2014	2015	2014	2015
	English		Math	
Treatment	0.252***	0.523***	0.425***	0.616***
	(0.0866)	(0.0989)	(0.0870)	(0.0881)
Lee bounds	[0.129, 0.411]	[0.360, 0.665]	[0.287, 0.533]	[0.453, 0.730]
Observations	7,202	7,904	7,202	7,904
No. of schools	183	206	183	206

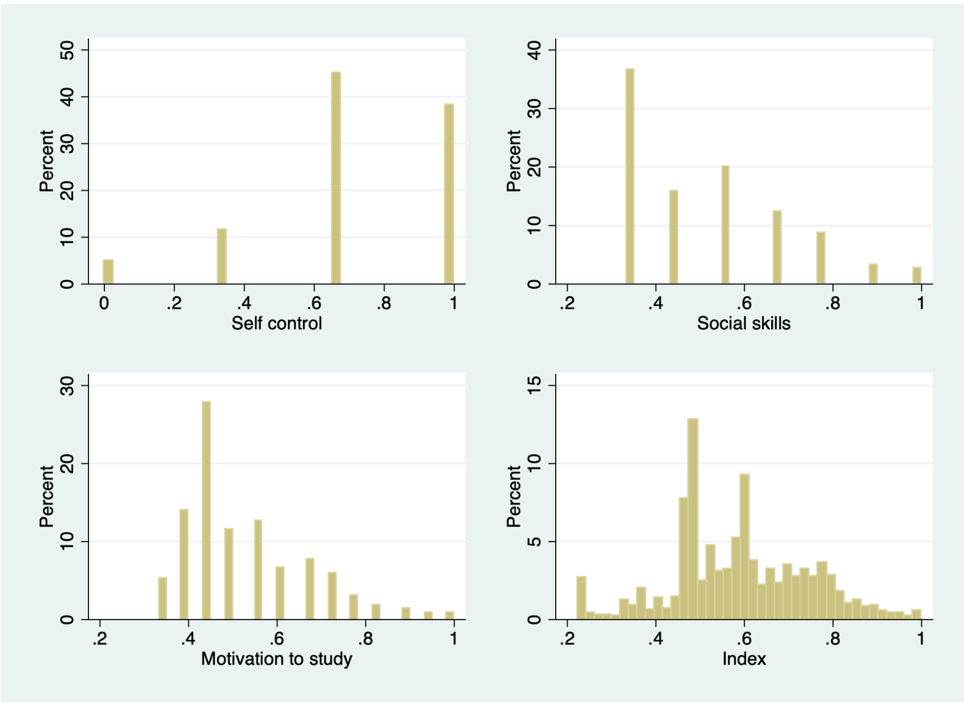
Notes: Treatment is an indicator variable for attending one of the three treatment (T1, T2, or T3) schools. All columns include controls for binary covariates including baseline test scores (above/below median), student gender, grade level, parent age (above/below median), parent occupation, parent income (above/below median) and number of children (above/below median).

A.7 Definition of noncognitive skills

Table A11: Definition of noncognitive skills

	Definition	Measurement
Self Control	Is the student regular in class?	1=Yes, 0=No
	Does the student regularly do his homework?	1=Yes, 0=No
	Does the student regularly do his homework?	1=Yes, 0=No
Social skills	How adaptive is the student in making friends?	1=Moderate, 2=Much, 3=Very much
	How adaptive is the student in sustaining relationships?	1=Moderate, 2=Much, 3=Very much
	How does the student behave with classmates?	1=Competitive, 2=Irreconcilable 3=Friendly
Motivation to study	How much attentive is the student in class?	1=Moderate, 2=Much, 3=Very much
	How desirous of study is the student?	1=Moderate, 2=Much, 3=Very much
	How eager is the student in learning lessons?	1=Moderate, 2=Much, 3=Very much
	Involvement in taking challenging job	1=Moderate, 2=Much, 3=Very much
	How much does the student enjoy doing challenging jobs	1=Moderate, 2=Much, 3=Very much
	Does the student take challenging work?	1=Sometimes, 2=Often, 3=Very often

Figure A4: The distribution of noncognitive skills



Note: The sample includes all students in T1, T2, T3 and control schools. Each non-cognitive skill is measured at $t = 2$. Index is the average of the three non-cognitive skills.

A.8 Balance check of noncognitive skills sample

Table A12: Balance checks of characteristics within noncognitive skills sample

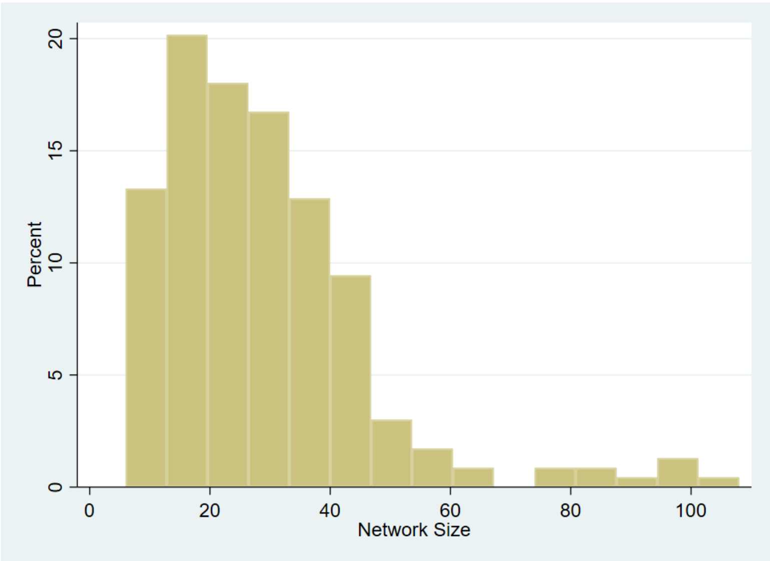
Variables	T1	T2	T3	Control	T1=T2=T3=C
Baseline english score	0.22	0.21	0.31	0.26	0.634
Baseline math score	0.31	0.21	0.30	0.31	0.710
Student gender	0.49	0.49	0.49	0.48	0.788
Age of father	38.98	39.16	38.72	39.14	0.943
Literacy of father	0.64	0.64	0.63	0.62	0.493
Occupation of father	0.60	0.62	0.60	0.59	0.891
Income of father	6440.4	6309.3	6421.3	6396.9	0.430
Number of Children	2	2	2	2	0.229
Observations	3,928	2,627	2,398	2,038	

A.9 Network statistics

Table A13: Network size

Mean	29
Standard deviation	17
Median	26
Variance	304
Skewness	1.7
Minimum	6
Maximum	108
Observations	232

Figure A5: The distribution of network size



A.10 Eigenvector centrality exposure ($\overline{E}^{T}_{g,s,0}$)

Figure A6: The distribution of centrality exposure ($\overline{E}^{T}_{g,s,0}$) in T1 and T2 schools

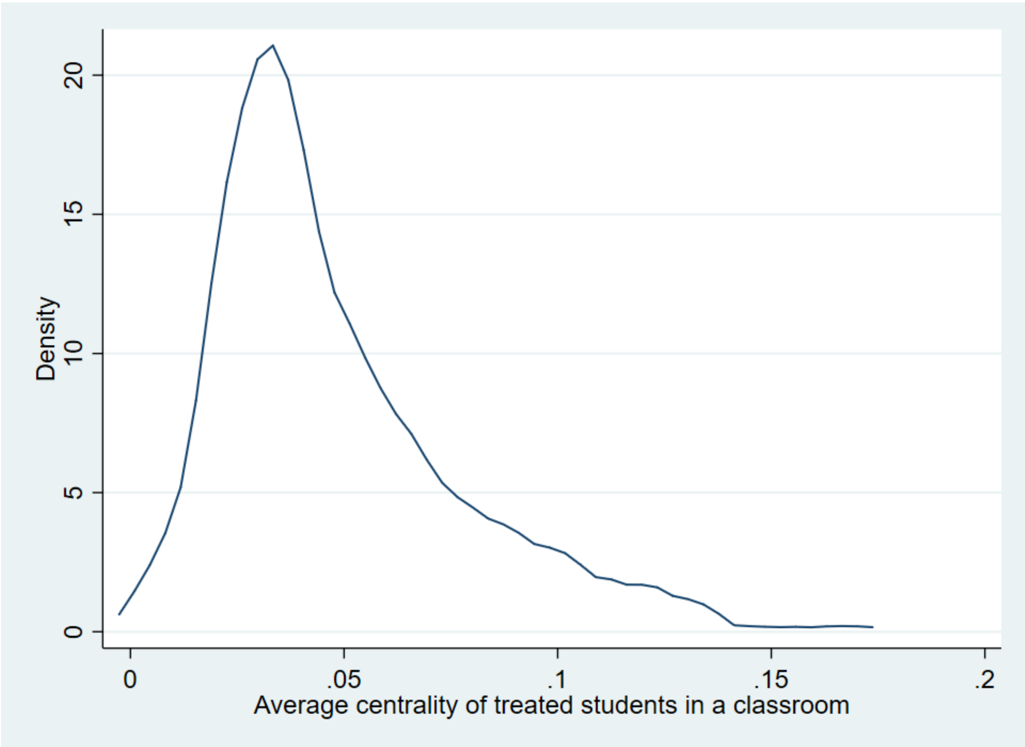


Table A14: Balancing test of centrality exposure ($\overline{E}^T_{g,s,0}$)

	(1)	p-value
$y_{i,g,s,0}$ (Eng)	0.851	0.890
	(6.145)	
$y_{i,g,s,0}$ (Math)	-1.839	0.705
	(4.839)	
Grade	7.692	0.694
	(19.47)	
T_2	6.951	0.177
	(5.096)	
Student Gender	8.184	0.60
	(4.288)	
Age of Household Head	0.528	0.314
	(0.521)	
Literacy of Household Head	-0.525	0.950
	(8.418)	
Occupation of Household Head	7.424	0.142
	(5.005)	
Number of Children	9.702*	0.06
	(4.889)	
Household Income	-0.004**	0.03
	(0.001)	
Observations	1,263	
No. of schools	71	

Notes: The Table reports coefficients from a regression of normalized centrality exposure rate in a class separately on each of the variables reported in the rows. The regressions also include a control for whether there are any missing household characteristics. Standard errors are clustered at school level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1.

B Theoretical model

B.1 Explaining the correlation between test score or non-cognitive skills and eigenvector centrality

In equilibrium, each agent maximizes her utility and the best-response function is given by:

$$y_{i,t} = \alpha_1 \sum\nolimits_{j \in \mathcal{N}} g_{ij} y_{j,t} + \pi_i, \tag{B.1}$$

or, in matrix form,

$$\mathbf{y}_t := \mathbf{y}_t(\mathbf{g}) = \alpha_1 \mathbf{G} \mathbf{y}_t + \pi \tag{B.2}$$

where $\pi = (\pi_1, \cdots, \pi_n)'$ and π_i is defined in Equation (7). Let $\rho(\mathbf{G})$ denote the spectral radius of the adjacency matrix $\mathbf{G}$, and $\mathbf{I}_n$ denote the $n \times n$ identity matrix. Then, if $\alpha_1 \rho(\mathbf{G}) < 1$, then $\mathbf{I}_n - \alpha_1 \mathbf{G}$ is nonsingular and the network game with the utility function (6) has a unique Nash equilibrium in pure strategies with the equilibrium effort vector $\mathbf{y}_t^* = (y_{1,t}^*, ..., y_{n,t}^*)'$ given by:

$$\mathbf{y}_t^* = (\mathbf{I}_n - \alpha_1 \mathbf{G})^{-1} \pi = \mathbf{b}_\pi (\alpha_1, \mathbf{G}), \tag{B.3}$$

where $\mathbf{b}_\pi (\alpha_1, \mathbf{G})$ is the vector of *weighted Katz-Bonacich centralities.* Furthermore, if $\alpha_1 \geq 0$ and $\pi \geq \mathbf{0}$, then the elements of the Neumann series

$$\mathbf{y}_t^* = (\mathbf{I}_n - \alpha_1 \mathbf{G})^{-1} \pi = \sum\nolimits_{k=0}^{\infty} (\alpha_1 \mathbf{G})^k \pi$$

are nonnegative and, hence, the equilibrium efforts $\mathbf{y}_t^* \geq \mathbf{0}$.

Let $\rho(\mathbf{G})$ denote the spectral radius of the adjacency matrix $\mathbf{G}$ and $\mathbf{b}_\pi (\alpha_1, \mathbf{G})$ is the vector of *weighted Katz-Bonacich centralities.*

Proposition 1. *If $\alpha_1 < 1/\rho(\mathbf{g})$, there exists a unique Nash equilibrium for which $\mathbf{y}_t^* = \mathbf{b}_\pi (\alpha_1, \mathbf{G})$, that is, each student i at time t exerts an effort proportional to her weighted Katz-Bonacich centrality. If $\alpha_1 \rightarrow (1/\rho(\mathbf{g}))^-$, then the measure of* weighted Katz-Bonacich centralities *converges to the standard eigenvector measure of centrality defined in (3), that is*

$$\rho(\mathbf{g}) \mathbf{y}_t = \mathbf{G} \mathbf{y}_t$$

where $\mathbf{y}_t^* := \mathbf{y}_t^*(\mathbf{g}) := \mathbf{E}_t(\mathbf{g})$ is the vector of eigenvector centralities in network $\mathbf{g}$.

Proof of Proposition 1: From these equations, using Ballester et al. (2006), we can prove the first part of the proposition showing that there exists a unique Nash equilibrium for which $\mathbf{y}_t^* = \mathbf{b}_\pi (\alpha_1, \mathbf{G})$.

To prove the second part of the proposition, that is, if $\alpha_1 \rightarrow (1/\rho(\mathbf{g}))^-$, then the measure of *weighted Katz-Bonacich centralities* converges to the standard eigenvector measure of centrality, we can use Bonacich & Lloyd (2001). ∎

B.2 A model with treated and untreated students

The best-reply functions for the treated and untreated students are respectively given by (12) and (13). In matrix form, we have

$$\mathbf{y}_t = \boldsymbol{\pi} + \alpha_1 \mathbf{A}^{Intra}\mathbf{G}^{Intra}\mathbf{y}_t + \alpha_1 \mathbf{A}^{Inter}\mathbf{G}^{Inter}\mathbf{y}_t$$

where

$$\mathbf{y} = \begin{pmatrix} y_1^T \\ y_2^T \\ \vdots \\ y_n^{NT} \end{pmatrix}, \quad \boldsymbol{\pi} = \begin{pmatrix} \pi_1^T \\ \pi_2^T \\ \vdots \\ \pi_n^{NT} \end{pmatrix}, \quad \mathbf{A}^{Intra} = \begin{pmatrix} \theta & 0 & \dots & 0 \\ 0 & \theta & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \quad \mathbf{A}^{Inter} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \theta \end{pmatrix}$$

$\mathbf{A}^{Intra}$ is a diagonal matrix where, on the diagonal, the first n^T rows have a θ while the remaining n^{NT} rows have a 1. For $\mathbf{A}^{Inter}$, it is exactly the opposite, that is, on the diagonal, the first n^T rows have a 1 while the remaining n^{NT} rows have a θ.

Proposition 2. *If $\alpha_1 \rho \left(\mathbf{A}^{Intra}\mathbf{G}^{Intra} + \mathbf{A}^{Inter}\mathbf{G}^{Inter} \right) < 1$, the peer effect game with payoffs (10) and (11) has a unique interior Nash equilibrium in pure strategies given by:*

$$\mathbf{y}_t = \left(\mathbf{I} - \alpha_1 \mathbf{A}^{Intra}\mathbf{G}^{Intra} - \alpha_1 \mathbf{A}^{Inter}\mathbf{G}^{Inter} \right)^{-1} \boldsymbol{\pi} \tag{B.4}$$

Proof: We need to show that $\mathbf{I} - \mathbf{B}$ is non-singular (i.e. invertible), where $\mathbf{B} \equiv \alpha_1 \mathbf{A}^{Intra}\mathbf{G}^{Intra} - \alpha_1 \mathbf{A}^{Inter}\mathbf{G}^{Inter}$. We know that $\mathbf{I} - \mathbf{B}$ is non-singular if $\alpha_1 \rho \left(\mathbf{A}^{Intra}\mathbf{G}^{Intra} + \mathbf{A}^{Inter}\mathbf{G}^{Inter} \right) < 1$ (see, e.g., Meyer (2000), page 618). The interiority of the solution is straightforward since we assumed that $\pi_i^T > 0$ and $\pi_i^T > 0$, for all i. ∎

This is an interesting result because it connects the adjacency matrix $\mathbf{G}$ to the split structure of peer effects (inter and intra-peer effects) and it is directly comparable to the condition given in Ballester et al. (2006), i.e. $\theta \rho(\mathbf{G}) < 1$, where peer effects were assumed to be the same across all agents.