Teams, Organization and Education Outcomes: Evidence from a field experiment in Bangladesh[*]

Youjin Hahn[†] Asadul Islam[‡] Eleonora Patacchini[§] Yves Zenou[¶]

May 26, 2015

Abstract

We study the relationship between network centrality and educational outcomes using a field experiment in primary schools in Bangladesh. After obtaining information on friendship networks, we randomly allocate students into groups and give them individual and group assignments. We find that the groups that perform the best are those whose members have high Katz-Bonacich and key-player centralities. Leaders are mostly responsible for this effect, while bad apples have little influence. Group members' network centrality is also important in shaping individual performance. We show that network centrality captures non-cognitive skills, especially patience and competitiveness.

JEL Classifications: A14, C93, D01, I20.

Keywords: Network centrality, team work, leaders, soft skills.

---

[*]We thank Simone Bertoli, Margherita Comola, Vianney Dequiedt, Carlo L. Del Bello, Quoc-Anh Do, François Fontaine, Ben Golub, Nicolas Jacquemet and David Margolis as well as the participants of the seminars at Monash University, CERDI (Clermont-Ferrand) and TEMA (Paris) for helpful comments and suggestions.

[†]Monash University, Australia. E-mail: Youjin.Hahn@monash.edu.

[‡]Monash University, Australia. E-mail: Asadul.Islam@monash.edu.

[§]Cornell University, EIEF and CEPR, USA. E-mail: ep454@cornell.edu.

[¶]Stockholm University and IFN, Sweden, and CEPR. E-mail: yves.zenou@ne.su.se.

1 Introduction

In contrast to the pervasive use of teams in organizations, few empirical studies have systematically examined the impact of teams on outcomes. Consequently, relatively little is known about the determinants of group performance when individuals interact in groups. Moreover, economists have long speculated about the importance of the individual position in a network of social contacts for decision making processes, but the existing evidence is not conclusive. The current challenge stems from the endogenous allocation of individuals into groups. Indeed, it is difficult to identify network effects when the network and the outcome are simultaneously chosen.

This study is the first to study the importance of network centrality in shaping performance in teams using randomly allocated groups.

We conduct a large scale field experiment in rural primary schools in Bangladesh. In June 2013, we collect students' friendship networks and conduct a separate household survey in 80 schools for fourth-graders to obtain information on demographic characteristics, education, family background and friendship. The friendship information is based upon actual friends' nominations. Pupils are asked to identify their best friends from a school roster (up to 10 nominations). We end up with 3,406 students distributed over 80 networks and schools. In July 2013, we randomly allocate students into groups of four in each classroom. We balance the group characteristics so that they have on average the same ability level. We then ask students to perform different tests on general knowledge and math, both in the very short run (i.e. when the groups are just formed) and in the longer run (a week after the groups have been formed). Some of the tests are performed collectively by the group while some others are performed individually. The tests performed at the time work teams are formed allow us to disentangle the effect of network topology from network (peer) effects in team outcomes. Indeed, at this stage of our experiment, students have no time to interact with each other, so that no performance spillovers (peer effects) are at work. The tests performed at a later stage would then reveal if the idiosyncratic characteristic, which is measured by network centrality, is still relevant in shaping outcomes after a week of team work.

Our analysis mainly focuses on two questions. First, we examine whether and how the centrality of members of randomly assigned groups affects the group and the individual test scores. Second, having found evidence of the existence of some effects, we then investigate their underlying mechanisms.

Our study enables us to get inside the black box of network centrality measures. Our conjecture is that network centrality measures, especially Katz-Bonacich and key-player centralities, capture some of the *non-cognitive* or *soft skills* of a student. As a result, groups with high key-player and Katz-Bonacich centrality are the groups that include popular and social students who can promote study activities and learning practices amongst group members, enforce decisions and stimulate discussions. To investigate if, indeed, these centrality measures (and others) capture the soft skills of the students, we enrich further our experiment by making our students play some games that capture personal traits such as risk-taking behavior, patience and competitiveness.

The elaboration of the data collected from our experiment shows the following evidence. Firstly, the *Katz-Bonacich* and *key-player average centrality* of the group have a significant and positive impact on the *group performance* when the task is completed at the time of the team formation. The effect is not mediated by the presence of unobservable factors affecting both friendship formation and school performance as the results hold true when we repeat our analysis when eliminating the groups that (by chance) contain friends. Interestingly, we find that these two centrality measures also matter for the longer-run group outcomes, when longer run means here one week after the groups were randomly formed. The centrality of other students in a group is also important for *individual performance.* When looking at the individual improvement in math scores after a week of team work, we find a significant impact of the team members' centrality on individual performance.

We also find that leaders play an important role in group outcomes in the short and longer run. In other words, it is often the student with the highest centrality in the group who affects the collective performance of the group. Leaders can also weaken the effect of own centrality on individual outcomes. On the other hand, the student with the lowest centrality in the group (the "bad apple") has nearly no impact on both the group and the individual performances.

When looking at the data collected from our games on students' soft skills, we find strong correlations between patience and competitive behavior of our students and their key-player centrality as well as their Katz-Bonacich centrality. Importantly, we also test whether each centrality captures the *cognitive skills* of the student as measured by their *ability* (captured by the test score at the individual math test taken before the experiment). We find that, for all centrality measures, there is *no* significant correlation between centrality measures and ability, which seems to indicate that centrality measures do capture non-cognitive skills (or soft skills) rather than cognitive skills. These results suggest that the personal traits of students that are conductive of high-quality collective and individual works are patience and competitiveness. Our analysis does not provide any evidence suggesting that a student's risk-taking behavior could help students work efficiently both in groups and individually.

The rest of the paper unfolds as follows. In the next section, we relate our paper to the literature and highlight our contribution. In Section 3, we give some background of the educational system in Bangladesh and describe the experimental design. Section 4 describes our data. In Section 5, we detail the regressions that we estimate and discuss our empirical results. In Section 6, we investigate the mechanisms behind our results by studying the extent to which centrality measures capture cognitive or non-congnitive skills. Finally, Section 7 concludes.

2 Related literature

Our paper is related to different literatures.

Peer effects in education

An important literature on peer effects and education looks at the causal impact of peers on educational outcomes using field experiments (see e.g. Angrist and Lavy, 1999; Sacerdote, 2001; Zimmerman, 2003).[1] For example, Sacerdote (2001) and Carrell et al. (2009) study specific contexts in which first-year roommates (or hallmates or squadron mates in the case of military academies) are randomly assigned by the housing office. This creates exogenous variation in one's peer group, which is then used to ask how much peers matter, which peers matter, and for what outcomes. They find strong peer effects in education. A recent paper by Carrell et al. (2013) examine squadron mates in the case of military academies. They manipulate the groups by putting together low-ability and high-ability incoming cadets at the US Air Force Academy. They show that performance for the lower-ability students fell relative to lower-ability students in the randomly assigned control group.

Compared to this literature, we study *network topology effects* rather than *peer* effects. While adopting the same methodology that ensures a random allocation of the peers, we consider the impact of *network centrality measures* on educational outcomes.

Non-cognitive versus cognitive skills

Another important literature in education is the one about the distinction between *cognitive skills* (grades) and *non-cognitive skills* (soft skills). It is well documented that cognitive ability (or skill) is a strong predictor of education success. Economists, psychologists, and sociologists are now actively examining determinants of social and economic success beyond those captured by cognitive ability (see e.g. Borghans et al., 2008). For example, a recent analysis of the Perry Preschool Program shows that *personal traits* other than those measured by IQ and achievement tests causally determine life outcomes (see, in particular, Heckman

[1] For an overview of this literature, see Sacerdote (2011, 2014).

et al., 2010; 2011). Traits such as perseverance and preferences related to an interest in learning might lead people to attain more total years of schooling. Indeed some evidence suggests that this might be the case. For example, Heckman et al. (2006) show that better adolescent personality traits —as measured by locus of control and self-esteem— increases the probability of graduating from, and stopping at, high school for males at the lowest quantiles of the personality distribution. Cunha et al. (2010) use a dynamic factor model to investigate the development of both cognitive skills and personality traits during childhood, allowing for endogenous investment in skills and dynamic complementarities. They find that adolescent personality—as measured by a variety of behavior inventories— accounts for 12% of the variation in educational attainment, whereas adolescent cognitive ability accounts for 16% of the variation (see Almlund et al., 2011, for a review).

To summarize, this literature shows that soft skills (such as motivation, tenacity, trustworthiness, conscientiousness, honesty, and perseverance) predict success in life, that they causally produce that success, and that programs that enhance soft skills have an important place in an effective portfolio of public policies (Heckman and Kautz, 2012).

Soft skills, however, are difficult to measure. While the measures of soft skills used in these papers are self-reported, we adopt an *indirect* approach where the soft skills of the students are measured by their position (centrality) and thus their popularity in the network. We show, indeed, that the centrality of students in a network is correlated with student's soft or non-cognitive skills. Our results suggest that centrality and thus soft skills (mostly patience and competitiveness) do affect the performance of the students and that they should be taken into account in education policies.

Centrality in networks

Bavelas (1948) and Leavitt (1951) were among the first to use centrality to explain differential performance of communication networks and network members on a host of variables including time to problem solution, number of errors, perception of leadership, efficiency, and job satisfaction.[2] Following their work, many (non-economist) researchers have investigated the importance of the centrality of agents on different outcomes. Indeed, various studies from different disciplines have shown that centrality is important in explaining employment opportunities (Granovetter, 1974), peer effects in crime (Haynie, 2001), power in organizations (Brass, 1984), the success of open-source projects (Grewal et al., 2006) as well as workers' performance (Mehra et al., 2001). On the other hand, economists are familiar with the difficulty of ascertaining cause and effect in such complex constructs and

[2]The economics of networks is a growing field in economics. For recent overviews, see Ioannides (2012), Jackson (2008, 2011, 2014), Jackson et al. (2015), Jackson and Zenou (2015) and Topa and Zenou (2015).

have remained dubious about the identification of the effects. The main challenge is the endogeneity of the network.[3]

A possible way out is to use controlled experiments. In the field of networks, this has been implemented by either (*i*) fully controlling the network of relationships in the laboratory (Choi et al., 2005; Kearns et al., 2009) or (*ii*) assigning subjects in field experiments in a network through which they must communicate (Centola, 2010, 2011; Goeree et al., 2010; Babcock and Hartman, 2010; Cai et al., 2015). In the present paper, we use a *field experiment* where agents embedded in a friendship network are randomly allocated to groups. We believe this is the first paper that uses such a strategy in the context of networks to evaluate the *causal* impact of centrality on outcomes.

Leadership and social networks

There is a literature in management and psychology that increasingly recognizes the importance of social processes and relational linkages in shaping leadership. In addition to resources that stem from human capital, organizational capacities can also be derived from social relationships - the so-called *social capital* (Putnam, 2000) In economics, while studies on the effects of social networks on a variety of outcomes are pervasive, the intersection between leadership and social networks has received limited attention. Some recent papers look at some related issues. Tao and Lee (2014) define peer pressure using some extreme order statistics (i.e. maximum value of the peer performance) rather than using the average level of activity. Using data from German 9th-graders, Tatsi (2014) find that bad apples are more important than classroom stars in affecting educational outcomes. Mastrobuoni and Patacchini (2012) document that network centrality, especially eigenvalue centrality, is an important predictor of leadership in the US mafia organization.

In line with this latter strand of research, in this paper, we use student popularity, as measured by network centrality, to define leaders (i.e. those having the highest value of centrality in a friendship network). This paper is the first to assess the importance of the

[3] Some economists have addressed the question about 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 and key-player centrality (defined in Appendix 5) have proven useful in describing the activity of each agent (Ballester et al., 2006; Calvó-Armengol et al., 2009; Liu et al., 2012, 2015; Lindquist and Zenou, 2014; König et al., 2014; Zenou, 2015). 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 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. In all these papers, the endogeneity of the network is tackled using a structural approach.

presence of a leader in a group in shaping group performance, as well as the performance of each group member.

Peer effects in productivity

There is a limited literature on peer effects in productivity. Guryan et al. (2009) consider performance of professional golfing pairs, where their parings are randomly assigned and the pairings are competitors not teammates. Bandiera et al. (2009) investigate how social connections between workers and managers affect the productivities of fruit pickers in the United Kingdom. Their measure of social connectedness is based on similarities of worker/manager characteristics (e.g., nationality) and there are multiple managers whose worker assignments change daily. They also conduct a field experiment where they exogenously vary manager compensation schemes, allowing them to perform a more nuanced analysis of the effect of worker/manager connectedness on worker productivity.[4] Mas and Moretti (2009) consider peer effects in the performance of supermarket cashiers, where pairing of peers is assumed to be exogenous, but do not specifically employ teams or networks in their analysis. Hamilton et al. (2003) analyze the effect of team composition on clothing manufacturing, providing support for the view that teams utilize collaborative skills, which are less valuable in individual production. Moreover, they consider and explicitly model individual self-selection into teams. Using different field experiments and different outcomes for students at the University of California-Santa Barbara, Babcock et al. (2015) find large team effects that operate through social channels. Finally, Horrace et al. (2015) consider the estimation of productivity spillovers in a network production function where the allocation of workers into teams is strategic and done by a manager.

Our study is the first to investigate the impact of team composition in terms of individual popularity (i.e. centrality) on individual and group performance.

To sum-up, our contributions to these literatures are: (*i*) We show how the group composition in terms of group members' network centrality as well as the presence of the leader affect the individual and the group outcomes both in the short and long run; (*ii*) We show that these centrality measures, in particular Katz-Bonacich and key player centrality, capture individual soft skills, mostly patience and competitiveness.

[4] See also Bandiera et al. (2010) who test the value of social connections in an observational design. They show that workers are more productive if they have more social ties to their co-workers.

3 Experimental design

3.1 Background and context

We conduct field experiments in rural primary schools in Bangladesh. Primary education is free and compulsory for children aged six to ten years (grades 1-5).[5] However, it is not enforced, and most parents traditionally would prefer to send their kids to work or keep them at home, in particular females. This tendency has changed over recent decades. In 1993, the government introduced the food for education (FFE) Program to support poor children in completing primary schooling. In 2002, the primary education stipend project replaced the FFE Program. It was designed to provide cash transfers to families that kept children enrolled in primary school, with a minimum attendance level. In addition, a variety of policies such as the elimination of school fees and the provision of free textbooks have been put in place to encourage school enrolment (Mahmud et al., 2013). Nowadays, at the primary level most kids are going to school. According to Bangladesh Bureau of Educational Information and Statistics (2012), the net enrolment increased from 81.8 percent in 2000 to 96.4 percent in 2012. Bangladesh has also achieved tremendous progress over recent decades in reducing gender disparity in enrolment rates. From 1990 to 2009, the gender parity index (ratio of girls to boys) increased from 0.83 to 1.01 in primary schools and from 0.51 to 1.07 in secondary schools.[6]

The school drop-out rates and grade repetition rates, however, remain extremely high. In 2010, nearly 50 percent of primary school students drop out before completing grade five, and 12.5 percent repeat a grade. The quality of education is the greatest concern, as there is still a large gap between expectations and achievements. According to the Department of Primary Education of Bangladesh, around 70 percent of children who complete primary education are unable to read, write or count properly. As in many other developing countries, these deficiencies are attributed to high teacher and student absenteeism, low classroom teaching time, and high student-teacher ratio. At primary level, teacher–student ratio is 1 to 50.

[5] The education system in Bangladesh is broadly divided into three major levels: primary, secondary and tertiary education. Primary education is free and compulsory for children aged six to ten years (grades 1-5). The secondary education has three sub-stages: junior secondary (grade 6-8), secondary (grade 9 and 10), and higher secondary (grade 11 and 12). The main forms of provision include government primary schools, registered non-government primary schools (private schools), ibtidayee madrasahs (religious schools) and NGO schools. The majority of private schools are registered as non-government primary (RNGP) schools. The RNGP schools follow the same curriculum as government schools.

[6] In mid-1990s, the government also introduced a stipend program in secondary schools in rural area targeting female children. The effectiveness of this policy has been studied by Begum et al. (2013).

Any intervention aiming at improving our understanding of the factors fostering educational outcomes and teaching practices is thus of paramount importance.

3.2 Experimental plan

We conduct our experiments among *grade-four students* in 80 randomly chosen schools located in two districts, Khulna and Satkhira, in Bangladesh.[7]

Figure 1 describes the timing of our experiment.

In June 2013 (referred to as period $t-1$), we collect information on social contacts by asking each student to name up to 10 closest friends from a school roster, in order of importance. *All* students are interviewed in each of the 80 schools. Each of them also performs a math test (*individual pre-experiment math test*, IPEMT) to assess their ability. It is a multiple-choice test, which contains 15 questions measuring numbering and number-comparison skills, numeral literacy, mastery of number facts, calculation skills, and understanding of concepts. Questions also include arithmetical reasoning, data addition, deduction, multiplication and division. Children have 20 minutes to complete the test. The test is developed by local educators and experts in the field of education. A detailed description of the IPEMT is contained in Appendix 1. At the same time, we gather information on family background using an household survey, which contains questions on parent education, parent age, parent occupation, and other household characteristics.

In July 2013 (referred to as period t), one month after, random groups of 4 students are formed in each school. Specifically, we rank students according to their *individual pre-experiment math test* (IPEMT). We then randomly select a student from each quartile of the IPEMT empirical distribution to form a group of size four. For all 80 schools, the ANOVA F-test fails to reject at the 5% level the hypothesis that the average test score (IPEMT) is the same between groups. This indicates that groups are balanced in terms of ability. No information on friendship links is used for group formation.

Newly formed groups are then asked to solve *collectively* a general knowledge test (*group general knowledge test,* GGKT). The GGKT consists of 20 multiple choices items that aimed to explore students' knowledge regarding national and international affairs, geographical aspects, current affairs, and sports (see Appendix 2). Students are allowed to discuss for 20 minutes to elicit the best answer. It is an exam-like situation and the students are not informed about the content of the test (nor about the existence of the test itself).

After the group general knowledge test is performed, each group is given a math test to

[7] There are more than 800 primary schools in these two districts, and a total of 64 districts in Bangladesh.

be completed *collectively* in one week time (*group math test*, GMT). The test consists in 10 questions or problems. While the questions reflect the contents of a grade-four mathematics textbook, they are not taken from the school's textbook. Some international mathematics tests (e.g., the National Assessment Program of Literacy and Numeracy (NAPLAN)) for students of their age are used instead (see Appendix 3). The problems reflect the knowledge, skills, understandings and capacities that are essential for every child to learn mathematics at grade level four. The tests are developed in consultation with retired school teachers and local educational experts.

At the end of the week (referred to as period $t + 1$ in Figure 1), after each group had completed the GMT, each individual is also asked to take an *individual* post-experiment math t*est* (IPOMT). This test is developed on the basis of the group math test (GMT). Although none of the GMT questions is repeated, the structure of the problems is similar. Therefore, the completion of the GMT actually helps students to answer the IPOMT. There are 10 problems in the test and 1.5 hours is allocated to respond to these problems (see Appendix 4).

Finally, students are asked questions about their study behavior and their interactions as a group during the week, including the number of meetings and amount of time spent together to solve the GMT, as well as the total number of hours spent studying individually (all subjects).

Prizes are given to the most successful students. Three prizes are given in each class. The first prize is given to the *best performing group in the GGKT*. Each student of this group receives a pencil box scale (ruler). The two other prizes are based on the performance of the students in the individual math tests at the IPOMT. One prize is given to the *absolute best performing group in the IPOMT*, that is the group that has the highest aggregate grade (i.e. the sum of the grades of the team members) at the IPOMT. The other prize is given to the *relative best performing group in the IPOMT*, that is the group that has the highest average increase between the IPEMT (period t) and the IPOMT (period $t + 1$). These two prizes are given to make sure that students who are not that smart work hard in order to increase their individual performance between the two math tests while the smart kids also get one prize for the absolute marks they receive in the IPOMT. For both prizes, each member of these groups receive an instrument box (geometry box) or diary and scale.

All stages of the experiment are carried out under the close supervision of specialized researchers. The enumerators and the other personnel working in the field who actually run the experiment in schools received a week-long training.

We run a pilot experiment in a few schools to make sure the students understand the experimental plan and that the tests are appropriate for their education level. The project received enormous supports from the school teachers and administration.

[*Insert Figure* 1 *here*]

4 Data Description

Our data set consists of 80 schools, 924 groups and 3,406 students.[8] Table 1 presents some information about our data. We see that there are roughly as many female as male students. The majority of the households in this region of rural Bangladesh lacks access to electricity and only 28% of the sample students have access to electricity at home. Parental education is measured as the maximum of mother's years of education and father's years of education. Parental educational attainment on average is 5 years, and illiteracy rate is high; about 40 percent of the parents are either illiterate or can sign only. Looking at descriptive statistics about individual and group performance in the different tasks, it appears a notable dispersion in terms of performance in our sample. To make them comparable, all our test scores have mean zero and a standard deviation of 1. Table 1 also shows information on the frequency of interactions. On average, students met roughly 3 times for about an hour. The time spent studying with teammates for the GMT is an important portion of total study hours.

[*Insert Table* 1 *here*]

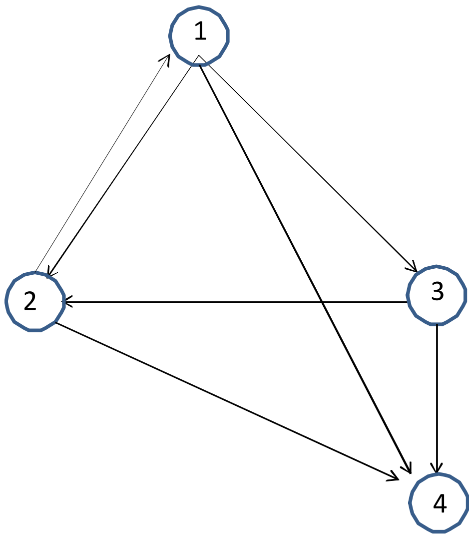
Regarding our definition of network centrality, we consider that the most popular students would be the ones who are nominated the most by other students. However, central to the design of our survey is the fact that the students are instructed to name friends in order of importance. Friends nominated first receive more importance. Formally, we denote a link from i to j as $g_{ij} \in [0, 1]$ if j has nominated i as his/her friend, and $g_{ij} = 0$, otherwise. Let us denote by d_i the number of nominations student i receives from other students, that is $d_i = \sum_j g_{ij}$. For each network, we can then define a matrix $\mathbf{W} = [w_{ij}]$, where each generic entry is defined as:

$$w_{ij} = 1 - \frac{(\vartheta - 1)}{d_i} \tag{1}$$

where ϑ denotes the order of nomination given by individual j to friend i in his/her nomination list. For example, consider a network of four students with the following nominations:

[8] There are some groups of size 3 for networks where the number of students could not be divided into groups of size 4.

(i) individual 1 nominates individual 2 first, then 3 and then 4; individual 2 nominates individual 1 first and then 4; individual 3 nominates individual 2 first and then individual 4; individual 4 nominates nobody. This network can be represented as follows:



The associated **W** matrix is given by:

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 2/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 1/2 & 0 \end{pmatrix}$$

Column 1 shows which person individual 1 has nominated and in which order. We can see that individual 1 has nominated first individual 2 (weight 1), then individual 3 (weight 2/3) and then individual 4 (weight 1/3). The same interpretation can be given for each column. By doing so, we are able to measure the (weighted) popularity of each individual. For example, individual 4 has nominated nobody but has been nominated by everyone, although never as the first person. If, as a measure of popularity, we just count the number of weighted links, then, even if 4 has the highest number of links, his/her popularity is lower than that of individual 2 who has only two links since $2 > 1/3 + 1/2 + 1/2$.

Using this definition of popularity, we calculate various measures of individual centrality that are used in the literature to capture the position of each individual in the friendship network (Wasserman and Faust, 1994; Jackson, 2008). There are many centrality measures and we will focus on the most prominent ones, which are formally defined Appendix 5. Let us give here some intuitive description. *Degree centrality* measures the number of links each agent has. As a result, it captures a simple measure of popularity. *Betweenness centrality* of a given agent is equal to the number of shortest paths between all pairs of agents that pass through the given agent. In other words, an agent is central if s/he lies on several shortest paths among other pairs of agents. Betweenness centrality thus captures the importance as an intermediary. Agent popularity as captured by betweenness centrality is related to the notion of *structural holes* developed by Burt (1992). He postulates that social capital is created by a network in which people can broker connections between otherwise disconnected segments of the network. Central agents according to Betweenness centrality have control over the flow of information in the network. *Closeness centrality* is a measure of how close an agent is to all other agents in the network. The most central agents can quickly interact with all others because they are close to all others. This measure of centrality captures how easily an individual reaches others. For example it captures how informed a given individual is in the context of information flows. *Eigenvector centrality* is a measure of the influence of an agent in a network. It takes all possible paths in a network (not only the shortest ones) and assigns relative scores to all agents in the network based on the concept that connections to high-scoring agents contribute more to the score of the agent in question than equal connections to low-scoring agents. It thus captures indirect reach so that being well-connected to well-connected others makes you more central. For example, Google's PageRank is a variant of the eigenvector centrality measure. The *Katz-Bonacich centrality* (due to Katz, 1953, and Bonacich, 1987) assigns a lower weight to nodes that are further away. As a result, Katz-Bonacich centrality captures the influence of friends and of their friends, with a discount rate. If there are strong network externalities (i.e. if the discount rate is close to one), it can be shown that Katz-Bonacich centrality becomes proportional to the eigenvector centrality (see Wasserman and Faust, 1994, Chap. 5.2). Finally, the *key-player centrality* (Ballester et al., 2006, 2010; Ballester and Zenou, 2014; Zenou, 2015) proposes a normative view of centrality. The key player is the agent who, once removed, generates the highest reduction is total activity in the network. In some sense, the key-player centrality (or *intercentrality*) shows how crucial an agent is in terms of the stability of the network. The main implication of this centrality is that the planner should target the key players in a network in order to change effectively the aggregate outcome.

Table 2a (Panel A) shows that there are large variations in individual centrality measures. It appears that the betweenness, closeness and degree centralities have quite small average values with large dispersion around this mean value. The maximum betweenness centrality is equal to 0.36, which means that one student has 36% of shortest paths that go through him/her.

[*Insert Table* 2*a here*]

Panel B of Table 2a collects summary statistics of network characteristics. One can see that the average size of our networks is roughly 51, which corresponds to the average size of the classroom. Indeed, in our sample every student is usually path-connected to any other student in the same classroom. Because there is one class per grade in each school, in our sample all fourth grades are in the same network and there is one network in each school. In fact, we see that the density of the network is quite low (20%), which is due to the relatively large size of networks. However, both the diameter and the average path length are quite small (4.7 and 1.7, respectively), indicating small world properties of these networks.

Table 2b, Panel A, reports summary statistics of the *average*, *maximum* and *minimum* centrality across groups. Panel B shows the values of these centralities for the subsample of groups that do not contain friendship links. This subsample will be used in a robustness check in Section 5.1. Figure 2 depicts the empirical distributions of the different centrality measures for all groups. It appears a notable dispersion, even though the distribution is skewed to the left for most of the centrality measures.

[*Insert Table* 2*b and Figure* 2 *here*]

The information on friendships is not used for the formation of our working teams in the experiment (period t in Figure 1). As a result, two friends could randomly be allocated to the same working team. This does not happen in the majority of groups. Indeed, Figure 3 shows the distribution of students by within-group nominations. Since there are 4 students in each group, each person can be nominated at most 3 times and thus has a maximum of 3 nominations. The figure shows that nearly 60% of the students has not been nominated by anyone else in the group. At the group level, Figure 4 shows that in roughly 20% of the groups there is not even one friendship relationship.[9]

[*Insert Figures* 3 *and* 4 *here*]

[9] More specifically, Figure 2 shows that there are 1,362 students (40% of 3,406) with no friends within the working team. Figure 3 shows that there are 166 groups (18% of 924) with no friends at all, which corresponds to roughly 664 students. The remaining 698 students with no friends within the working team are thus in groups where there is at least one friendship link (between other group members).

5 Empirical analysis

The aim of our empirical analysis is to investigate the extent to which group and individual performance is affected by the social skills of the members of the team, as measured by the individual position in the network of social contacts.

5.1 Group outcomes and network centrality

We estimate regressions of the form:

$$\overline{y}_{rs} = \beta_0 + \beta_1 \overline{C}_{rs} + \beta_2 \overline{A}_{rs} + \beta_3 \overline{X}_{rs} + \eta_s + \epsilon_{rs} \quad (2)$$

where $\overline{y}_{rs}$ is the *group test score*[10] of group r in school s in either the GGKT or the GMT, $\overline{C}_{rs}$ is the average centrality measure of group r (we look separately at each of the six centrality measures defined in Appendix 5), $\overline{A}_{rs}$ is the average ability of group r in school s (i.e. the average test score at the IPEMT), $\overline{X}_{rs}$ corresponds to the average observable *characteristics* of group r in school s (which includes gender, parent's education, access to electricity, etc.; see Table 1), η_s is the school fixed effects and ϵ_{rs} is an error term. Standard errors are clustered at the school level.

Observe that, as mentioned in the data description, in Bangladesh class size is large and there is only one fourth-grader class for each school. As a result, school fixed effects are here equivalent to network fixed effects since there is one (path-connected) network in each school.[11] School fixed effects capture all unobserved school specific factors. For example, if teacher quality differs between schools, then this is captured by school fixed effects- all fourth graders in a school will face the same teacher.

The OLS estimation results are displayed in the first two columns of Table 3. Each coefficient on average centrality is obtained from a different regression.

First, we see that the average *betweenness centrality* of the group has no impact on both the GGKT and the GMT. This suggest that the betweenness centrality of students is not a relevant factor in shaping educational outcomes in our schools in Bangladesh, a result also obtained by Calvó-Armengol et al. (2009) for students in the United States. Second, we see that the remaining centrality measures instead show a positive and significant impact. This is an important result, which indicates that, keeping ability constant, being randomly exposed to a group whose members have high centrality increases the group performance

[10] Since the test is performed *collectively* there is indeed one grade for each group.

[11] The use of network fixed effects to control for unobserved factors common to all network members in the analysis of peer effects is a traditional practice in the network literature (see e.g. Bramoullé et al., 2009).

both in the short run (when the group is just formed for the GGKT) and in the longer run (a week after the group is formed for the GMT).[12]

[*Insert Table* 3 *here*]

For robustness check, we also repeat our analysis when eliminating groups that (by chance) contain friends. Indeed, our identification strategy hinges on the random allocation of students into groups: students do not choose their team mates. Individual (and hence group) centrality is a pre-determined characteristics, which is exogenous to team performance. A threat to this identification strategy would be the possible presence of unobserved factors driving both group formation and school performance if two (or more) friends end up (by chance) in the same working team. Figure 4 shows that this happens in a few cases. We thus check whether our results hold true when removing those cases from our sample. Figure 4 shows that there are almost 20% of the groups where no friendship links exist. The last two columns of Table 3 contain the estimation results of equation (2) only for this subsample, i.e. groups for which members do not know each other directly. Although the variance of the centrality measures is largely reduced in this sub-sample (see Table 2b, panel B), the positive and significant impact of most centrality measures on the GGKT remain. Since the GGKT is performed just after the formation of the groups, when no interactions took place, these regressions identify a genuine effect of network topology on group outcomes as distinct from peer effects. When looking at the results for the GMT (performed a week after the teams were formed), we find that the effect of network centrality is not significant any more. This suggests that other factors, such as peer effects, may be more important in the longer run for group outcomes.

To further investigate the relevance of the number of friendship links on group outcomes, we estimate equation (2) by adding an extra control variable, that is the *fraction of friendship links within a group*. The results are reported in Table 4 when we run this exercise on the whole sample (columns (1) and (2)) and also on the subsample that excludes groups with no friendship links (columns (3) and (4)). For both samples, it appears that the estimated effects of network centrality on outcomes are similar to those in Table 3, both for the GGKT and the GMT. We also find that the fraction of friendship links in the group show no significant impact on the group outcomes GGKT and GMT.[13] As a whole, this evidence points towards an important role of network centrality on group outcomes as distinct from friends' influence.

[12]To save space, we do not report the effect of the average ability $\overline{A}_{rs}$ on the group outcomes $\overline{y}_{rs}$ (see (2)). We find that the average ability of the group has no significant impact on the GGKT or the GMT. This is not surprising given that the groups are balanced on their average ability.

[13]In order to capture non linear effects of friendship links, we also run a similar regression by introducing

[*Insert Table* 4 *here*]

The next interesting question is which centrality measure has the highest predictive power. To answer this question, we test the explanatory power of each centrality measure against each other. In Table 5, we report the correlations between our six measures of centralities. We see that the Katz-Bonacich and the key-player centrality show an almost perfect correlation (correlation of 0.954). The correlations between the other centrality measures range between 0.2 and 0.8. Observe also that Katz-Bonacich and key-player centralities are the only centrality measures that are microfounded through a model of social interactions (see Appendix 5). While their microfoundation is somehow different, both of them are a function of the strength of interactions within a network and consider the entire network topology in shaping individual centrality in a recursive manner (Ballester et al., 2006, 2010). The extremely high correlation of these two measures in our case implies that we cannot really distinguish their relative importance. However, our analysis can shed some light on the relative importance of measures stemming from a behavioral foundation (Katz-Bonacich and key-player centralities) and those that merely depend on network topology (degree, eigenvector, closeness and betweenness centralities).

[*Insert Table* 5 *here*]

Since we have 5 out of 6 centrality measures that are statistically significant in Table 3, we focus only on them and perform 10 different regressions for the GGKT and the GMT where we include two centrality measures in each regression. In Table 6, we only report the results for the measures that show significant results most frequently. Table 6 reveals that it is the *Katz-Bonacich* or the *key-player centrality* of the group that has the highest predictive power, both on the GGKT (short run test) and on the GMT (long run test).

[*Insert Table* 6 *here*]

three dummy variables (when the total number of links in the group is between 1 and 3, between 4 and 6 and greater than 6). The results stay the same, i.e. the average centrality measures have still a significant impact on the GGKT and the GMT and the dummy variables have no significant impact. Also, we explored the importance of interaction terms between friendship links and centrality measures. No relevant cross effects are detected. These results are available upon request.

5.2 Leadership versus weakest link

In this section, we investigate the role of leaders (stars) and weakest links (bad apples) in shaping outcomes. Even if two groups have the same average centrality, it is possible that, in one group, there is a person with a very high centrality and another person with a very low centrality while, in the other group, all members have the same centrality.

We estimate the following equation:

$$\overline{y}_{rs} = \beta_0 + \beta_1 \max_{j \in r} C_{jrs} + \beta_2 \overline{A}_{rs} + \beta_3 A_{\max_{j\in r} C_{jrs}} + \beta_4 \overline{X}_{rs} + \eta_s + \epsilon_{rs} \tag{3}$$

where the variables $\overline{y}_{rs}$, $\overline{A}_{rs}$, $\overline{X}_{rs}$, η_s and ϵ_{irs} are the same as in (2). To test the leadership effect, we introduce $\max_{j\in r} C_{jrs}$, which is the student with the highest centrality within the group, including individual i. This variable replaces $\overline{C}_{irs}$ in (2). We also control for the ability of the student with the highest centrality in the group by adding $A_{\max_{j\in r} C_{jrs}}$ in the regression.

To study the *weakest-link* effect on *group outcomes*, we estimate equation (3) where we replace $\max_{j\in r} C_{jrs}$ by $\min_{j\in r} C_{jrs}$ and $A_{\max_{j\in r} C_{jrs}}$ by $A_{\min_{j\in r} C_{jrs}}$.

Table 7 displays the results. Panel A contains the evidence on the effects of the *leaders*. As above, even if all centralities are reported in the same column, we test the impact of each average centrality on outcomes *separately*.

Interestingly, the results are similar to that of Table 3 when we looked at the impact of the average group centrality on group outcomes. Indeed, we find that the impact of a leader, including myself, is positive and significant for five centrality measures (Katz-Bonacich, key-player, closeness, eigenvector and degree centrality). This means that having a leader in a group, i.e. a student with a high centrality, increases the group performance both in the short run (when the group is just formed for the GGKT) and in the longer run (a week after the group is formed for the GMT). In terms of magnitude, the effects of maximum centrality and of average centrality are comparable. For example, one can see that one standard deviation increase in the average Katz-Bonacich centrality (from Table 3) translates into about 12% of a standard deviations of the GGKT. This is about the same effect that is obtained for the maximum Katz-Bonacich centrality (from Table 8).

Panel B contains the evidence on the effects of the least central individuals, i.e. the weakest links (or bad apples) in a group. Interestingly, contrary to the leadership effects where all centrality measures had a significant impact on both the GGKT and the GMT (Table 8, Panel A), we see that the weakest link has nearly no impact on the GGKT and on the GMT. In other words, if the weakest link is measured as the person (including myself) with the lowest centrality in my group, then this person has nearly no impact on the group's short-run (GGKT) and long-run outcome (GMT).

$[Insert\ Table\ 7\ here]$

In Table 8, we compare the explanatory power of alternative centrality indicators as a measure of leadership. Not surprisingly, in Table 8, when we consider one centrality against the other, we see that again the *Katz-Bonacich centrality* and the *key-player centrality* have the highest explanatory power. In other words, leaders, as measured by a high Katz-Bonacich and key-player centrality, are the students that positively affect the outcomes of the group both in the short and long(er) run.

$[Insert\ Table\ 8\ here]$

5.3 Individual outcomes and network centrality

Our analysis so far has shown that the team members' skills captured by the individual position in own social networks are important in enhancing performance in collective tasks. The question we address in this section is whether the team members' network centralities are also important in shaping individual performance. The influence of group members' outcomes and characteristics on individual outcomes is a question still open in the vast literature on peer effects because of the empirical difficulties associated with an endogenous allocation of individuals into peer groups. The random allocation of students in groups in our experiments allows us to answer this question.

As explained in Section 3.2, in our experiment, at the end of the week (referred to as period $t+1$ in Figure 1), each student is also asked to take an *individual post-experiment math test* (IPOMT). This test is taken individually, although the questions and structure of this test are similar to the ones of the group math test (GMT).

We investigate whether the individual performance is affected by the teammates skills after a week of interactions using the model:

$$\Delta y_{irs} = \beta_0 + \beta_1 \overline{C}_{-irs} + \beta_2 C_{irs} + \beta_3 X_{irs} + \eta_s + \epsilon_{irs} \tag{4}$$

where Δy_{irs} is the difference in the *individual test score* of individual i belonging to group r in school s between the pre-experimental math test IPEMT and the post-experimental math test IPOMT, C_{irs} is the centrality of individual i, and $\overline{C}_{-irs}$ denotes the average centrality of group r to which i belongs to, which excludes individual i. Similarly to model (2), X_{rs}

denotes observable *characteristics* of individual i in school s (which includes gender, parent's education, access to electricity, etc.; see Table 1), η_s is the school fixed effects and ϵ_{rs} is an error term. By using differences in the individual performance, we control for the influence of unobserved individual factors that are constant over time.

The empirical results are displayed in Table 9. We report in column (1) the results on the entire sample and in column (2) the results on the groups with no friendship links. It appears that, after a week of interactions, there is a significant influence of the network centrality of the peers each individual has been randomly exposed to on the individual gains in educational performance.

[*Insert Table* 9 *here*]

Finally, we investigate in Table 10 whether and to what extent own centrality and group centralities generate cross effects. Column (1) reveals that cross effects between average centrality and own centrality are not significant for most centrality measures. The only exception is eigenvector centrality, for which the effect is statistically significant and positive. This indicates that the impact of own centrality on own performance is higher the higher the centralities of the other members of the group. In column (2), we investigate whether the relationship between individual centrality and individual performance is affected by the presence of a leader in the groups. The results show that the cross effect, when significant, is negative. This suggests that the presence of a leader weakens the impact of own centrality on own performance. On the other hand, column (3) shows that the weakest link in the group does not interfere with the positive impact of own centrality on IPOMT.

[*Insert Table* 10 *here*]

6 Inspecting the mechanisms

We conjecture that network centrality measures, especially the *Katz-Bonacich* and the *key player* centrality, capture some *non-cognitive* or *soft skills* of the students.

In order to investigate our conjecture, we further enrich our experiment. We *randomly* select 16 schools out of the 80 schools. Mean baseline math test score of these 16 schools is 6.896, which is very similar to the mean test score of 6.898 for the same test among the students in 80 schools. We ask all students of grade four in these schools (N=512) to play standard games about risk-taking behavior, patience and competitiveness.[14]

The first game is a *risk taking* game where we measure the degree of riskiness of each student. We prepare a jar with five pencils in it, out of which one pencil has a red mark on the bottom. Students cannot see the mark on the pencil until they take it out of the jar. The rule of the game is that a student can take out from the jar as many pencils as s/he wants, as long as the pencil with red mark is not included in the selection. S/he can keep all the pencils if there is no red mark on them, otherwise s/he needs to return all of them. We ask students to decide how many pencils to take out of the jar. Thus, the more pencils a student decides to take out of the jar (out a total of five), the more s/he enjoys risk. The second game is a *time preference* game where we measure how patient the students are. This game consists of asking students to decide between having a plate of 4 candies tomorrow or a plate of 6 candies after two days. We use whether the student choose to wait as a measure of patience.[15] Finally, we propose a *competition game* to assess how competitive the students are. In this game, students are asked to sum series of three randomly chosen two-digit numbers in a five minutes time window. They can choose between two different payment schemes. If a student chooses Option 1, s/he gets 1 candy for each problem that s/he solves correctly in the 5 minutes. His/her payment does not decrease if s/he provides an incorrect answer to a problem. If s/he chooses Option 2, s/he is randomly paired with another person and his/her payment depends on his/her performance relative to that of the person that s/he is paired with. If s/he solves more problems correctly than the person s/he is paired with, s/he receives 2 candies per correct answer. If both of them solve the same number of problems, they will receive 1 candy per correct answer. If s/he solves less than the person s/he is paired with, s/he will not receive any candy. The students are not allowed to use a calculator to do the sums; however they are allowed to make use of the provided scratch paper. Our indicator of competitiveness is whether the student chooses to compete, that is, takes option 2.

These experimental measures of risk, time preference and competitiveness are closely related to soft skills or non-cognitive skills in the literature on economics of education (Koch, et al., 2015). Previous literature found that risk attitudes play a significant role in educa-

[14]See Andersen et al. (2013), Bettinger and Slonim (2007), Gneezy et al. (2003, 2009) and Samak (2013) for studies using those games to elicit attitudes among children of various ages. Also, Cameron et al. (2013) use similar games for similar purposes for a population of adults.

[15]Specifically, we run 4 rounds and we extract randomly the round that counts. Students take home the choice they made in that chosen round. They are told that they should make decisions in each round as if it is the round that counts.

tion and labor market outcomes (Castillo et al., 2010; Liu, 2013). For instance, Belzil and Leonardi (2007, 2013) find that students with high risk aversion invest less in higher education. Competition or confidence is also claimed to have a strong impact on labor market outcomes. Gneezy et al. (2003) and Niederle and Vesterlund (2007) study a gender gap in competitiveness as a potential explanation for a gender gap in wage.

In order to investigate whether centrality measures capture these non-cognitive skills, we estimate a series of regressions where each of the six centralities is a function of the game outcomes and ability of the students. The estimation results are displayed in Table 11. First, in column (1) (and also column (4)), we regress each centrality on the *ability* of each student, captured by the test score at the individual math test taken before the experiment, i.e. the IPEMT. It should be clear that ability, as measured by the test score at the IPEMT, is a measure of *cognitive skill*. Second, in columns (2) and (3) (and also columns (5) and (6)), controlling for ability (or IPEMT), we look at the impact of the different non-cognitive skills (risk taking, time preference and competitiveness) on centrality measures without and with other controls.

[*Insert Table* 11 *here*]

First, looking at all columns (1) and (4), we see that, for all the centrality measures, there is *no* significant correlation between centrality measures and ability (or IPEMT). This seems to indicate that centrality measures do *not* capture cognitive skills.

Second, when we look at the other columns, we see that there are strong correlations between time preference or patience and competitive behaviors of our students and their intercentrality measure (or key-player centrality) and their Katz-Bonacich centralities. This is suggestive evidence that, indeed, these two centrality measures capture some of the personal traits (or soft skills) of the students, namely their degree of patience and competitiveness. If we look at the other results, we also see that competitiveness is strongly related with the four other centrality measures. In other words, the more students are competitive, the higher is their centrality. Finally, it seems that the eigenvector centrality and the closeness centrality are correlated with the risk-taking behavior of students. Overall, these results suggest that centrality measures could provide important information about students' non-cognitive ability.

7 Conclusion

This paper documents that the position in a network of social contacts signals skills that enhance educational performance in collective and individual tasks. We reveal that some of these skills are personal traits, namely patience and competitiveness. Our results thus support the literature on non-cognitive skills, which shows that soft skills predict success in life and that programs that enhance soft skills have an important place in an effective portfolio of public policies (Heckman and Kautz, 2012). Also, our evidence on the influence of leaders in working groups suggests that individuals who are perceived as leaders because of their social network position can be used to specifically target and diffuse opinions as well as accelerate the diffusion of innovations. We leave this promising area of research for further studies.

References

[1] Almlund, M., Duckworth, A.L., J. Heckman and T. Kautz (2011), "Personality psychology and economics," In: *Handbook of the Economics of Education, Vol. 4*, Amsterdam: Elservier, pp. 1-181.

[2] Andersen, S., Ertac, S., Gneezy, U., List, J.A. and S. Maximiano (2013), "Gender, competitiveness and socialization at a young age: Evidence from a matrilineal and a patriarchal society," *Review of Economics and Statistics* 95, 1438-1443.

[3] Akerlof, G.A. (1997), "Social distance and social decisions," *Econometrica* 65, 1005-1027.

[4] Akerlof, G.A. (1997), "Social distance and social decisions," *Econometrica* 65, 1005-1027.

[5] Angrist, J. and V. Lavy (1999), "Using Maimonides' rules to estimate the effect of class size on scholastic achievement," *Quarterly Journal of Economics* 114, 533-575.

[6] Babcock, P.S. and J.L. Hartman (2010), "Networks and workouts: Treatment size and status specific peer effects in a randomized field experiment," NBER Working Paper No. 16581.

[7] Babcock, P.S., Bedard, K., Charness, G., Hartman, J.L. and H. Royer (2015), "Letting down the team? Social effects of team incentives," *Journal of the European Economic Association*, forthcoming.

[8] Ballester, C., Calvó-Armengol, A. and Y. Zenou (2006), "Who's who in networks. Wanted: the key player," *Econometrica* 74, 1403-1417.

[9] Ballester, C., Calvó-Armengol, A. and Y. Zenou (2010), "Delinquent networks," *Journal of the European Economic Association* 8, 34-61.

[10] Ballester, C. and Y. Zenou (2014), "Key player policies when contextual effects matter," *Journal of Mathematical Sociology* 38, 233-248.

[11] Bandiera, O., Barankay, I. and I. Rasul (2009), "Social connections and incentives in the workplace: Evidence from personnel data," *Econometrica* 77, 1047-1094.

[12] Bandiera, O., Barankay, I. and I. Rasul (2010), "Social incentives in the workplace," *Review of Economics Studies* 77, 417-458.

[13] Banerjee, A., Chandrasekhar, A.G., Duflo, E. and M.O. Jackson (2013), "Diffusion of microfinance," *Science* 341 (6144).

[14] Bangladesh Bureau of Educational Information and Statistics [BANBEIS] (2012), "Basic education statistics 2011," http://www.banbeis.gov.bd/webnew.

[15] Bavelas, A. (1948), "A mathematical model for group structures," *Human Organization* 7, 16-30.

[16] Begum, L., A. Islam and R. Smyth (2013), "Girls' education, stipend programs and their effects on the education of younger siblings," Monash Economics Working Papers 17-13, Monash University, Department of Economics.

[17] Belzil, C. and M. Leonardi (2007), "Can risk aversion explain schooling attainments? Evidence from Italy," *Labour Economics* 14 (6), 957-970.

[18] Belzil, C. and M. Leonardi (2013), "Risk aversion and schooling decisions," *Annals of Economics and Statistics*, 35-70

[19] Bettinger, E. and R. Slonim (2007), "Patience among children," *Journal of Public Economics* 91, 343-363.

[20] Bonacich P. (1987), "Power and centrality: A family of measures," *American Journal of Sociology* 92, 1170-1182.

[21] Borghans, L.,Duckworth,A.L.,Heckman,J.J., and B, ter Weel (2008), "The economics and psychology of personality traits," *Journal of Human Resources* 43, 972-1059.

[22] Brass, D. J. (1984), "Being in the right place: A structural analysis of individual influence in an organization," *Administrative Science Quarterly* 29, 518-539.

[23] Bramoullé, Y., Djebbari, H. and B. Fortin (2009), "Identification of peer effects through social networks," *Journal of Econometrics* 150, 41-55.

[24] Burt, R.S. (1992), *Structural Holes: The Social Structure of Competition*. Harvard: Harvard University Press.

[25] Cai, J., de Janvry, A. and E. Sadoulet (2015), "Social networks and the decision to insure," *American Economic Journal: Applied Economics* 7, 81-108.

[26] Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009), "Peer effects and social networks in education," *Review of Economic Studies* 76, 1239-1267.

[27] Calvó-Armengol, A. and Y. Zenou (2004), "Social networks and crime decisions. The role of social structure in facilitating delinquent behavior," *International Economic Review* 45, 939-958.

[28] Cameron, L.A., Erkal, N., Gangadharan, L. and X. Meng (2013), "Little emperors: Behavioral impacts of China's one-child policy," *Science* 339 (6122), 953-957.

[29] Carrell, S.E., Fullerton, R.L. and J.E. West (2009), "Does your cohort matter? Estimating peer effects in college achievement," *Journal of Labor Economics* 27, 439-464.

[30] Carrell, S.E., Sacerdote B.I. and J.E. West (2013), "From natural variation to optimal policy? The importance of endogenous peer group formation," *Econometrica* 81, 855-882.

[31] Castillo, M., Petrie, R. and M. Torero (2010), "On the preferences of principals and agents," *Economic Inquiry* 48, 266-273.

[32] Centola, D. (2010), "The spread of behavior in an online social network experiment," *Science* 329, 1194-1197.

[33] Centola, D. (2011), "An experimental study of homophily in the adoption of health behavior," *Science* 334, 1269-1272.

[34] Choi, S., Gale, D., and S. Kariv (2012), "Social learning in networks: A quantal response equilibrium analysis of experimental data," *Review of Economic Design* 16,93-118.

[35] Cunha, F., Heckman,J.J., and S.M. Schennach (2010), "Estimating the technology of cognitive and noncognitive skill formation," *Econometrica* 78, 883-931.

[36] Dequiedt, V. and Y. Zenou (2014), "Local and consistent centrality measures in networks", CEPR Discussion Paper No. 10031.

[37] Gneezy, U, Niederle, M. and A. Rustichini (2003), "Performance in competitive environments: Gender differences," *Quarterly Journal of Economics* 143, 1049-1074.

[38] Gneezy, U, Leonard, K.L. and J.A. List (2009), "Gender differences in competition: Evidence from a matrilineal and a patriarchal society," *Econometrica* 77, 1637-1664.

[39] Goeree, J.K., M.A. McConnell, T. Mitchell, T. Tromp, and L. Yariv (2010), "The 1/d Law of Giving," *American Economic Journal: Micro-Economics* 2, 183-203.

[40] Granovetter, M.S. (1974), *Getting a Job: A Study in Contacts and Careers, Cambridge*, MA: Harvard University Press.

[41] Grewal, R., Lilien, G. L. and G. Mallapragada (2006), "Location, location, location: How network embeddedness affects project success in open source systems," *Management Science* 52, 1043-1056.

[42] Guryan, J., Kroft, K. and M.J. Notowidigdo (2009), "Peer effects in the workplace: Evidence from random groupings in professional golf tournaments," *American Economic Journal: Applied Economics* 1, 34-68.

[43] Hamilton, B.H., Nickerson, J.A. and H. Owan (2003), "Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation," *Journal of Political Economy* 111, 465-497.

[44] Haynie, D. (2001), "Delinquent peers revisited: Does network structure matter?" *American Journal of Sociology* 106, 1013-1057.

[45] Heckman, J.J. and T. Kautz (2012), "Hard evidence on soft skills," *Labour Economics* 19, 451-464.

[46] Heckman, J.J.,Malofeeva, L., Pinto, R. and P.A. Savelyev (2011), "Understanding the mechanisms through which an influential early childhood program boosted adult outcomes," Unpublished manuscript, Department of Economics, University of Chicago.

[47] Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A. and A. Yavitz (2010), "The rate of return to the High Scope Perry Preschool Program," *Journal of Public Economics* 94, 114-128.

[48] Heckman, J.J., Stixrud, J. and S. Urzua (2006), "The effects of cognitive and noncognitive abilities on labor-market outcomes and social behavior," *Journal of Labor Economics* 24,411-482.

[49] Holbrook, J. (2005), "Report on organizing the ROSE survey in Bangladesh," http://www.uv.uio.no/ils/english/research/projects/rose/partners/bangladesh/report-bgd.pdf.

[50] Horrace, W., Liu, X. and E. Patacchini (2015), "Endogenous network production function with selectivity," *Journal of Econometrics*, forthcoming.

[51] Ioannides, Y.M. (2012), *From Neighborhoods to Nations: The Economics of Social Interactions*, Princeton: Princeton University Press.

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

[53] Jackson, M.O. (2011), "An overview of social networks and economic applications," In: J. Benhabib, A. Bisin and M.O. Jackson (Eds.), *Handbook of Social Economics Volume 1A*, Amsterdam: Elsevier Science, pp. 511-579.

[54] Jackson, M.O. (2014), "Networks in the understanding of economic behaviors," *Journal of Economic Perspectives* 28, 3-22.

[55] Jackson, M.O., Rogers, B.W. and Y. Zenou (2015), "The economic consequences of social network structure", CEPR Discussion Paper 10406.

[56] Jackson, M.O. and Y. Zenou (2015), "Games on networks", In: P. Young and S. Zamir (Eds.), *Handbook of Game Theory, Vol. 4*, Amsterdam: Elsevier Publisher, pp. 91-157.

[57] Katz, L. (1953), "A new status index derived from sociometric analysis," *Psychometrika* 18, 39-43.

[58] Kearns, M.J., Judd, S., Tan, J. and J. Wortman (2009), "Behavioral experiments on biased voting in networks," *Proceedings of the National Academy of Sciences of the USA* 106, 1347-1352.

[59] Koch, A., Nafziger, J. and H. Skyt Nielsen (2015), "Behavioral economics of education," *Journal of Economic Behavior and Organization*, forthcoming.

[60] König, M.D., Liu, X. and Y. Zenou (2014), "R&D networks: Theory, empirics and policy implications," CEPR Discussion Paper No. 9872.

[61] Leavitt, H.J. (1951), "Some effects of certain communication patterns on group performance," *Journal of Abnormal and Social Psychology* 46, 38-50.

[62] Lindquist, M.J. and Y. Zenou (2014), "Key players in co-offending networks," CEPR Discussion Paper No. 9889.

[63] Liu, E.M. (2013), "Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China," *Review of Economics and Statistics* 95, 1386-1403.

[64] Liu, X., Patacchini, E. and Y. Zenou (2014), "Endogenous peer effects: Local aggregate or local average?" *Journal of Economic Behavior and Organization* 103, 39-59.

[65] Liu, X., Patacchini, E., Zenou, Y. and L-F. Lee (2012), "Criminal networks: Who is the key player?" CEPR Discussion Paper No. 8772.

[66] Liu, X., Patacchini, E., Zenou, Y. and L-F. Lee (2015), "Who is the key player? A network analysis of juvenile delinquency," Unpublished manuscript, Stockholm University.

[67] Mahmud, W, Asadullah, N. and A. Savoia (2013), "Bangladesh's achievements in social development indicators. Explaining the puzzle," *Economic and Political Weekly* 48, no 44.

[68] Mas, A. and E. Moretti (2009), "Peers at work," *American Economic Review* 99, 112-145.

[69] Mehra, A., Kilduff, M. And D. J. Brass (2001) , "The social networks of high and low self-monitors: Implications for workplace performance," *Administrative Science Quarterly* 46, 121-146.

[70] Niederle, M., and L. Vesterlund (2007), "Do women shy away from competition? Do men compete too much?," *Quarterly Journal of Economics* 122, 1067-1101.

[71] Mastrobuoni, G. and E. Patacchini (2012), "Organized crime networks: An application of network analysis techniques to the American mafia," *Review of Network Economics* 11, Article No. 10.

[72] Putnam, R. (2000), *Bowling Alone: The Collapse and Revival of American Community*, New York: Simon and Schuster.

[73] Sacerdote, B. (2001), "Peer effects with random assignment: Results from Dartmouth roomates," *Quarterly Journal of Economics* 116, 681-704.

[74] Sacerdote, B. (2011), "Peer effects in education: How might they work, how big are they and how much do we know thus far?", In: E.A. Hanushek, S. Machin and L. Woessmann (Eds.), *Handbook of Economics of Education, Vol. 3*, Amsterdam: Elevier Science, pp. 249-277.

[75] Sacerdote, B. (2014), "Experimental and quasi-experimental analysis of peer effects: Two steps forward," *Annual Review of Economics* 6, 253-272.

[76] Samak, A.C. (2013), "Is there a gender gap in preschoolers' competitiveness? An experiment in the U.S.," *Journal of Economic Behavior and Organization* 92, 22-31.

[77] Tao, J. and L.-F. Lee (2014), "A social interaction model with an extreme order statistic," *Econometrics Journal* 17, 197–240.

[78] Tapan, M.S.M. (2010), "Science education in Bangladesh," In: Y.-J. Lee (Ed.), *World of Science Education: Science Education Research in Asia*, Rotterdam: Sense, pp. 17–34.

[79] Tatsi, E. (2014), "Endogenous social interactions: Which peers matter?" Unpublished manuscript, Goethe University, Frankfurt.

[80] Topa, G. and Y. Zenou (2015), "Neighborhood and network effects", In: G. Duranton, V. Henderson and W. Strange (Eds.), *Handbook of Regional and Urban Economics, Vol. 5*, Amsterdam: Elsevier Publisher, pp. 561-624.

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

[82] Zenou, Y. (2015), “Key players,” In: Y. Bramoullé, B.W. Rogers and A. Galeotti (Eds.), *Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press, forthcoming.

[83] Zimmerman, D. (2003), “Peer effects in academic outcomes: Evidence from a natural experiment,” *Review of Economics and Statistics* 9-23.

Appendix 1:
Individual Pre-Experiment Math Test (IPEMT)

1. In a case, the dividend is 7363, quotient is 49 and remainder is 13. What is the divisor?
   a) 130
   b) 140
   c) 150
   d) 160

2. Write the smallest number using the digits 2, 3, 6, 1?
   a) 2326
   b) 1236
   c) 6321
   d) 1362

3. The price of a book is 17 Taka. What would be the total price of three of these books?
   a) 50 Taka
   b) 51 Taka
   c) 61 Taka
   d) 71 Taka

4. Which number is divisible by 1, 3, 6, 9?
   a) 19
   b) 20
   c) 17
   d) 18

5. Calculate the L.C.M. of 25 and 30.
   a) 300
   b) 200
   c) 150
   d) 250

6. 28 + 7 = 3 + 8 – 20. What is this called?
   a) Number
   b) Symbol
   c) Number series
   d) Mathematical statement

7. Which number needs to be added with 37 to get a sum of 50?
   a) 13
   b) 14
   c) 5
   d) 12

8. How many types of triangles are there based on the sides?
a) 2
b) 3
c) 4
d) 5

9. What does the symbol $\leq$ mean?
a) Smaller
b) Greater
c) Equal
d) Smaller and equal

10. What is the previous number to the smallest number with three digits?
a) 101
b) 112
c) 99
d) 100

11. What is the sum of the place values of 4, 7, 2 in the number of 947231?
a) 47231
b) 47200
c) 40072
d) 4720

12. What are the symbols of greater and smaller?
a) $>, =$
b) $<, =$
c) $>, <$
d) None of the above

13. Sum of three numbers is 9890. Two of these numbers are 620 and 1260. What is the third number?
a) 8100
b) 590
c) 8010
d) 8770

14. How many hours are equal to 5 weeks 6 days 9 hours?
a) 993 hours
b) 990 hours
c) 940 hours
d) 949 hours

15. 1 *Mon* = how many *Ser*?

a) 56 *Ser*
b) 40 *Ser*
c) 39 *Ser*
d) 45 *Ser*

Appendix 2: Group General Knowledge Test (GGKT)

Direction: Please answer ALL of the following questions. You will get 1 (one) mark for each correct answer. Total time is 20 minutes.

1. Which of the following is the independence day of Bangladesh?

a) 21 February	b) 26 March
c) 17 April	d) 16 December

2. In terms of population, what is the position of Bangladesh in the world?

a) 5th	b) 7th
c) 8th	d) 10th

3. Which is the longest sea beach in the world?

a) Cox's Bazar	b) Kuakata
c) Deegha	d) Pataya

4. Which is the greatest delta in the world?

a) India	b) China
c) Bangladesh	d) Australia

5. What is the area of Bangladesh?

a) 54501 sq miles	b) 56501 sq miles
c) 57401 sq miles	d) 58501 sq miles

6. Which is the oldest place in Bengal?

a) Horikel	b) Samatal
c) Pundra	d) Rarh

7. Which of the following district was called 'Jahanabad'?

a) Satkhira	b) Khulna
c) Dhaka	d) Barisal

8. Which of the following is regarded as the national children day of Bangladesh?

a) 17 January	b) 17 February
c) 17 March	d) 17 April

9. Who is the only Nobel Prize winner of Bangladesh?

a) Joynul Abedin	b) Kamrul Hassan
c) Dr. Muhammad Younus	d) Kazi Nazrul Islam

10. For which book did Rabindranath Tagore win the Nobel Prize?

a) Sonar Tori	b) Geetanjali
c) Sanchaeeta	d) Balaka

11. Who is the first Everest Winner of Bangladesh?

a) Musa Ibrahim	b) Sajal Khaled
c) Sakib Al Hassan	d) Mohammad Ashraful

12. Which of the following is not a part of folk music of Bangladesh?

a) Baul music	b) Keertan music
c) Jari music	d) Band music

13. What is the national sport event of Bangladesh?

a) Football	b) Cricket
c) Hockey	d) Kabadi

14. Which country is the maximum winner of World Cup Cricket?

a) India	b) Pakistan
c) Australia	d) England

15. Which country was the winner of 2010 World Cup Football?

a) Brazil	b) Argentina
c) Italy	d) Spain

16. Which is the first artificial Earth satellite?

a) Asterix	b) Sputnik 1
c) Sputnik 2	d) Apollo 11

17. How many continents are there in the world?

a) 5	b) 6
c) 7	d) 9

18. In terms of population, which is the largest continent in the world?

a) America	b) Asia
c) Europe	d) Africa

19. Which is the longest river in the world?

a) Padma	b) Jamuna
c) Hoangho	d) Yangsikian

20. Which part of Asia is Bangladesh situated?

a) North-East	b) South-East
c) North-West	d) South-West

Appendix 3: Group Math Test (GMT)

Problem 1: Arrange the numbers in the following Table in Ascending and Descending order using symbol. One is done for you.

Number	Ascending	Descending
65032, 8973, 26940, 53278, 80149, 84256, 9856	8973 < 9856 < 26940 < 53278 < 65032 < 80149 < 84256	84256 > 80149 > 65032 > 53278 > 26940 > 9856 > 8973
88457, 45682, 23412, 780021, 100000, 45789, 65231		
78921, 12356, 98213, 238593, 45123, 636336, 24789		
9874, 87412, 23145, 89564, 98741, 45621, 32100		
654646, 3265, 7841565, 568984, 56874, 89586, 656898		

Problem 2: Without repeating any digit, arrange the following groups of numbers to make the greatest and smallest numbers possible. Calculate the difference between the greatest and smallest number in each set.

(a) 7, 2, 3, 0, 1
(b) 4, 2, **3**, 8, 1
(c) 6, 0, 7, 8, 5
(d) 2, 3, 7, 0, 9

Problem 3: Here is part of a wall chart that lists numbers from 1 to 100.

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25					

Below is part of the same wall chart.

43	
53	
	?

Look at the charts carefully and find out what number should be in the box with the question mark inside. How do you find this?

Problem 4: In which pair of numbers is the second number 100 more than the first number? Please show how you solve this problem.

A. 199 and 209
B. 4236 and 4246
C. 9635 and 9735
D. 51863 and 52863

Problem 5: Ajay wanted to use his calculator to add 1463 and 319. He entered 1263 + 319 by mistake. What could he do to correct his mistake?

A. Add 20
B. Add 200
C. Subtract 200
D. Subtract 20

Please show how you solve this problem.

Problem 6: Rahim had 100 mangoes. He sold some and then had 50 left. □ represents the number of mangoes that he sold. Which of these is a number sentence that shows this?

A. □ – 50 = 100
B. 50 – □ = 100
C. □ – 100 = 50
D. 100 – □ = 50

Problem 7: Rahim had 100 mangoes. He sold some and then had 50 left. He found some rotten mangoes and threw them away. Finally he had 45 mangoes left. □ represents the number of mangoes that he sold and # represents the number that was rotten. Which of these is a number sentence that shows this?

A. □ + 50 – # = 100
B. □ + 50 + # = 100
C. □ + 45 + # = 100
D. 100 – □ = 45

Problem 8: The sum of ages of a mother and a daughter is 65 years. The mother's age is 4 times as much as the daughter's. What are the ages of the mother and the daughter? What will be their ages after 6 years?

Problem 9: Tina has Tk. 125 more than Bina and Tk. 45 less than Rina. Tina has Tk. 300. How much does each of Bina and Rina have? How much do the three persons have altogether?

Problem 10: In 2012, there were 95 members in a cooperative society. In 2013 25 new members joined in the society. Each of the members has paid 200 for a picnic in 2013. How much money was collected as subscription?

Appendix 4: Individual Post-Experiment Math Test (IPOMT)

Problem 1: Arrange the following numbers in Ascending and Descending order using symbol.

5238, 4132, 8725, 6138, 7201

Problem 2: Without repeating, arrange the following digits to make the smallest number possible.

4, 3, 9, 1

Problem 3: Subtract the greatest number with 3 digits from the smallest number with 5 digits.

Problem 4: The difference between two numbers is 425. If the greater number is 7235, find out the smaller number.

Problem 5: When you subtract one of the following numbers from 900, the answer is greater than 300. Which number is it?

A. 823

B. 712

C. 667

D. 579

Problem 6: What is 3 times 23?

A. 323

B. 233

C. 69

D. 26

Problem 7: Mr. Rahim drew eight 100 Taka notes, four 50 Taka notes and two 10 Taka notes from the bank. What is the amount he drew from the bank?

Problem 8: Fill the blank in the following number sentence.

2000 + ____________ + 30 + 9 = 2739

Problem 9: Kamal had 50 mangoes. He sold some and then had 20 left. Which of these is a number sentence that shows this?

A. □ – 20 = 50

B. 20 – □ = 50

C. □ – 50 = 20

D. 50 – □ = 20

Problem 10: If we equally distribute Taka 7642 among 52 people, how much will each of them receive? What will be the remaining amount?

Appendix 5: Network Centrality Measures

5.1. Definitions

In this appendix, we give the formal definition of the centrality measures used in the paper. We consider a finite set of individuals (or nodes) $N = \{1, \ldots, n\}$ who are connected in a network. A *network* (or *graph*) is a pair $(N, \mathbf{g})$, where $\mathbf{g}$ is a network on the set of nodes N. A network $\mathbf{g}$ is represented by an $n \times n$ adjacency matrix $\mathbf{G}$, with entry g_{ij} denoting whether i is linked to j and can also include the intensity of that relationship. In this paper, we consider *indegree weighted directed networks*, which are defined in Section 3.2.

The *distance* $\delta_{ij}(\mathbf{g})$ between two nodes i and j in the same component of a network $\mathbf{g}$ is the length of a shortest path (also known as a *geodesic*) between them.

The *diameter* is the largest distance between two nodes i and j in a network $\mathbf{g}$, i.e. $\max \delta_{ij}(\mathbf{g})$, $\forall i, j$.

The *neighbors* of a node i in a network $(N, \mathbf{g})$ are denoted by $N_i(\mathbf{g})$.

The *degree* of a node i in a network $(N, \mathbf{g})$ is the number of neighbors that i has in the network, i.e., $|N_i(\mathbf{g})|$. As a result, the degree centrality is the degree of node i divided by the number of feasible links, i.e.,

$$d_i(\mathbf{g}) = \frac{|N_i(\mathbf{g})|}{n-1} = \frac{\sum\limits_{j=1}^{n} g_{ij}}{n-1} \tag{5}$$

It has values in $[0, 1]$.

The *betweenness centrality* a measure of a node's centrality in a given network. It is equal to the number of *shortest paths* from all nodes to all others that pass through that node. It is calculated as follows:

$$B_i(\mathbf{g}) = \frac{1}{(n-1)(n-2)} \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{m^i_{jk}(\mathbf{g})}{m_{jk}(\mathbf{g})}, \tag{6}$$

where $m_{jk}(\mathbf{g})$ is the number of shortest paths between node j and node k in network $\mathbf{g}$ and $m_{jk}(\mathbf{g})$ is the number of shortest paths between node j and node k trough i in network $\mathbf{g}$. It has values in $[0, 1]$.

The *closeness centrality* is defined as follows:

$$C_i(\mathbf{g}) = \frac{n-1}{\sum\limits_{j \neq i} \delta_{ij}(\mathbf{g})} \tag{7}$$

where $\delta_{ij}(\mathbf{g})$ is the shortest path between nodes i and j in network $\mathbf{g}$. It has values in $[0,1]$.

The *eigenvector centrality* is defined using the following recursive formula:

$$v_i(\mathbf{g}) = \frac{1}{\lambda_1} \sum_{j=1}^{n} g_{ij} v_j(\mathbf{g}) \tag{8}$$

where λ_1 is the largest eigenvalue of $\mathbf{G}$. By the Perron-Frobenius theorem, using the largest eigenvalue guarantees that v_i is always positive. Eigenvector centrality $v_i(\mathbf{g})$ is the leading eigenvector of $\mathbf{G}$. This centrality measure is different from the others above because, in measuring a node's centrality, it gives a specific weight to each connected node by considering its relevance in terms of centrality. In matrix form, we have:

$$\lambda_1 \mathbf{v}(\mathbf{g}) = \mathbf{G}\mathbf{v}(\mathbf{g}) \tag{9}$$

The *Katz-Bonacich centrality* is a generalization of (8), which allows the Katz-Bonacich centrality to depend on a parameter ϕ. We have:

$$b_i(\phi, \mathbf{g}) = 1 + \phi \sum_{j=1}^{n} g_{ij} b_j(\phi, \mathbf{g}) \tag{10}$$

where $\phi < 1/\lambda_{\max}(\mathbf{G})$, where $\lambda_{\max}(\mathbf{G})$ is the spectral radius of $\mathbf{G}$. The parameter ϕ is usually interpreted as a discount factor of each node. In matrix form, we have:

$$\mathbf{b}(\phi, \mathbf{g}) = \mathbf{1}_n + \phi \mathbf{G}\mathbf{b}(\phi, \mathbf{g}) \tag{11}$$

where $\mathbf{1}_n$ is a $n-$vector of 1. The Katz-Bonacich centrality has a closed form solution, which is:

$$\mathbf{b}(\phi, \mathbf{g}) = (\mathbf{I}_n - \phi \mathbf{G})^{-1} \mathbf{1}_n \tag{12}$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix. The condition $\phi < 1/\lambda_{\max}(\mathbf{G})$ guarantees that $(\mathbf{I}_n - \phi \mathbf{G})$ is invertible.

We show below that the Katz-Bonacich centrality can be obtained as a Nash equilibrium of a network game. The *key player centrality* captures instead a normative view of centrality.

5.2. Foundation of centrality measures

Consider a simple game on networks with strategic complementarities (Jackson and Zenou, 2015). Following Calvó-Armengol and Zenou (2004) and Ballester et al. (2006), consider the following linear-quadratic utility function

$$u_i(\mathbf{y}, \mathbf{g}) = \alpha_i y_i - \frac{1}{2} y_i^2 + \phi \sum_{j=1}^{n} g_{ij} y_i y_j \tag{13}$$

where each student i decides how much effort $y_i \in \mathbb{R}_+$ to exert in terms of education (i.e. how many hours to study) given the network $\mathbf{g}$ s/he belongs to. In (13), α_i captures the *observable* characteristics of student i (gender, parent's education, etc.) and $\phi > 0$ is the intensity of interactions between students. Remember that $g_{ij} = 1$ if two students are friends and zero otherwise. In our weighted directed network, $g_{ij} \in [0, 1]$ if j has nominated i as his/her friend, and $g_{ij} = 0$, otherwise, where the weight g_{ij} is given by (**??**). Ballester et al. (2006) have shown that, if $\phi < 1/\lambda_{\max}(\mathbf{G})$, then there exists a unique interior Nash equilibrium of this game with utility (13), which is given by:

$$\mathbf{y}^* \equiv \mathbf{y}^*(\mathbf{g}) = \mathbf{b}_{\boldsymbol{\alpha}}(\phi, \mathbf{g}), \tag{14}$$

where $\mathbf{b}_{\boldsymbol{\alpha}}(\phi, \mathbf{g})$ is the weighted Katz-Bonacich centrality defined as:

$$\mathbf{b}_{\boldsymbol{\alpha}}(\phi, \mathbf{g}) = (\mathbf{I}_n - \phi \mathbf{G})^{-1} \boldsymbol{\alpha} = \sum_{k=0}^{\infty} \phi_1^k \mathbf{G}^k \boldsymbol{\alpha} \tag{15}$$

Observe that (15) is just a generalization of (12) when, $\mathbf{1}_n$, the $n-$vector of 1, is replaced by $\boldsymbol{\alpha}$, the $n-$vector of α_i. This result shows that, in any game with strategic complementarities and linear-quadratic utility function where agents choose effort, there is a unique Nash equilibrium in pure strategies such that each agent provides effort according to his/her Katz-Bonacich centrality. This is gives a micro-foundation for the Katz-Bonacich centrality.[16]

Let us now define the *key-player centrality*. For that, consider the game with strategic complements developed above for which the utility is given by (13) and denote $Y^*(\mathbf{g}) = \sum_{i=1}^{n} y_i^*$ the total equilibrium level of activity in network $\mathbf{g}$, where y_i^* is the Nash equilibrium effort given by (14). Also denote by $\mathbf{g}^{[-i]}$ the network $\mathbf{g}$ without individual i. Then, in order to determine the *key player*, the planner will solve the following problem:

$$\max\{Y^*(\mathbf{g}) - Y^*(\mathbf{g}^{[-i]}) \mid i = 1, ..., n\} \tag{16}$$

Assume that $\phi < 1/\lambda_{\max}(\mathbf{G})$. Then, the *intercentrality* or the *key-player centrality* $c_i(\phi, \mathbf{g})$ of agent i is defined as follows:

$$c_i(\phi, \mathbf{g}) = \frac{b_{\alpha_i}(\phi, \mathbf{g}) b_{1_i}(\phi, \mathbf{g})}{m_{ii}} \tag{17}$$

[16] Dequiedt and Zenou (2014) propose an *axiomatic approach* to derive the degree, eigenvector and Katz-Bonacich centralities. In other words, they show which axioms are crucial to characterize centrality measures for which the centrality of an agent is recursively related to the centralities of the agents she is connected to (this includes the degree, eigenvector and Katz-Bonacich centralities).

where m_{ii} is the (i,i) cell of matrix $\mathbf{M}(\mathbf{g},\phi_1)=(\mathbf{I}_n-\phi\mathbf{G})^{-1}$, $b_{\alpha_i}(\phi,\mathbf{g})$ and $b_{1_i}(\phi,\mathbf{g})$ is the weighted and unweighted Katz-Bonacich centrality of agent i. Ballester et al. (2006, 2010) have shown that the player i^* that solves (16) is the key player if and only if i^* is the agent with the highest intercentrality in $\mathbf{g}$, that is, $c_{i^*}(\phi,\mathbf{g}) \geq c_i(\phi,\mathbf{g})$, for all $i=1,...,n$. The intercentrality measure (17) of agent i is the sum of i's centrality measures in $\mathbf{g}$, and i's contribution to the centrality measure of every other agent $j \neq i$ also in $\mathbf{g}$. It accounts both for one's exposure to the rest of the group and for one's contribution to every other exposure. This means that the key player i^* in network $\mathbf{g}$ is given by $i^* = \arg\max_i c_i(\phi,\mathbf{g})$, where

$$c_i(\phi,\mathbf{g}) = Y^*(\mathbf{g}) - Y^*\left(\mathbf{g}^{[-i]}\right). \tag{18}$$

As a result, we can rank all students in our networks by their intercentrality or key-player centrality using the formula (17) or (18).

5.3. Parameter choice for the Katz-Bonacich and the key-player centrality

The parameter ϕ is crucial in any empirical application since the centrality ranking in a given network is sensitive to this parameter's choice. It can be obtained by estimating a spatial model (as in Calvó-Armengol et al., 2009 or Liu et al., 2014). However, such estimation is unreliable with small networks and depends on the available covariates. We rely on a simple heuristic algorithm that mimics our theoretical model (given in Section 5.2), where given linear-quadratic preferences (13), individual effort in a network is equal to his/her Katz-Bonacich centrality (see (14)). Indeed, for each network $r=1,\cdots,R$, with N_r members each, we need to find a ϕ_r such that the Euclidean distance between the Grade Point Average or GPA of a student i (which measures y_i^* in our model; see Calvó-Armengol et al., 2009) and his/her Katz-Bonacich centrality is minimized. For each network r, a grid search is performed to find ϕ_r such that:

$$\min_{\phi_r}\left\{\sum_{i=1}^{N_r}\left[y_{i,r}^* - b_{\alpha_i}(\phi_r,\mathbf{g})\right]^2\right\}, \quad r=1,\cdots,R \tag{19}$$

where $b_{\alpha_i}(\phi,\mathbf{g})$ is defined by (15). Once each parameter ϕ_r that satisfies the problem (19) is found, individual Katz-Bonacich and key-player centralities can be calculated in each network using formulas (15) and (17), respectively.

Figure 1: Timeline of the experiment

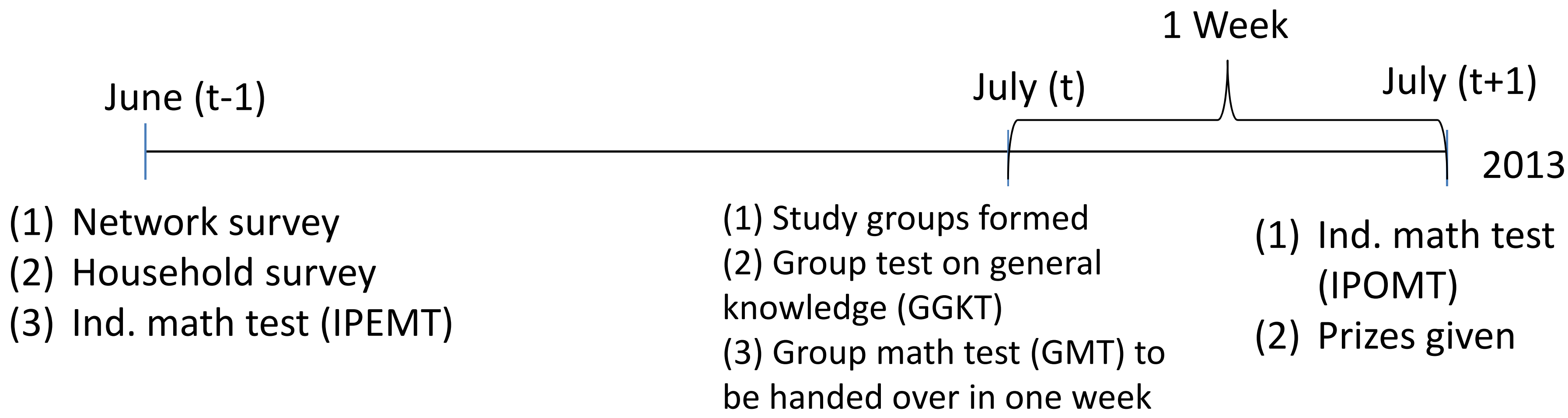


Figure 2: Distribution of group-average network centrality (N=924)

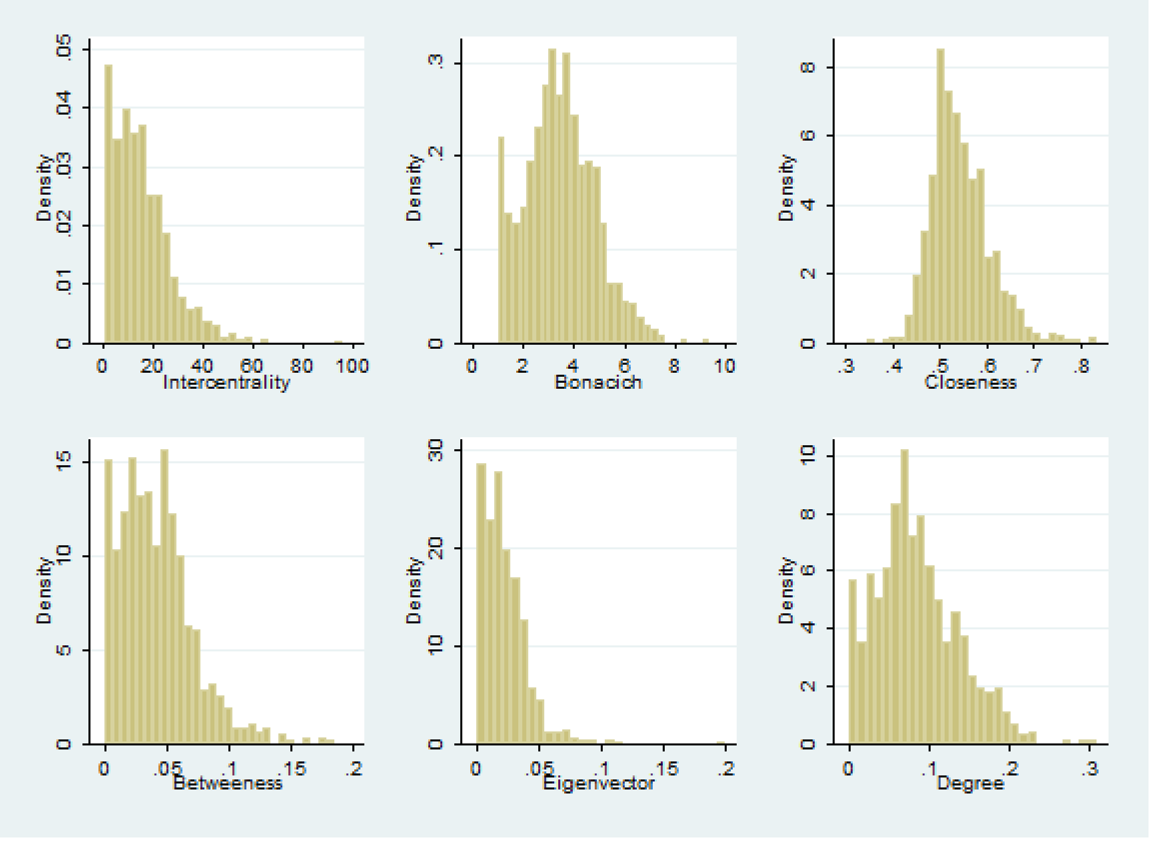


Figure 3: Distribution of individuals by number of within-group nominations

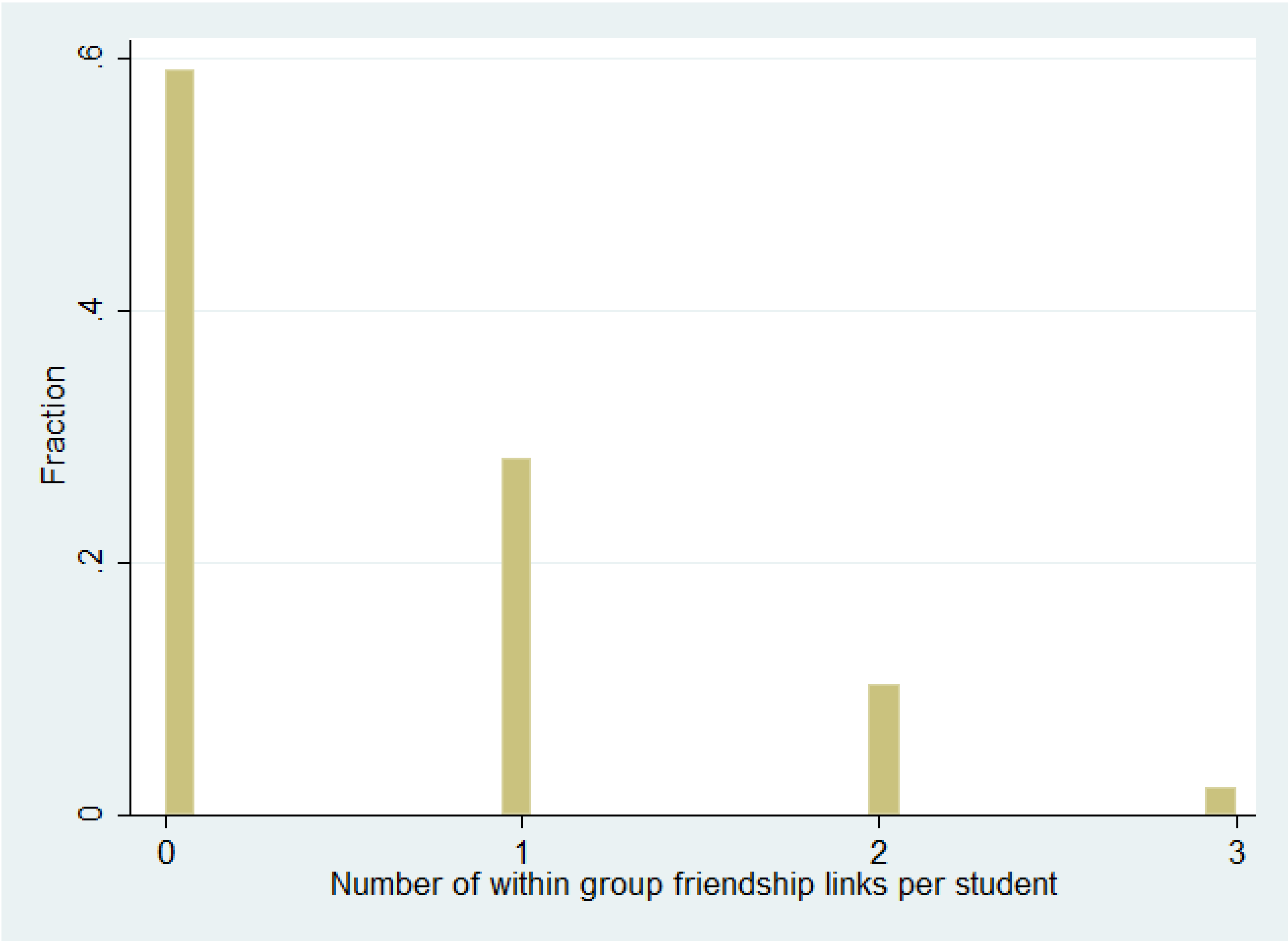


Figure 4: Distribution of groups by number of within-group nominations

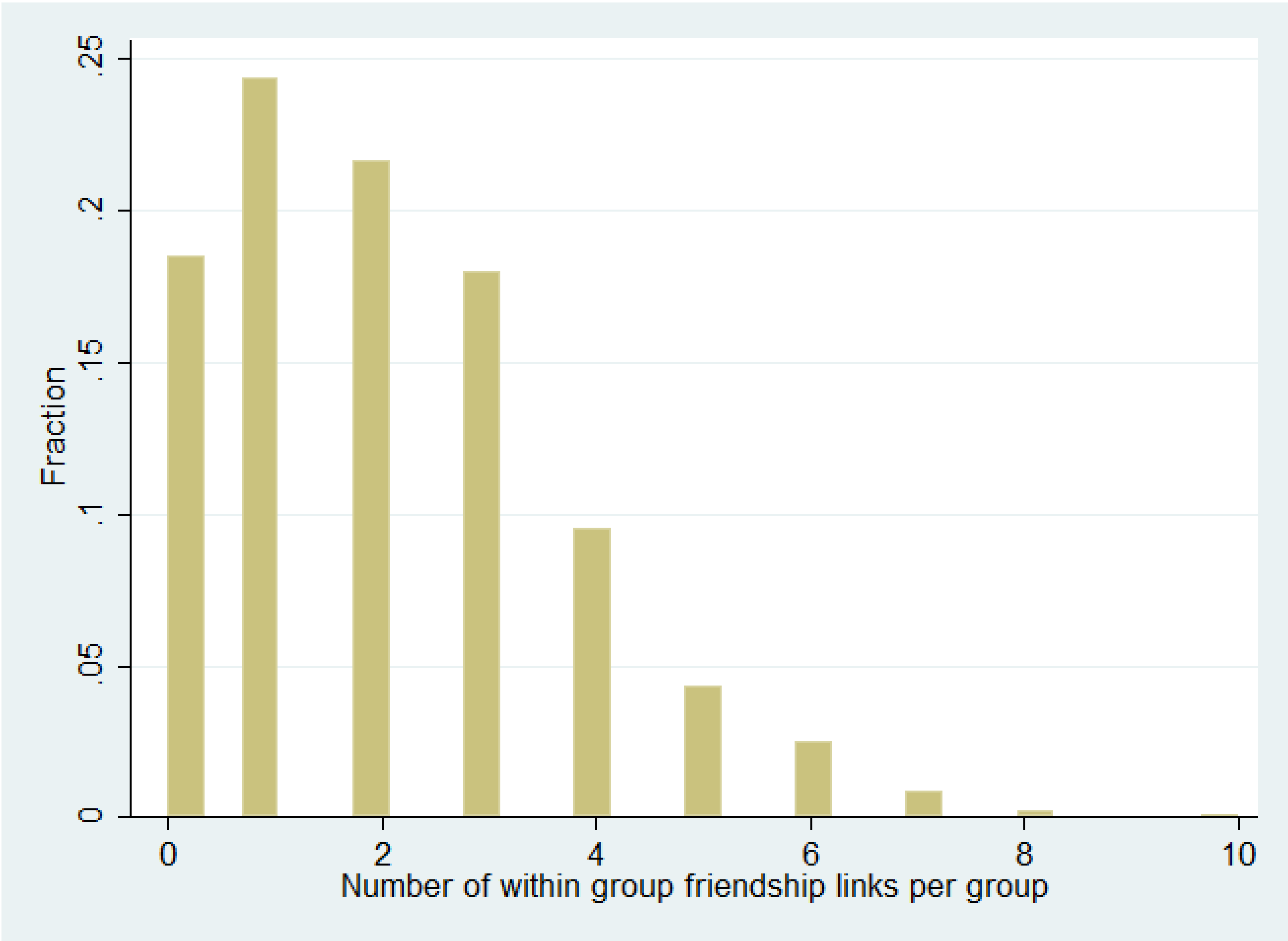


Table 1: Data Description

	Obs	Mean	Std. Dev.	Min	Max
Individual characteristics					
Female	3406	0.508	0.500	0	1
Household income per cap	3406	4297.603	836.862	921.182	10000
Household has electricity	3406	0.280	0.449	0	1
Parent education in years	3406	4.939	3.741	0	17
Parent age	3406	39.673	6.760	22.5	80
Performance indicators					
Individual pre-experiment math test (IPEMT)	3406	-0.003	1.003	-2.322	2.727
Individual post-experiment math test (IPOMT)	3406	0.003	1.002	-1.578	2.364
Group score on general knowledge (GGKT)	924	-0.002	1.006	-2.811	3.054
Group score on math assignment (GMT)	924	-0.036	1.020	-2.863	1.969
*Frequency of interactions**					
Number of meetings as a team	3405	3.408	1.717	1	6
Number of hours spent as a team for math assignment	3403	3.021	1.851	0.5	6
Number of hours spent individually studying	3400	9.426	3.794	3	14

***** "Number of meetings as a team" takes a value of 1 for "met once", 2.5 for "met 2 to 3 times", 4.5 for "met 4 to 5 times", and 6 for "met more than 5 times". "Number of hours spent as a team for math assignment" takes a value of 0.5 for "less than 1 hour", 2.5 for "2 to 3 hours", 4.5 for "4 to 5 hours" and 6 for "more than 5 hours". "Number of hours spent individually studying" is total study hours including non-math subjects. It takes a value of 3 for "less than 6 hours", 8 for "7 to 9 hours", 11 for "10 to 12 hours", and 14 for "more than 12 hours".

Table 2a: Descriptive statistics of network measures

	Obs	Mean	Std. Dev.	Min	Max
Panel A: Student level					
Intercentrality	3406	15.753	22.22	1	194.01
Katz-Bonacich	3406	3.460	2.500	1	16.12
Closeness	3406	0.546	0.078	0	2.51
Betweeness	3406	0.042	0.051	0	0.36
Eigenvector	3406	0.022	0.032	0	0.40
Degree	3406	0.085	0.086	0	0.83
Panel B: Network level					
Size	80	51.637	21.712	25	125
Density	80	0.207	0.069	0.071	0.42
Path length	80	1.740	0.170	1.314	2.18
Diameter	80	4.713	1.105	1	8

Table 2b: Descriptive statistics on average, maximum and minimum centrality measures

	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
	Panel A: All groups (N=924)				Panel B: Groups with no friendship links (N=171)			
Avg Intercentrality	15.69	11.62	1	95.98	8.47	8.06	1	46.10
Max Intercentrality	36.35	30.98	1	194.01	18.37	20.63	1	106.53
Min Intercentrality	3.18	3.8	1	47.98	2.20	2.11	1	17.26
Avg Katz-Bonacich	3.45	1.39	1	9.39	2.45	1.19	1	6.09
Max Katz-Bonacich	5.92	2.85	1	16.12	3.86	2.45	1	11.30
Min Katz-Bonacich	1.63	0.83	1	8.08	1.39	0.57	1	4.32
Avg Closeness	0.55	0.06	0.34	0.83	0.51	0.04	0.39	0.66
Max Closeness	0.1	0.38	0	2.51	0.54	0.05	0.45	0.71
Min Closeness	0.07	0.28	0	1.84	0.48	0.05	0.30	0.61
Avg Betweenness	0.04	0.03	0	0.19	0.03	0.03	0	0.12
Max Betweenness	0.09	0.06	0	0.36	0.06	0.05	0	0.25
Min Betweenness	0.01	0.01	0	0.13	0.01	0.01	0	0.07
Avg Eigenvector	0.02	0.02	0	0.2	0.01	0.01	0	0.04
Max Eigenvector	0.05	0.04	0	0.4	0.02	0.02	0	0.13
Min Eigenvector	0	0.01	0	0.07	0.00	0.00	0	0.02
Avg Degree	0.08	0.05	0	0.31	0.04	0.03	0	0.14
Max Degree	0.16	0.11	0	0.83	0.07	0.06	0	0.26
Min Degree	0.03	0.03	0	0.2	0.01	0.02	0	0.10

Table 3: Group outcomes and Network centrality

	(1) GGKT	(2) GMT	(1) GGKT	(2) GMT
	All Groups		Groups with no friendship links	
Avg. Intercentrality	0.009	0.011	0.030	-0.000
	(0.003)***	(0.002)***	(0.010)***	(0.013)
Avg. Bonacich	0.082	0.093	0.228	-0.043
	(0.024)***	(0.021)***	(0.096)**	(0.093)
Avg. Closeness	2.450	2.364	9.436	-0.863
	(0.980)**	(0.815)***	(4.619)**	(2.936)
Avg. Betweenness	-0.307	-1.334	-3.236	-7.524
	(0.937)	(1.049)	(5.517)	(4.148)
Avg. Eigenvector	4.915	6.709	25.388	4.270
	(1.764)***	(1.754)***	(11.125)**	(10.616)
Avg. Degree	1.668	2.553	6.392	-4.088
	(0.648)**	(0.643)***	(3.899)	(3.829)

Notes: Columns (1) and (2) contain the results of regressions on the whole sample while columns (3) and (4) contain the results of regressions on the subsample of groups with no friendship links. N=924 for columns (1) and (2); N=171 for columns (3) and (4). Each coefficient on average centrality is obtained from a different regression. All regressions control for individual characteristics and school fixed effects. Standard errors are clustered at the school level. * p<0.10; ** p<0.05; *** p<0.01.

Table 4: Group outcomes and Network centrality controlling for the fraction of friendship links in the group

	(1) GGKT	(2) GMT	(3) GGKT	(4) GMT
	All Groups		Groups with at least one friendship link	
Avg. Intercentrality	0.009	0.011	0.008	0.011
	(0.003)***	(0.003)***	(0.003)***	(0.003)***
Fraction of links in a group	-0.055	0.011	-0.086	-0.135
	(0.247)	(0.271)	(0.288)	(0.293)
Avg. Bonacich	0.086	0.092	0.080	0.101
	(0.029)***	(0.025)***	(0.028)***	(0.026)***
Fraction of links in a group	-0.087	0.007	-0.132	-0.160
	(0.257)	(0.276)	(0.294)	(0.300)
Avg. Closeness	2.365	2.076	1. 716	1.498
	(1.148)**	(0.917)***	(1.175)	(1.069)
Fraction of links in a group	0.062	0.209	0.033	0.100
	(0.245)	(0.265)	(0.285)	(0.296)
Avg. Betweenness	-0.500	-1.630	0.281	-1.285
	(0.928)	(1.138)	(1.024)	(1.246)
Fraction of links in a group	0.291	0.446	0.153	0.251
	(0.201)	(0.263)	(0.262)	(0.290)
Avg. Eigenvector	4.857	6.561	4.422	5.674
	(2.095)**	(2.120)***	(2.332)*	(2.454)**
Fraction of links in a group	0.021	0.052	-0.038	-0.043
	(0.249)	(0.283)	(0.285)	(0.311)
Avg. Degree	1.643	2.559	1.443	2.547
	(0.832)*	(0.809)***	(0.856)*	(0.929)***
Fraction of links in a group	0.018	-0.004	-0.037	-0.140
	(0.264)	(0.287)	(0.295)	(0.315)

Notes: Columns (1) and (2) contain the results of regressions on the whole sample while columns (3) and (4) contain the results of regressions on the subsample that excludes the groups with no friendship links. N=924 for columns (1) and (2); N=753 for columns (3) and (4). Each estimated coefficient on average centrality is obtained from a different regression. All regressions control for average group characteristics and school fixed effects. Standard errors are clustered at the school level. * p<0.10; ** p<0.05; *** p<0.01

Table 5: Correlation across network centrality measures

	Intercentrality	Bonacich	Closeness	Betweeness	Eigenvector	Degree
Intercentrality	1					
Bonacich	0.954	1				
Closeness	0.438	0.477	1			
Betweeness	0.170	0.254	0.361	1		
Eigenvector	0.768	0.796	0.564	0.184	1	
Degree	0.794	0.854	0.669	0.275	0.810	1

Table 6: Relative importance of centrality measures

	(1)	(2)	(3)	(4)
Outcome: GGKT (N=924)				
Average centrality in the group based on				
Intercentrality	0.006	0.008	0.009	0.013
	(0.006)	(0.003)**	(0.004)**	(0.004)***
Bonacich	0.032			
	(0.066)			
Closeness		0.877		
		(1.239)		
Eigenvector			0.317	
			(2.097)	
Degree				-1.376
				(1.008)
Outcome: GGKT (N=924)				
Intercentrality	0.006			
	(0.006)			
Bonacich	0.032	0.070	0.081	0.152
	(0.066)	(0.033)**	(0.036)**	(0.052)***
Closeness		0.788		
		(1.334)		
Eigenvector			0.131	
			(2.124)	
Degree				-2.346
				(1.311)*
Outcome: GMT (N=924)				
Intercentrality	0.015	0.011	0.008	0.010
	(0.009)	(0.003)***	(0.004)**	(0.004)**
Bonacich	-0.035			
	(0.088)			
Closeness		0.199		
		(1.021)		
Eigenvector			2.476	
			(2.650)	
Degree				0.192
				(1.306)

Notes: All regressions control for individual characteristics and school fixed effects. Standard errors are clustered at the school level. * $p<0.10$; ** $p<0.05$; *** $p<0.01$

Table 7: Group outcomes, Maximum and Minimum centrality

	(1) GGKT	(2) GMT
Panel A -Maximum centrality in the group based on		
Intercentrality	0.003	0.004
	(0.001)***	(0.001)***
Bonacich	0.040	0.042
	(0.010)***	(0.011)***
Closeness	1.066	1.419
	(0.458)**	(0.478)***
Betweenness	-0.679	-0.849
	(0.471)	(0.508)*
Eigenvector	1.775	2.543
	(0.669)***	(0.702)***
Degree	0.744	1.059
	(0.303)**	(0.345)***
Panel B -Minimum centrality in the group based on		
Intercentrality	0.000	0.005
	(0.006)	(0.008)
Bonacich	-0.004	0.032
	(0.033)	(0.040)
Closeness	1.239	0.210
	(0.836)	(0.702)
Betweenness	3.583	0.279
	(1.928)*	(1.857)
Eigenvector	1.733	6.197
	(3.367)	(3.507)*
Degree	-0.344	1.140
	(0.912)	(0.991)

Notes: Each estimated coefficient on centrality is obtained from a different regression. All regressions control for group average characteristics and school fixed effects. Leader's (panel A) or weakest link's (panel B) math baseline test, IPEMT, is also included in the model specification. Standard errors are clustered at the school level. * p<0.10; ** p<0.05; *** p<0.01

Table 8: Relative importance of maximum centrality measures

	(1)	(2)	(3)	(4)
Outcome: GGKT (N=924)				
Max centrality in the group based on				
Intercentrality	-0.004			
	(0.004)			
Bonacich	0.080	0.040	0.051	0.062
	(0.043)*	(0.010)***	(0.015)***	(0.019)***
Closeness		-0.028		
		(0.575)		
Eigenvector			-0.865	
			(0.819)	
Degree				-0.791
				(0.475)
Outcome: GMT (N=924)				
Max centrality in the group based on				
Intercentrality	0.002	0.004	0.003	0.003
	(0.004)	(0.001)***	(0.001)*	(0.002)*
Bonacich	0.015			
	(0.054)			
Closeness		0.551		
		(1.069)		
Eigenvector			1.068	
			(1.075)	
Degree				0.141
				(0.658)
Outcome: GMT (N=924)				
Max centrality in the group based on				
Intercentrality	0.002			
	(0.004)			
Bonacich	0.015	0.042	0.030	0.040
	(0.054)	(0.011)***	(0.017)*	(0.024)*
Closeness		0.516		
		(1.059)		
Eigenvector			0.986	
			(1.092)	
Degree				0.070
				(0.674)

Notes: All regressions control for group average characteristics and school fixed effects. Leader's math baseline test, IPEMT, is also included in the model specification. Standard errors are clustered at the school level. * p<0.10; ** p<0.05; *** p<0.01.

Table 9: Individual outcomes and network centrality

	(1) IPOMT-IPEMT	(2) IPOMT-IPEMT
	All groups	Groups with no network links
Avg. Intercentrality	0.004	0.009
	(0.001)***	(0.004)**
Own Intercentrality	0.003	0.003
	(0.001)***	(0.003)
Avg. Bonacich	0.043	0.064
	(0.011)***	(0.037)*
Own Bonacich	0.030	0.018
	(0.011)***	(0.032)
Avg. Closeness	0.894	4.371
	(0.471)*	(2.020)**
Own Closeness	0.442	0.198
	(0.408)	(1.847)
Avg. Betweeness	0.153	-0.071
	(0.612)	(2.119)
Own Betweeness	0.404	0.600
	(0.412)	(1.306)
Avg. Eigenvector	2.115	6.606
	(0.951)**	(4.710)
Own Eigenvector	1.127	3.222
	(0.937)	(3.537)
Avg. Degree	1.145	3.200
	(0.323)***	(1.629)*
Own Degree	0.761	0.912
	(0.299)**	(1.396)

Notes: Columns (1) contain the results of regressions on the whole sample while columns (2) contain the results of regressions on the subsample of groups with no friendship links. N=3401 for column (1) and N=579 for column (2). Each coefficient on average centrality is obtained from a different regression. Average centrality is defined by group average centrality, excluding own centrality. All regressions control for individual characteristics and school fixed effects. Standard errors are clustered at the school level. * p<0.10; ** p<0.05; *** p<0.01

Table 10: Individual outcomes, maximum and minimum centrality

	(1)		(2)		(3)
	IPOMT-IPEMT		IPOMT-IPEMT		IPOMT-IPEMT
Avg. Intercentrality	0.007	Max Intercentrality	0.002	Min Intercentrality	0.006
	(0.002)***		(0.001)***		(0.009)
Own Intercentrality	-0.000	Own Intercentrality	0.004	Own Intercentrality	0.003
	(0.000)		(0.002)		(0.001)**
Avg. Intercentrality	0.003	Max Intercentrality	-0.000	Min Intercentrality	0.000
*Own Intercentrality	(0.002)	*Own Intercentrality	(0.000)	* Own Intercentrality	(0.000)
Avg. Bonacich	0.052	Max Bonacich	0.021	Min Bonacich	0.011
	(0.022)**		(0.008)**		(0.060)
Own Bonacich	0.002	Own Bonacich	0.027	Own Bonacich	0.021
	(0.005)		(0.031)		(0.022)
Avg. Bonacich	0.005	Max Bonacich	-0.001	Min Bonacich *	0.003
*Own Bonacich	(0.032)	*Own Bonacich	(0.003)	Own Bonacich	(0.011)
Avg. Closeness	3.469	Max Closeness	2.721	Min Closeness	-0.166
	(2.128)		(1.377)*		(2.082)
Own Closeness	-4.326	Own Closeness	2.745	Own Closeness	0.370
	(3.998)		(1.983)		(2.060)
Avg. Closeness	2.644	Max Closeness	-3.766	Min Closeness	0.112
* Own Closeness	(2.514)	* Own Closeness	(2.634)	* Own Closeness	(3.713)
Avg. Betweenness	-0.324	Max Betweenness	-0.173	Min Betweenness	-1.047
	(1.032)		(0.450)		(2.087)
Own Betweenness	7.371	Own Betweenness	0.224	Own Betweenness	0.233
	(11.206)		(1.029)		(0.481)
Avg. Betweenness	-0.144	Max Betweenness	1.377	Min Betweenness	16.865
* OwnBetweenness	(0.912)	* Own Betweenness	(5.563)	* Own Betweenness	(16.384)
Avg. Eigenvector	3.860	Max Eigenvector	0.980	Min Eigenvector	5.946
	(1.341)***		(0.516)*		(5.759)
Own Eigenvector	-42.121	Own Eigenvector	3.095	Own Eigenvector	0.957
	(12.954)***		(1.481)**		(1.050)
Avg. Eigenvector	2.639	Max Eigenvector	-16.492	Min Eigenvector	-34.865
* Own Eigenvector	(1.419)*	* Own Eigenvector	(6.630)**	* Own Eigenvector	(93.181)
Avg. Degree	2.173	Max Degree	0.677	Min Degree	0.896
	(0.597)***		(0.253)***		(1.305)
Own Degree	-7.687	Own Degree	1.317	Own Degree	0.669
	(4.833)		(0.710)*		(0.385)*
Avg. Degree	1.374	Max Degree	-2.685	Min Degree	-1.149
* Own Degree	(0.827)	* Own Degree	(1.769)	* Own Degree	(7.885)

Notes: All regressions control for individual characteristics and school fixed effects. Average centrality is defined by group average centrality, excluding own centrality. Standard errors are clustered at the school level. * $p<0.10$; ** $p<0.05$; *** $p<0.01$

Table 11: Network Centrality, cognitive and non-cognitive skills

	(1)	(2)	(3)	(4)	(5)	(6)
		Intercentrality			Katz-Bonacich	
IPEMT	0.251	0.181	0.204	0.014	0.005	0.011
	(0.294)	(0.292)	(0.294)	(0.037)	(0.037)	(0.037)
Risk-taking		-0.396	-1.579		-0.059	-0.239
		(3.117)	(3.100)		(0.391)	(0.386)
Time preference		3.210*	3.261*		0.465**	0.461**
		(1.749)	(1.739)		(0.220)	(0.216)
Competition		15.98***	15.13***		2.078***	1.977***
		(5.036)	(5.045)		(0.632)	(0.628)
Controls	No	No	Yes	No	No	Yes
		Closeness			Betweenness	
IPEMT	0.0015	0.001	0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Risk-taking		0.039***	0.037***		0.003	0.002
		(0.011)	(0.012)		(0.009)	(0.009)
Time preference		0.004	0.005		0.006	0.005
		(0.006)	(0.006)		(0.005)	(0.005)
Competition		0.135***	0.131***		0.032**	0.0312**
		(0.019)	(0.019)		(0.015)	(0.015)
Controls	No	No	Yes	No	No	Yes
		Eigenvector			Degree	
IPEMT	0.0005	0.0003	0.0004	0.0019	0.001	0.001
	(0.0005)	(0.0005)	(0.0005)	(0.0014)	(0.001)	(0.001)
Risk-taking		0.012**	0.009*		0.015	0.009
		(0.005)	(0.005)		(0.015)	(0.015)
Time preference		0.002	0.003		0.010	0.010
		(0.003)	(0.003)		(0.008)	(0.008)
Competition		0.049***	0.046***		0.112***	0.106***
		(0.008)	(0.008)		(0.023)	(0.023)
Controls	No	No	Yes	No	No	Yes

Notes: The pre-experiment math test (IPEMT) is used as a measure of cognitive ability. Risk-taking=1 if the student chooses to draw the highest number of pencils. Time preference=1 if the student chooses to get 6 candies two days later as opposed to 4 candies tomorrow. Competition=1 if the student chooses to compete the competition game with an anonymous classmate. Controls include student and household characteristics. N=512. Standard errors are clustered at the school level. * p<0.10; ** p<0.05; *** p<0.01