Teaching Practices and Friendship Networks*

Euncheol Shin† Syngjoo Choi‡ Booyuel Kim§ Eungik Lee¶ Yoonsoo Park‖

First version: January 31st, 2022 This version: January 31st, 2022

Preliminary Version

Abstract

Promoting social capital has long been an important issue in social sciences. This paper argues that teaching practices can stimulate social capital at both the individual and the classroom levels by evaluating the impact of project-based learning, a student-centered teaching pedagogy program. The sample in this paper consists of 1,239 7th grade students from 12 middle schools in the Republic of Korea. We measure changes in students' friendship network and directed altruism with comprehensive friendship surveys and incentivized dictator game experiments conducted before and after the intervention. We find that the project-based learning program positively affects social capital by expanding students' friendship networks and being more generous toward their peers, especially those not in direct friendship and without homophilous traits. Moreover, structural estimations suggest that the program also reduces friendship formation costs among the students, especially for those studying in the same classroom. Our results support the idea that teaching practice focusing on student-centered learning can be considered an effective educational policy to support social capital formation among students.

Keywords: Social network; Social capital; Teaching practices; Dictator game; Lab-in-the-field experiments

*We have benefited from conversations with Matthew Chao, Federico Echenique, Donggyu Kim, Sanghyun Kim, Youngse Kim, Kirill Pogorelskiy, Adam Wierman, Leeat Yariv, and Chamna Yoon. We also thank Youngjae Hwang, Jieun Hong, Wonsik Ko, Jahyeon Koo, Eunjae Lee, Jiyeong Lee, Myeonghwan Na, Minseon Park, and Myungkou Shin for their superb research assistance. This paper is financially supported by Korea Development Institute Research Grant, KAIST Faculty Research Grant, and Creative-Pioneering Researchers Program through Seoul National University. Author names are in alphabetical order except for the first author.

†KAIST College of Business, 85 Hoegiro, Dongdaemun-Gu, Seoul 02455, Republic of Korea. Email: eshin.econ@kaist.ac.kr; Website: https://sites.google.com/site/euncheolshin00/.

‡Department of Economics, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea. Email: syngjooc@snu.ac.kr.

§Corresponding author. Graduate School of Environmental Studies, Seoul National University1, Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea. Email: booyuel@snu.ac.kr.

¶Department of Economics, New York University, 19 West 4th Street, New York, NY 10003, United States. Email: lee.eungik@gmail.com.

‖Department of Economics, Sookmyung Women's University 53-12, Chungpa-dong Yongsan-gu, Seoul 04310, Republic of Korea. Email: yoonpark@sm.ac.kr.

1 Introduction

School education is undeniably the best means for individuals to accumulate human capital and can also promote the creation of social capital in a society. A growing numbers of studies confirm that social capital is positively associated with years of schooling (Milligan et al. 2004; Glaeser et al. 2007; Helliwell and Putnam 2007) and is influenced by the teaching practices with which students are taught (Algan et al. 2013). In particular, teaching practices stimulating student group work provide students opportunities to engage with their peers as part of learning in the classroom. Such interactive experiences can change students' perception regarding the value and cost of their relationship with peers and affect the social connections among them. Because peer relationships in school can affect academic achievement (Fletcher et al. 2020), adult mental health and life satisfaction (Narr et al. 2019; Powdthavee 2008), and economic performances in markets (Cohen et al. 2008), educational programs that enable a change in social cohesion among students is one tool that policymakers can resort to for enhancing social capital. In this paper, we aim to have a deeper understanding of the potential mechanisms through which an education intervention works toward building social capital.

We examine the impact of experiencing a student-centered teaching pedagogy program on students' value of friendship and the formation of friendship networks in school. A student-centered teaching pedagogy program, called *project-based learning* (henceforth abbreviated as PBL), was introduced in six schools by the local education authority in a large city in the Republic of Korea. The program trains and encourages teachers to change their teaching practices from the conventional lecture-oriented classes to horizontal teaching practices in which students work in groups and do projects together, and thereby stimulating interpersonal interactions among students in a classroom. For comparison purposes, we carefully selected six comparison schools that were not practicing PBL, located in proximity to treatment schools. There were no observable differences between PBL treatment schools and comparison schools, thereby mimicking balanced samples that are achieved by random assignment.

To evaluate the impacts of the PBL program on students' assessment of friendship and friendship formation, we conducted a school-wide survey of friendship nomination and an incentivized lab-in-the-field experiment on directed altruism before and after the intervention in both treatment and comparison schools. All students were asked to list up to 10 friends among the entire list of students in the same grade at the same school, thereby enabling us to construct an entire friendship network at the school level. In order to elicit the value of peer relationships, each student was asked to divide endowed money between self and a randomly matched student whose name and student ID was revealed. Students repeated this decision with 10 randomly matched students, allowing us to measure their directed altruism toward the matched students. The combination of this experiment with the friendship survey enables us to comprehend the variations of directed altruism over distances defined in the friendship network, as done in Goeree et al. (2010).

We find that the exposure to horizontal teaching practices influences the value structure of friendship across network distances and the formation of friendship. First, introduction of PBL increases directed altruism by about 6 percentage points, which amounts an increase of 23% in giving, compared to the baseline level of giving in the comparison schools. By dissecting this treatment effect along the dimension of friendship network distances, we further establish that the PBL effect on giving is driven by its effect on the value of indirect friendship, that is, the values of the relationship with partners who are not nominated as a friend by the giver but connected through friendship networks. The magnitude of the PBL effect ranges from 5 percentage points to 7 percentage points across distances. In contrast, we find no PBL effect on the value of direct friendship, that is, the value of the relationship with partners who are nominated as a friend by the giver and the value of strangers who are not connected in the network.

Furthermore, the PBL also affects the homophilic composition of friendship formation along with classroom and gender. Focusing on co-education schools with mixed-gender classrooms, the PBL increases the ratio of friendship nomination within the same classroom by 8 percentage points but decreases the ratio of friendship nomination for the same gender by 2.5 percentage points. Because of the feature that it was introduced at the classroom level, the PBL intervention substantially reduced gender homophily within the classroom by 4.6 percentage points. A similar pattern is established in other types of schools.

While adopting the economics approach on the formation of networks, we consider a simple model in which individuals assess the cost and benefit of friendship formation and make decisions accordingly. By implementing a model-based structural estimation on the cost of making friends and applying the difference-in-differences method in the estimated costs of friendship formation, we infer the PBL effect on the cost of friendship formation. We find that the stimulation of interactions through the PBL program significantly reduces the cost of friendship formation not only within the classroom but also outside the classroom. Moreoever, there also appears to be a reduction in the linking cost when making friends with the opposite gender; however, it is imprecisely estimated.

2 The Daegu PBL Program

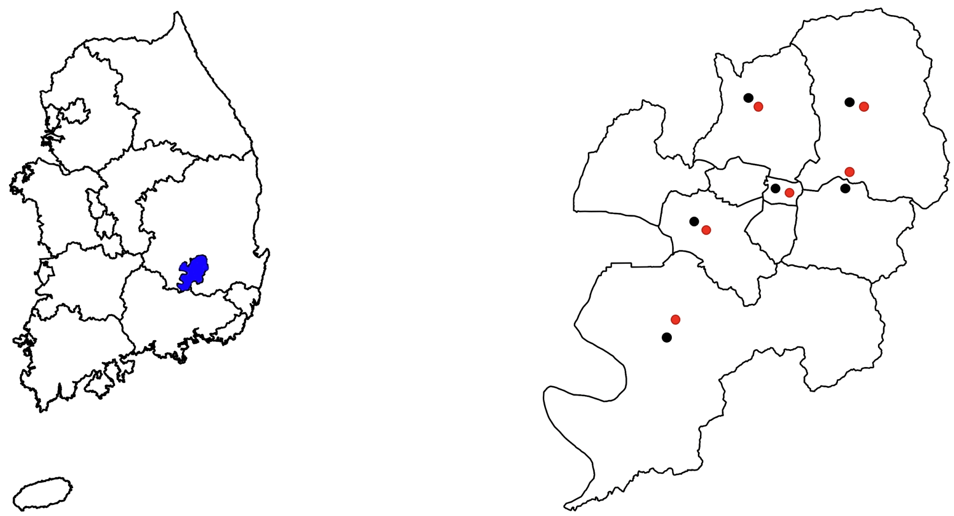
To examine the effect of introducing the PBL-type classes[1] on the value of friendship among students in the seventh grade (first-year students in middle schools in South Korea), we collaborated on the design of the PBL intervention and the implementation of a survey and experiments with the Daegu Metropolitan Office of Education (henceforth, DMOE) and the Korea Development Institute (henceforth, KDI). Six treatment schools (indicated by the black dots in the right panel of Figure 1) in Daegu, the fourth largest city in Korea (indicated by the blue colored region in the left panel of Figure 1), were selected to introduce PBL-type classes in the fall semester of 2016. Six comparison schools (indicated by red dots in the right panel of Figure 1) were chosen based on their establishment type (private or public), gender type (all-boy, co-educational with mixed-gender classrooms, or co-educational with single-gender classrooms), and proximity to the treatment schools.[2] As will be shown later, this effort of selecting sample schools guarantees that there is no statistical difference between PBL treatment schools and comparison schools in observed characteristics.

The DMOE requested teachers in the treatment schools to implement PBL-type classes during the fall semester of 2016. Specifically, the teachers were required to implement at least two in-class group projects, each of which was required to last for five or more class hours. To assist teachers in the treatment schools in running PBL-type classes, the DMOE and KDI offered a four-day (or 30-hour) training workshop on how to design and prepare PBL-type classes approximately two months prior to the fall semester of 2016, as well as provided ongoing consultation and coaching services throughout the semester. Table 1 summarizes the timing of the training and consultation programs. In contrast, no such program was offered to teachers in the comparison schools.

[1]PBL can be defined as learning that focuses on group projects in which students investigate solutions through asking questions, debating ideas, and communicating with other students. In a typical PBL class, the teacher initially introduces the project that students need to address in the group. This consists of providing the background information, the main question of the project, and the instructions on the specific tasks that students need to accomplish. Then, students develop a group plan for the project including brainstorming ideas, collecting information, and assigning different roles and tasks among group members. Finally, their end product is presented to other students (Helle et al. 2006).

[2]Among the six comparison schools, five were chosen within the same administrative district as the treatment schools. Only one comparison school was chosen from a neighboring administrative district because the administrative district in which the treatment school is located does not have a school with the same characteristics (all-boy school) as the treatment school. The average distance between a treatment school and its matched comparison school is approximately 1.2 kilometers.

Figure 1: Location of Daegu City and the Sample Schools



(a) Location of Daegu City (b) Location of the sample schools

Notes: The left panel (a) shows the map of South Korea and the right panel (b) displays the city map of Daegu with 9 districts.

Table 1: Timeline for 2016 PBL intervention and surveys

	Before 2016 Fall Semester (June–July)	During 2016 Fall Semester (August–December)
PBL intervention	four-day teacher training and workshop School 1: 6/11 (9), 6/18 (9), 6/29 (4), 7/6 (8) School 2: 6/11 (8), 6/12 (7), 7/22 (8), 7/23 (7) School 3: 6/25 (8), 7/21 (7), 7/22 (8), 7/23 (7) School 4: 7/1 (7), 7/2 (8), 7/19 (7), 7/20 (8) School 5: 7/9 (8), 7/16 (7), 7/19 (8), 7/20 (7) School 6: 7/20 (8), 7/21 (7), 7/22 (8), 7/23 (7) * *Training and workshop hours in parentheses*	Regular consulting and coaching
Baseline and endline surveys		Baseline: 5/22–8/31 Endline 12/14–12/23

3 Survey and Experimental Measurement

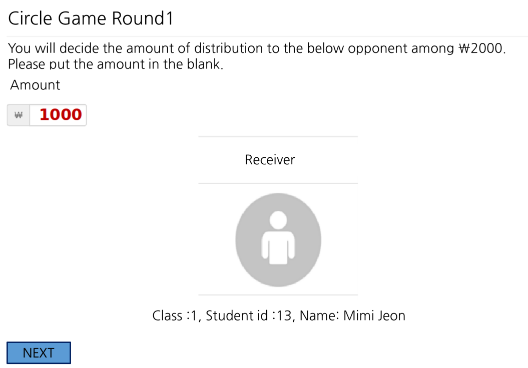
For both the treatment and comparison schools, a student survey with lab-in-the-field experiments was conducted at the beginning and end of the 2016 fall semester, respectively. The baseline survey, which was conducted between August 22 and August 31, surveyed 1,239 students from four randomly selected seventh grade classrooms at each of the treatment and comparison schools. The endline survey was conducted between December 14 and December 23 and collected the same information as that collected in the baseline survey. Of the 1,239 students who completed the baseline survey, 1,130 (91.2%) completed the endline survey, and the rate of survey attrition was statistically indifferent between the treatment and comparison schools. The final sample for this paper consists of 1,130 students who participated in both the baseline and endline surveys.

The baseline and endline surveys consist of three components: 1) background survey, 2) lab-in-the-field experiments, and 3) friendship network survey. Students participated in the surveys while sitting in their classrooms and using the laptops provided to each of them. First, the background survey includes questions about students' demographic and socio-economic status, five math problems to evaluate students' cognitive abilities, and various questionnaires to assess students' non-cognitive traits such as the Rosenberg self-esteem scale (Rosenberg 1965) and the big-5 personality test (Gosling et al. 2003).

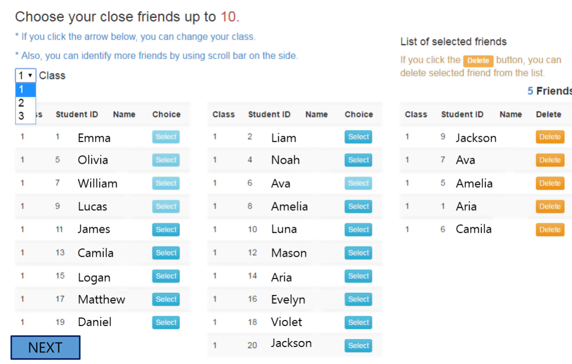
Second, the lab-in-the-field experiments include the dictator game designed to measure directed altruism when combined with friendship survey information. We interpret the amount that is given to a matched student as the value of friendship that a dictator assigns to that matched student. Each student participated in 10 rounds of the dictator game with randomly assigned peers in the same grade at the same school. For each round of the dictator game, students were given KRW 2,000 (approximately USD $2) and were asked to allocate the endowment between themselves and a randomly matched recipient whose name and student ID were displayed on the decision screen, as illustrated in Figure 2-(a). At the end of the survey, we randomly selected one dictator game that counts for the actual payments for each participant, which was selected at random from the set of games in which that participant joined as a dictator (10 rounds of the game) or as a recipient. For the selected game that counts for actual payments, the participant was paid according to a decision made by a dictator. Because the dictator game was repeated in both the baseline and endline surveys, we used the same set of randomly matched recipients from the baseline for the endline dictator game. In the endline survey, the order of recipients for the dictator game was randomly shuffled.

Third, the friendship network survey asked each student to list up to 10 friends among the entire list of students (clustered by classrooms on the survey screen) in the same grade at the same school, as illustrated in Figure 2-(b). We implemented the friendship network survey not only for the randomly selected four classrooms but also for all the remaining seventh grade classrooms in each school in order to examine the complete picture of school-wide friendship network at a school level. This yields a network sample of 2,792 students who participated in both the baseline and endline surveys. There is no systematic attrition bias between the treatment and comparison schools. This enables us to make the computation of friendship distance between two matched students in the dictator game based on the entire friendship network at the school level.

Figure 2: Example Screens of the Dictator Game and the Friendship Survey



(a) dictator game screenshot



(b) friendship network survey screenshot

Notes: In (a), each subject playing the dictator's role can see the recipient's name with his/her classroom number and in-class identification number. Then, the subject inserts an amount between 0 and 2000 to be given to the recipient selected on the screen. When the subject clicks the "next" button on the left-bottom corner of the screen, a new randomly chosen recipient appears in the next round. In (b), each subject first selects a class number at the attending school. Then, all the student names in the selected class appear on the screen. The subject identifies the student as a close friend in our friendship survey by clicking the blue button next to a student's name. We also allowed the subjects to cancel their choices; by clicking the orange button next to a chosen student's name, the corresponding student is deleted from the list of close friends. When the subject clicks the "next" button on the left-bottom corner of the screen, a new survey screen appears. All the names in the above screenshots are pseudonyms.

4 Balance Checks and Empirical Strategy

4.1 Balance Checks

We first examine whether the baseline characteristics between the treatment and comparison schools are well balanced. Although our research design is not a randomized controlled trial, the comparison school matching process led to a treatment-comparison balance in the baseline characteristics, including (pre-intervention) outcome variables. From the participating students' perspective, this PBL program implemented by the DMOE was exogenous to them, and there was no endogenous selection of students into the program. Moreover, the middle school admission process in the Republic of Korea is based on within-district random assignment with the shortest distance priority. Since the comparison school matching process ensures that we have the closest school to each treatment school, while holding school characteristics (male-only or co-educational with mixed or single gender classrooms) constant, the baseline characteristics between the treatment and comparison schools are balanced.

Table 2 confirms that we have a well-balanced sample in terms of students' demographic and socioeconomic status, cognitive and non-cognitive traits including personalities, average offers of baseline dictator game, and distances and composition of friendship. Column (1) presents the sample means of variables for students in the comparison schools and Column (2) presents the mean differences for each of the variables between the treatment and comparison schools, reporting the standard errors clustered at school level in parentheses. Furthermore, Panel A of Table 2 reports the demographic and socio-economic information. 42.2% of the sample is female students and students' average height and weight are 162 cm and 52.5 kg, respectively; 45.6 percent of the students are first-born children. The parents are in their mid-40s (43.8 years for mothers and 46.3 years for fathers) and more than half of them have completed two-year college education or above (56.6 percent for mothers and 61.5 percent for fathers). Panel B of Table 2 includes the baseline outcomes of dictator games. The proportion of direct friends (distance 1) among 10 matched peers was 9.9 percent in the baseline. The proportions of indirect friends were 18.5 percent (distance 2), 26.9 percent (distance 3), 22.1 percent (distance 4), 15.0 percent (distance 5), and 7.5 percent (distance 6–11), respectively. The baseline average offer in the dictator experiment was 24.2 percent (approximately KRW 500 of the KRW 2,000), which is similar to the average amount (28.4 percent) reported in other dictator experiments (Engel 2011).[3]

4.2 Difference-in-Differences Estimations

Value of friendship. To estimate the causal effect of the PBL intervention on the value of friendship, we employ the difference-in-differences (or fixed-effect) strategy. Specifically, we consider the following regression equations:

$$Share_{ijst} = \beta_0 + \beta_1 PBL_{st} + \gamma_s + \delta_t + X_{ijst} + \varepsilon_{ijst}, \tag{4.1}$$

$$Share_{ijst} = \beta_0 + \beta_1 PBL_{st} + \gamma_s + \delta_t + X_{ijst} + dist_{ijst} + \varepsilon_{ijst}, \tag{4.2}$$

where $Share_{ijst}$ represents the share of offer in the dictator game by student i (dictator) to student j (recipient) in school s at time t. PBL_{st} denotes the indicator variable that takes the value of 1 if school s was under the PBL intervention at time t. Specifically, PBL_{st} is coded to take 1 for the treatment schools at the endline survey, and 0 otherwise. γ_s refers to school fixed effects, a list of dummy variables indicating each of the treatment and comparion schools (i.e., 11 dummies excluding the reference

[3]Our subjects socially know the recipients from the same grade of their school and the setting we have is comparable to that of Goeree et al. (2010), where fifth and sixth grade students in the same school make average offers of 34 percent to their friends, friends-of-friends, and strangers (defined as those of distance 3 or greater). One of the reasons why the average offer in our dictator game, which is 24.2%, is smaller than that in Goeree et al. (2010), which is 34%, is that Goeree et al. (2010) matched each dictator with three direct friends (distance 1), three indirect friends (distance 2), and four others (distance 3 or higher), whereas much lesser direct friends and distance-two friends were matched (9.9% and 18.5%, respectively) in our setting.

Table 2: Descriptive Statistics and Baseline Balancing Test

	Control	Treated		Control	Treated
Panel A: Individual characteristics (n= 1,130)			Panel B: Dictator game (n= 11,300)		
Female (=1)	0.422	0.021	Proportion of distance 1 matched	0.099	-0.003
	[0.494]	(0.119)		[0.299]	(0.018)
Height (cm)	162.064	0.614	Proportion of distance 2 matched	0.185	0.020
	[7.480]	(0.980)		[0.389]	(0.041)
Weight (kgs)	52.520	0.909	Proportion of distance 3 matched	0.269	0.023
	[10.443]	(1.316)		[0.444]	(0.052)
Mother's age	43.772	0.164	Proportion of distance 4 matched	0.221	0.020
	[4.042]	(0.392)		[0.415]	(0.035)
Father's age	46.290	0.194	Proportion of distance 5 matched	0.150	-0.050
	[3.681]	(0.398)		[0.357]	(0.043)
Mother's education	0.566	0.092	Proportion of distance 6-11 matched	0.075	-0.024
	[0.496]	(0.067)		[0.263]	(0.031)
Father's education	0.615	0.052	Proportion of distance 12 matched	0.002	0.015
	[0.487]	(0.076)		[0.043]	(0.008)
Birth order	1.634	-0.081	Dictator's share	0.242	-0.014
	[0.657]	(0.059)		[0.274]	(0.024)
First born	0.456	0.088	Panel C: Friendship composition (n= 1,130)		
	[0.499]	(0.050)			
Math score	2.721	-0.061	Same classroom (p_1)	0.571	-0.008
	[1.577]	(0.284)		[0.281]	(0.046)
Self esteem	3.164	-0.035	nomination rate	0.886	-0.016
	[0.495]	(0.026)		[0.187]	(0.019)
Personality (outgoing)	3.598	-0.057	Different gender & Different classroom ($P(dg, dc)$)	0.025	-0.014
	[1.006]	(0.040)	(6 coed schools only, n = 505)	[0.143]	(0.018)
Personality (agreeableness)	3.322	0.058	Same gender & Different classroom ($P(sg, dc)$)	0.469	0.005
	[0.746]	(0.048)	(6 coed schools only, n = 505)	[0.282]	(0.072)
Personality (conscientiousness)	3.408	0.002	Different gender & Same classroom ($P(dg, sc)$)	0.065	0.021
	[0.875]	(0.057)	(6 coed schools only, n = 505)	[0.137]	(0.022)
Personality (stability)	3.016	-0.053	Different gender & Same classroom ($P(sg, sc)$)	0.441	-0.011
	[0.825]	(0.066)	(6 coed schools only, n = 505)	[0.283]	(0.074)
Personality (openness)	3.544	0.013	Nomination rate ($P(nomination)$)	0.858	0.020
	[0.872]	(0.079)	(6 coed schools only, n = 505)	[0.202]	(0.033)

Notes: The above table reports descriptive statistics in the baseline survey. Standard deviations are given in square brackets, and standard errors clustered at school level are in parentheses. Treated columns display the difference in means between the treatment and comparison groups. Mother's education and Father's education are dummy variables that equal one when parents' education level is above two-year college graduation. First born is a dummy variable that equals to one when birth order is one. Math score counts the number of correct answers from five math questions. Self-esteem measures the average of 10 Rosenberg Self-esteem scale questions (4-point likert scale). Personality comes from TIPI (Ten Item Personality Inventory) with 5-point likert scale.

group). γ_s controls any confounding factors of each school, either observable or unobservable, that are stable between the baseline and endline surveys. δ_t denotes time fixed effect, a dummy variable indicating the timing of survey (coded to take 1 for the endline survey, treating the baseline survey as the reference group). δ_t controls any temporal changes of the outcome variable ($Share_{ijst}$) that are common across each of the treatment and comparison schools. X_{ijst} further controls students' demographic and socio-economic status, cognitive (math score) and non-cognitive (self-esteem and personalities) traits, and friendship homophily. ε_{ijst} is an error term clustered at individual level. β_1, which is of main interest in this study, estimates the difference in the temporal changes in the outcome variable between the treatment and comparison schools. Given that the treatment and comparison schools are well balanced in terms of various characteristics, including the average offer in the dictator game as reported in Table 2, β_1 likely identifies the causal effect of the PBL intervention on the outcome variable.

Network formation. In order to measure the causal effects of the PBL intervention on students' network formation, we consider the following regression equations of difference-in-differences analysis:

$$Characteristic_{it} = \beta_0 + \beta_1 PBL_{st} + \gamma_s + \delta_t + X_{it} + \varepsilon_{ist}, \tag{4.3}$$

where $Characteristic_{it}$ represents a network characteristic of student i at time t. We consider five network characteristics: (1) $P(dg, dc)$, which represents the proportion of friends of different gender in different classrooms; (2) $P(sg, dc)$, which represents the proportion of friends of the same gender in different classrooms; (3) $P(dg, sc)$, which represents the proportion of friends of different gender in the same classrooms; (4) $P(sg, sc)$, which represents the proportion of the friends of the same gender in the same classrooms; and (5) $P(nomination)$, which calculates the nomination rate, the ratio of outdegree to ten, the upper limit of the number of friendship nominations.[4] For the explanatory variables, γ_s and δ_t refer to the school fixed effects and time fixed effects, respectively. X_{it} includes student i's gender, height, and weight. Different from the difference-in-difference analysis, other control variables, such as demographic and SES information, are not included because these variables are collected only from the students who participated in the dictator game experiment. ε_{ijst} is an error term clustered at school level. Again, β_1 identifies the difference in the temporal changes in the network characteristic variable between the treatment schools and the comparison schools.

5 Results

5.1 Value of Friendship

First, we analyze the average treatment effects of the PBL intervention on students' value of friendship as measured by the dictator game. Table 3 shows our educational intervention increased students' offers by 5.8 percentage points (22.6 percent) overall and the results are robust when we control individual demographic and socioeconomic status characteristics, cognitive and non-cognitive traits, and homophily patterns (from column (2) to column (4) of Table 3). Since the coefficient of

[4] We calculate $\frac{1}{|N|}\sum_{i \in N} \frac{\text{outdegree}(i)}{10}$, where N is the set of all nodes in the network, and outdegree(i) is the outdegree of node $i \in N$. We consider the ratio to 10 as we allow the students to nominate up to 10 of their friends in our friendship survey.

PBL_{st} captures the total effect, including both the direct effect of PBL intervention on the value of friendship and indirect effect through the change in the friendship network distance, we further control time-varying friendship distance for randomly matched peers in the dictator games in order to disentangle direct and indirect effects of the PBL program. Column (5) of Table 3 reports the direct effect of the intervention after controlling for the friendship distance as shown in Equation (4.2). The average offer increased by 6.1 percentage points, which is similar to the total effect (5.8 percentage points increase).

Table 3: Effects of PBL Intervention on Average Offer in the Dictator Game

	Share				
Variables	(1)	(2)	(3)	(4)	(5)
PBL_{st}	0.058***	0.058***	0.058***	0.058***	0.061***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Demo and SES	No	Yes	Yes	Yes	Yes
Cog and Non-cognitive	No	No	Yes	Yes	Yes
Homophily	No	No	No	Yes	Yes
Friendship Distance	No	No	No	No	Yes
Observations	22,565	22,565	22,565	22,565	22,565
R-squared	0.014	0.020	0.025	0.058	0.077
Mean of dependent variable (t = 0 & treatment = 0)	0.257	0.257	0.257	0.257	0.257

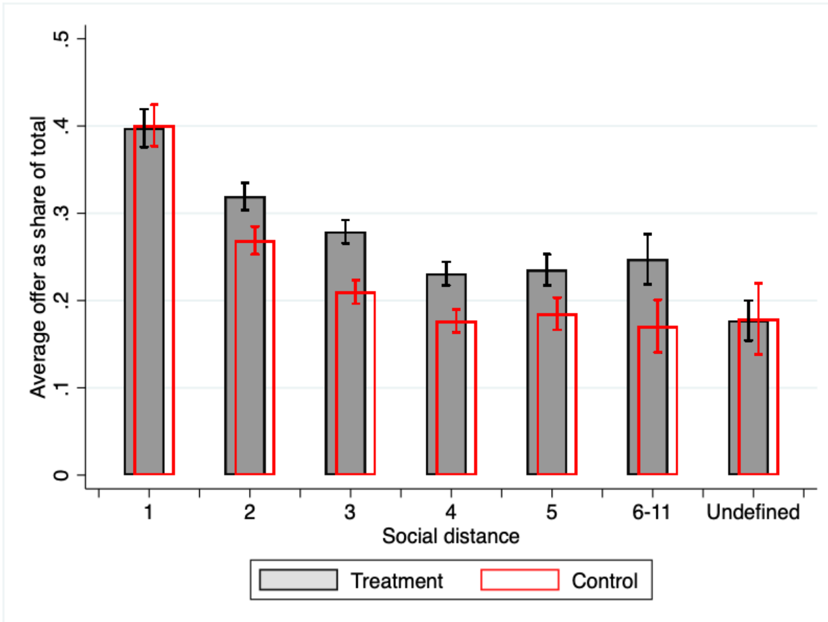
Notes: Demographic and SES controls include gender, height, weight, birth order, first-born dummy, father's age, mother's age, father's education, and mother's education. Cognitive and non-cognitive controls include math score, Rosenberg's self-esteem score, and big 5 personalities (extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience). Homophily controls include dummies for same height, same weight, same gender, and same classroom. Friendship distance ranges from 1–11 and the undefined distance was coded as 12. Robust standard errors in parentheses are clustered at individual ID. School fixed effects are included.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

5.2 Effects of PBL Intervention

Since the value of friendship is heavily dependent on social distance, we attempt to dissect these average treatment effects by friendship distances based on our extensive friendship network survey. We divide the sample by friendship distances ranging from one to eleven (we pooled the data for distances greater than six for which we have relatively fewer observations). First of all, Figure 3 shows that the average offers significantly decline with social distance, which follows the $1/d$ law of giving suggested by Goeree et al. (2010). Second, we do not observe the treatment effects for direct friends (distance 1) and strangers (distance undefined). Third, treated students give more toward indirect friends (distance 2-11). This means that the decaying effects of social distance for indirect friends are significantly reduced in the treatment schools.

The difference-in-differences regression analysis conditional on friendship distance in Table 4 con-

Figure 3: Illustration of PBL Effects on Average Offer by Social Distance



Notes: The height of the gray bars represents the average offer as a proportion of the sharing of the total by the social distance of the students in treatment schools after the intervention. The black bar at the top of each bar represents the 95% confidence interval. The bars with red borders represent the average offer by the social distance of the students in comparison schools, and the red bars indicate the corresponding 95% confidence intervals.

firms the mitigating effect of the PBL intervention on the decay rate of average offers only for indirect friends. We find the average treatment effects of an increase of 6.3 percentage points (for distance 2), increase of 7.9 percentage points (for distance 3), increase of 7.2 percentage points (for distance 4), increase of 5.3 percentage points (for distance 5), and increase of 6.0 percentage points (for distance 6-11), respectively.

Table 4: Effects of PBL Intervention on Average Offer by Social Distance

	Share						
	$dist$ = 1	$dist$ = 2	$dist$ = 3	$dist$ = 4	$dist$ = 5	$dist$ = 6–11	$dist$ = undefined
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PBL_{st}	-0.006	0.063***	0.079***	0.072***	0.053***	0.060**	0.038
	(0.025)	(0.020)	(0.018)	(0.018)	(0.020)	(0.026)	(0.040)
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,939	4,866	5,591	4,652	2,581	1,200	736
R-squared	0.050	0.056	0.069	0.059	0.078	0.082	0.123

Notes: Full controls include demographic and SES controls, cognitive and non-cognitive controls. Robust standard errors in parentheses are clustered at school level. School fixed effects are included.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

5.3 Effects on Network Formation

We now present how the PBL intervention affects formation of friendship among students. Since students in the same classroom have a greater likelihood of interacting with one another under the PBL intervention, the students in treatment schools are naturally expected to form more friendships with classmates than the students in comparison schools.

Panel A in Table 5 presents the results of the difference-in-differences analysis for six coed schools with mixed-sex classrooms. We do observe treatment effects on the friendship composition. First, in columns (3) and (4), we find that students make more friends with classmates due to the PBL intervention with statistical significance. In particular, in column (3), we find an increase of 4.6 percentage points in the nomination rate between classmates with different genders in treatment schools. This increase is higher than the 3.5 percentage points increase between classmates of the same gender. Third, we do not find any statistically significant changes in the nomination ratio, which indicates that the PBL intervention has an impact on reshaping students' network formation but does not lead to an increase in the volume of networks.

Table 5: PBL Impacts on Homophilic Network Formation

	$P(dg, dc)$	$P(sg, dc)$	$P(dg, sc)$	$P(sg, sc)$	$P(sc)$	$P(nomination)$
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: six coed schools with mixed-gender classrooms						
PBL_{st}	-0.020	-0.060***	0.046***	0.034*		0.012
	(0.010)	(0.012)	(0.006)	(0.014)		(0.030)
Controls	Yes	Yes	Yes	Yes		Yes
Observations	2,213	2,213	2,213	2,213		2,256
R-squared	0.081	0.065	0.029	0.048		0.039
Panel B: four coed schools with separate-gender classrooms						
PBL_{st}	-0.020***	0.017			0.002	0.036***
	(0.002)	(0.016)			(0.016)	(0.004)
Controls	Yes	Yes			Yes	Yes
Observations	2,382	2,382			2,382	2,414
R-squared	0.040	0.046			0.030	0.035

Notes: $P(dg, dc)$ represents the proportion of friends of different genders in different classrooms. $P(sg, dc)$ represents the proportion of friends of the same gender in different classrooms. $P(dg, sc)$ represents the proportion of friends of different gender in the same classrooms. $P(sg, sc)$ represents the proportion of the friends of the same gender in the same classrooms. $P(sc)$ represents the proportion of the friends in the same classroom. $P(nomination)$ is the ratio of outdegree to ten, the upper limit of the number of friendship nominations. Control variables include students' gender, height, and weight. Other controls used in previous analyses, such as demographic and SES information, are not included because they are available only for the students who participated in the dictator game experiment. School fixed effects are included. Robust standard errors in parentheses are clustered at the school level.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Panel B in Table 5 summarizes the results of the difference-in-differences analysis for six coed schools with separated-sex classrooms. Among three possible types of friendship compositions, we only observe a decrease of 2 percentage points in the proportion of friends of different genders in other classrooms. This result is different from the observations from the coed schools with mixed-sex classrooms. Moreover, we identify an increase of 3.6 percentage points in the friendship nomination rate, which shows that the PBL intervention has an impact on increasing the volume of networks in these schools.[5]

6 Effect of PBL on Network Formation

6.1 The Network Formation Model and Structural Estimation

We first present a simple network formation model for structural estimation. We here present the model used to analyze the results for the six coed schools with mixed-sex classrooms. The models for the schools of other types (i.e., the two male schools and the four coed schools with separated-sex classrooms) are different in terms of the number of parameters to estimate, as shown in Appendix A.

A network formation model. We consider a network formation model in which nodes simultaneously create or delete direct links to other nodes that belong to the same network. Let g be a network. We assume that the utility of each node i from node j in network g is represented by

$$u_{ij}(g) = \underbrace{\alpha_{ij} + \delta dist_{ij}(g)^{\gamma}}_{\text{value of friendship}} - \underbrace{c_{ij}\mathbf{1}\{ij \in g\}}_{\text{cost of direct link}}, \tag{6.1}$$

where $\alpha_{ij} > 0$ is the *intrinsic value* of nodes, $\delta > 0$ is a coefficient for the value driven through the indirect friendship in the network, $dist_{ij}(g_k) \in \mathbb{N} \cup \{\infty\}$ denotes the distance from node i to node j in network g_k,[6] $\gamma < 0$ is the decay rate in social distance, $c_{ij} > 0$ is the *cost* to node i of having link ij, and $\mathbf{1}\{ij \in g_k\}$ is the indicator function that represents whether there is link ij in network g_k.[7] $dist_{ij} = 1$ represents $ij \in g_k$, and $dist_{ij} = \infty$ implies that there is no path from node i to node j. Thus, if $ij \in g_k$, then $u_{ij}(g_k) = \alpha_{ij} + \delta - c_{ij}$. Moreover, if $ij \notin g_k$, then $u_{ij}(g_k) = \alpha_{ij} + \delta dist_{ij}(g_k)^{\gamma}$. Since $\gamma < 0$, if other things are equal, the value through the indirect friendship increases in γ. We

[5]We also conducted a similar difference-in-differences analysis for male schools. We find that the PBL intervention increased the proportion of the friendship nomination between classmates and the nomination rate with a statistical significance. The corresponding table is presented in Appendix B.

[6]We consider directed networks. As such, a distance between two nodes i and j is defined as the length of a shortest directed path from i to j consisting of links, provided at least one such path exists. Thus, if there is no path from node i to j, then the value of friendship only contains the intrinsic value. We refer to Jackson (2010) for terminologies and definitions related to networks in this paper.

[7]This utility function is similar to that of Jackson and Wolinsky (1996) in that the value part only depends on the distance between nodes in a given network. The difference is that they consider exponential discounting in distance, but we consider a polynomial discounting, as in Goeree et al. (2010).

assume that the cost depends on two characteristics of the nodes. Specifically, for each pair of nodes, the cost is randomly drawn from a distribution $f(\cdot;\theta(x,y))$, where $x \in \{dg, sg\}$ represents whether nodes are the same gender and $y \in \{dc, sc\}$ denotes whether nodes are classmates.[8]

Let $g+ij$ denote the network obtained by adding link ij to the the existing network g. Similarly, $g-ij$ denotes the network obtained by deleting link ij from the existing network g. Then, on one hand, if node i does not have a link to node j in network g, then i creates a link ij if and only if $u_{ij}(g+ij) \geq u_{ij}(g)$. On the other hand, if node i has a link to node j in network g, then i maintains the link ij if and only if $u_{ij}(g) \geq u_{ij}(g-ij)$. Note that the intrinsic value α_{ij} does not play any role in node i's decision as it appears on both sides of the inequalities. We assume that all the nodes make this link formation decision simultaneously.

Structural estimation. The above model is estimated by using the method of simulated moments. We choose parameter vector $\theta = (\theta(dg,dc), \theta(sg,dc), \theta(dg,sc), \theta(sg,sc)) \in \Theta = (0,10)^4$ to minimize a quadratic objective function that measures the Euclidean distance between moments driven from the observed data and the corresponding moments driven from the network formation model with random perturbations. In particular, for the network formation model, we take the observed networks as given and let the nodes create and delete links to other nodes after taking into account random costs. For parameters δ and γ, we use the estimates from the estimation from a non-linear regression in the $1/d$ law of giving framework by Goeree et al. (2010).[9] With regard to the family of parameterized cost distributions, we consider the beta distributions: depending on the nodes gender and classrooms, c_{ij} is drawn from the beta distribution $f(\cdot;\theta(x,y), 10-\theta(x,y))$ if the nodes' types correspond to $\theta(x,y)$. The sum of two parameters of each beta distribution is set to be ten. This choice is arbitrary, but it makes the interpretation of our results intuitive.[10]

In the estimation, we find the parameter θ so that the moments of the five network characteristics generated by the network formation model best match the observed corresponding moments from the observed data.[11] We let m_{emp} be the vector of empirical moments. For each possible choice of θ, we simulate the network formation model 100 times and each time we allow the nodes to create

[8]For example, if two nodes represent male students who study in a common classroom, then the cost of link formation between two nodes is drawn from a distribution $f(\cdot;\theta(sg,sc))$.

[9]See Appendix A for specifications and the estimates used for our structural estimations.

[10]For instance, each distribution is bell-shaped whenever the parameter is contained in $(1,9)$. In addition, $\frac{\theta(x,y)}{10} \in (0,1)$ represents the expected link formation cost.

[11]The five moments are (1) $P(dg,dc)$, (2) $P(dg,dc)$, (3) $P(dg,dc)$, (4) $P(dg,dc)$, and (5) $P(nomination)$.

and delete links with other nodes. Then, we calculate $m_{sim}(\theta)$, the vector of simulated moments, where each entry of the vector represents the corresponding moment in m_{emp}. Then, we choose $\widehat{\theta}$ that minimizes a quadratic criterion function as $\widehat{\theta} = \text{argmin}_{\theta \in \Theta} ||m_{sim}(\theta) - m_{emp}||^2$. Finally, in order to identify treatment effects, we calculate the difference-in-differences parameter vector $\widehat{\theta}_{did}$ defined as

$$\widehat{\theta}_{did} = (\widehat{\theta}(\text{treatment, after}) - \widehat{\theta}(\text{treatment, before})) - (\widehat{\theta}(\text{comparison, after}) - \widehat{\theta}(\text{comparison, before})),$$

where $\widehat{\theta}(\text{treatment, after})$ represents the parameter estimated from the networks in the treatment group after the PBL intervention; and other terms in the right-hand side are similarly defined. Hence, $\widehat{\theta}_{did}(x, y)$ measures the causal impact of the PBL intervention on friendship formation costs between the students corresponding to type (x, y).

To establish the distribution of $\widehat{\theta}$, we use a simple bootstrap algorithm. Specifically, we find a set of random nodes with replacement from the data and calculate new empirical moments. Then, we consider the link formation decisions for these nodes with other chosen nodes 20 times. We find $\widehat{\theta}^b$ that best matches the empirical moments. We repeat this process 500 times to obtain distributions. Then, we calculate the bootstrap difference-in-differences parameter vector of $\widehat{\theta}^b_{did}$ and use the 95% confidence interval of $\widehat{\theta}^b_{did}(x, y)$ as a confidence interval of $\widehat{\theta}_{did}(x, y)$ for $(x, y) \in \{(dg, dc), (sg, dc), (dg, sc), (sg, sc)\}$.

6.2 Results

Table 6 summarizes the impacts of the PBL intervention on friendship formation costs by homophily type. In the table, Panel A presents the results for the six coed schools with mixed-gender classrooms, and Panel B presents results for the four coed schools with separated-gender classrooms. Columns (1) and (2) show the estimates from the networks in comparison schools before and after the PBL intervention, respectively. Columns (3) and (4) are estimates from the networks in treatment schools before and after the PBL intervention, respectively. Column (5) includes the difference-in-differences estimates based on columns (1)–(4). Columns (6) and (7) gather the 95% confidence intervals of the difference-in-differences estimates based on the bootstrapping method.

In Panel A, estimates in column (5) suggest that the PBL intervention substantially reduced the cost of friendship formation within and outside the classroom. For instance, on average, the magnitude of the cost reduction between the classmates of the same gender corresponds to a 9.9 percentage points change in the dictator's offer as a share of the total. The bootstrap confidence interval indicates that this finding is statistically significant at the 95% level. In addition, the magnitude of cost reduction

Table 6: Parameter Estimates and Bootstrap Confidence Intervals for 10 Coed Schools

	Comparison		Treatment		DID	Lower CI	Upper CI
	Before	After	Before	After			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: six coed schools with mixed-gender classrooms							
different gender & different class ($\theta(dg, dc)$)	6.081	6.297	6.575	5.913	-0.878	-3.728	2.075
same gender & different class ($\theta(sg, dc)$)	4.313	4.459	5.031	4.307	-0.871	-1.172	-0.750
different gender & same class ($\theta(dg, sc)$)	4.975	5.119	5.281	4.136	-1.289	-4.814	2.308
same gender & same class ($\theta(sg, sc)$)	2.238	2.425	2.750	1.944	-0.994	-1.288	-0.494
Panel B: four coed schools with separated-gender classrooms							
different gender & different classroom ($\theta(dg, dc)$)	6.156	5.769	5.856	5.716	0.247	-3.322	3.346
same gender & different classroom ($\theta(sg, dc)$)	4.275	4.266	4.036	3.756	-0.270	-0.420	-0.167
same classroom ($\theta(sc)$)	2.631	2.442	2.452	1.938	-0.325	-0.375	-0.238

Notes: In Panel A, there were three co-educational schools with mixed-gender classrooms in each group of treatment and control schools. As such, there are four homophily types depending on whether a pair of students are (1) the same gender and (2) classmates. $\theta(dg, \cdot)$ represents the parameter for the pairs of students of different gender, and $\theta(sg, \cdot)$ denotes the parameter for the pairs of students of the same gender. $\theta(\cdot, dc)$ represents the parameter for the pairs of students in different classrooms, and $\theta(\cdot, sc)$ denotes the parameter for the pairs of students in the same classroom.
In Panel B, there were two co-educational schools with mixed-gender classrooms in each group of treatment and control schools. Since all the students in a classroom have the same gender, there are only three homophily types depending on whether a pair of students are (1) the same gender and (2) classmates. $\theta(dg, \cdot)$ represents the parameter for the pairs of students of different gender, and $\theta(sg, \cdot)$ denotes the parameter for the pairs of students of the same gender. $\theta(sc)$ represents the parameter for the pairs of students in the same classroom.

between students in different classrooms matches a change of 8.7 percentage points in the dictator's offer, and it is also statistically significant. Although the reduction in linking cost is present among the students of different genders, there is no statistical significance. The results in Panel B are qualitatively the same. However, quantitatively, the magnitudes of cost reduction are smaller.

References

Algan, Y., P. Cahuc, and A. Shleifer (2013). "Teaching practices and social capital". *American Economic Journal: Applied Economics* 5.3, pp. 189–210.

Cohen, L., A. Frazzini, and C. Malloy (2008). "The small world of investing: Board connections and mutual fund returns". *Journal of Political Economy* 116.5, pp. 951–979.

Engel, C. (2011). "Dictator Games: A Meta Study". *Experimental Economics* 14.4, pp. 583–610.

Fletcher, J. M., S. L. Ross, and Y. Zhang (2020). "The consequences of friendships: Evidence on the effect of social relationships in school on academic achievement". *Journal of Urban Economics* 116, p. 103241.

Glaeser, E. L., G. A. Ponzetto, and A. Shleifer (2007). "Why does democracy need education?" *Journal of economic growth* 12.2, pp. 77–99.

Goeree, J. K. et al. (2010). "The 1/d Law of Giving". *American Economic Journal: Microeconomics* 2.1, pp. 183–203.

Gosling, S. D., P. J. Rentfrow, and W. B. S. Jr. (2003). "A Very Brief Measure of the Big-Five Personality Domains". *Journal of Research in Personality* 37.6, pp. 504–528.

Helle, L., P. Tynjälä, and E. Olkinuora (2006). "Project-Based Learning in Post-Secondary Education–Theory, Practice and Rubber Sling Shots." *Higher Education* 51.2, pp. 287–314.

Helliwell, J. and R. Putnam (2007). "Education and social capital Eastern Economic Journal".

Jackson, M. O. (2010). *Social and Economic Networks*. Princeton University Press.

Jackson, M. O. and A. Wolinsky (1996). "A Strategic Model of Social and Economic Networks". *Journal of Economic Theory* 71.1, pp. 44–74. ISSN: 0022-0531.

Milligan, K., E. Moretti, and P. Oreopoulos (2004). "Does education improve citizenship? Evidence from the United States and the United Kingdom". *Journal of public Economics* 88.9-10, pp. 1667–1695.

Narr, R. K. et al. (2019). "Close friendship strength and broader peer group desirability as differential predictors of adult mental health". *Child development* 90.1, pp. 298–313.

Powdthavee, N. (2008). "Putting a price tag on friends, relatives, and neighbours: Using surveys of life satisfaction to value social relationships". *The Journal of Socio-Economics* 37.4, pp. 1459–1480.

Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton University Press.

Appendix

A Model and Estimation

A.1 Network Formation Model

We here formally present the network formation model. Let $\mathscr{G} = \{g_1, \ldots, g_M\}$ be a collection of friendship networks, where M denotes the number of networks to consider. For instance, this collection may represent the friendship networks of treatment schools before the PBL intervention. As such, we consider four collections of friendship networks in the estimation explained later: (1) comparison schools before the intervention, (2) comparison schools after the intervention, (3) treatment schools before the intervention, and (4) treatment schools after the intervention. For each network g_k, let $N_k = \{1, \ldots, n_k\}$ be the set of nodes (i.e., students) in the network. Let ij denote the subset of N_k containing nodes i and j, and it is said to be the *link* ij. If $ij \in g_k$, then we say that there is a link from node i to node j. The interpretation is that node i identified node j as a close friend in our friendship survey. If $ij \notin g_k$, $g_k + ij$ denotes the network obtained by adding link ij to network g_k. Similarly, if $ij \in g_k$, $g_k - ij$ denotes the network obtained by deleting link ij from network g_k. Hence, a network g_k can be represented by a set of links among $|N_k|$ nodes. Note that since we consider *directed* networks, ignoring the nomination limit of ten, there are $(n_k - 1)^2$ possible links among nodes in network g_k.

Consider two nodes i and j in network g_k. We assume that the utility of each node i from node j in network g_k is represented by

$$u_{ij}(g_k) = \alpha_{ij} + \delta dist_{ij}(g_k)^{\gamma} - c_{ij}\mathbf{1}\{ij \in g_k\}, \tag{A.1}$$

where $\alpha_{ij} > 0$ is the *intrinsic value* of the nodes, $\delta > 0$ is a coefficient for the value driven through the indirect friendship in the network, $dist_{ij}(g_k) \in \{1, \ldots, \infty\}$ denotes the distance from node i to node j in network g_k, $\gamma < 0$ is the decay rate in social distance, $c_{ij} > 0$ is the *cost* to node i of having link ij, and $\mathbf{1}\{ij \in g_k\}$ is the indicator function representing whether there is link ij in network g_k or not. $dist_{ij} = \infty$ means that there is no path from node i to node j. $ij \in g_k$ implies that $u_{ij}(g_k) = \alpha_{ij} + \delta - c_{ij}$, and $ij \notin g_k$ implies that $u_{ij}(g_k) = \alpha_{ij} + \delta dist_{ij}(g_k)^{\gamma}$. Since $\gamma < 0$, if other things are equal, the value through the indirect friendship increases in γ.

We now explain the homophilic random costs. We consider two-dimensional types of nodes depending on their gender and classroom. In particular, for each pair of nodes i and j, we assume that there are four possible types for each pair of nodes: (1) node i and node j have neither the same gender nor classmates, (2) node i and node j have the same gender but not classmates, (3) node i and node j have classmates but not the same gender, and (4) node i and node j have both the same gender and classmates. For the sake of notation simplicity, we let θ_{00}, θ_{10}, θ_{01}, and θ_{11} be the corresponding types of node pairs, respectively. In this regard, the first subscript represents whether two given nodes have the same gender (i.e., it becomes 1 if the nodes are the same gender), and the second subscript represents whether the nodes are classmates (i.e., it takes the value of 1 if the nodes are in the same classroom).

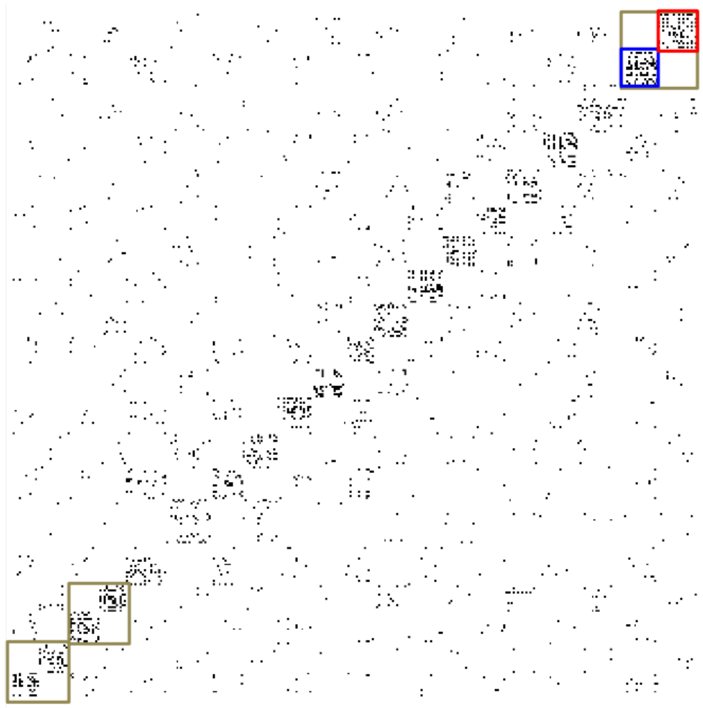
We consider random link costs: when the type of nodes i and j is θ_{st}, then the cost in equation (A.1) is drawn from a distribution parameterized by $f(\cdot;\theta_{st})$, where $\theta_{st} \in \{\theta_{00}, \theta_{10}, \theta_{01}, \theta_{11}\}$. Note that the first two terms in equation (A.1) do not rely on types of nodes i and j. Therefore, for given other model parameters $\{\alpha_{ij}\}_{i,j\in N_k}$, δ, and γ, the cost distribution parameter $\theta = (\theta_{00}, \theta_{10}, \theta_{01}, \theta_{11})$ determines the utilities among the nodes.

The above network formation model is estimated using the method of simulated moments. We choose parameter θ to minimize a quadratic objective function that measures the distance between moments observed in the data and the moments predicted by the model for a selected parameter.

Remark 1 The two-dimensional homophily type vector is well-motivated by adjacency matrices generated by the data. Figure 4 is an illustration of the adjacency matrix of a co-educational school with mixed-gender classrooms in our data. Students are ordered according to their student IDs, from the bottom to the top. The student IDs are ordered according to their classroom number and gender. In the same manner, students are ordered according to their IDs, from left to right. The figure represents the manner in which the students corresponding to the vertical axis identified the students corresponding to the horizontal axis. Specifically, each black dot denotes that the vertical student identified the horizontal student in the friendship survey.

There are several notable features in the figure. First, we observe relatively denser black dots on the diagonal entries. These black dots on the diagonal entries are clustered. Specifically, there are 10 classrooms in this school, where each classroom is again segregated into two groups by gender. For instance, at the bottom left corner of the figure, two colored squares represent two classrooms in the school. At the right top corner of the figure, a classroom has two red and blue squares which presents a set of students have a single-gender. Black dots are highly clustered within the same-gender students, especially in the same classroom. As such, we consider a two-dimensional type vector.

Figure 4: Illustration of the Adjacency Matrix of a Co-Educational School with Mixed-Gender Classrooms



Remark 2 Among six schools in the treatment schools, four schools are coed with mixed-gender schools. However, among the other three schools, two schools are coed with separated-gender classrooms, and one school is a male-only school. This diversity of school types is symmetric to the comparison schools.

For the above reason, for the network formation model of co-educational schools with separated-gender, we consider three-dimensional type vector, $\theta = (\theta_{00}, \theta_{01}, \theta'_1)$, where θ_{00} is for the pairs of nodes with a different gender in different classrooms, θ_{10} is for the pairs of nodes with the same gender in the same classroom, and θ'_1 is for the pairs of nodes in the same classroom regardless of their genders. For the network formation model of male-only schools, we consider a two-dimensional type vector, $\theta = (\theta'_0, \theta'_1)$, where θ'_0 represents the pairs of nodes in different classrooms, and θ'_1 represents the pairs of nodes in the same classroom.

A.2 Structural Estimation

Let $\mathscr{G} = \{g_1, \ldots, g_M\}$ be a collection of friendship networks. There are four different collections: (1) networks of the comparison schools before intervention, (2) networks of the comparison schools after intervention, (3) networks of the treatment schools before intervention, (4) networks of the treatment schools after the intervention.

Other model parameters. Recall that from our parametric estimation of the $1/d$ law of giving, we obtain estimates for δ and γ from the data consisting of subsamples corresponding to the collection of networks. For instance, for the networks of the control schools before the intervention, we can calculate the estimates for δ and γ from the subjects' giving behavior in these schools. We consider the following regression equation:

$$Share_{ijs} = \beta_0 + \delta dist_{ijs}^{\gamma} + X_{ijs} + \gamma_s + \varepsilon_{ijs},, \tag{A.2}$$

where X_{ijs} collects information regarding demographic, SES, and homophily between subject i (dictator) and j (recipient), γ_s refers to school fixed effects (i.e., a list of dummy variables indicating each school in the subsample dataset), and ε_{ijs} is an error term. The results are summarized in Table 7, and we use these values to replace the corresponding parameters in equation (A.1).

As we explain soon, the intrinsic value α_{ij} does not matter when we consider calculations of unilateral pair-wise marginal utilities of link formation. Therefore, it suffices to use the estimates for δ and γ and ignore all the other covariates from the $1/d$ law of giving estimation because they do not rely on the network distance. We will be more explicit about this when we present the marginal utility calculation.

Simulation. The link formation cost c_{ij} depends on whether node i and node j have the same gender and are classmates, as in the model. With regard to the family of parameterized distributions, we consider the beta distributions. Specifically, when a pair of nodes corresponds to type θ_{st}, we assume that c_{ij} is drawn from the beta distribution $f(\cdot; \theta_{st}, 10 - \theta_{st})$. The sum of two beta distribution parameters is set to be 10 for all the distributions. This choice is arbitrary, but it makes the interpretation of results intuitive. Specifically, each distribution is bell-shaped whenever the parameter is contained in $(1, 9)$. In addition, $\frac{\theta_{st}}{10} \in (0, 1)$ represents the expected link formation cost of pair type θ_{st}.

In the estimation, we estimate $\theta = (\theta_{00}, \theta_{10}, \theta_{01}, \theta_{11}) \in \Theta = (0, 10)^4$ from the data from six co-

Table 7: Estimates for the Simulation under $1/d$ Law of Giving Specification

	Comparison		Treatment	
	Before	After	Before	After
Parameter	(1)	(2)	(3)	(4)
Panel A: six coed schools with mixed-gender classrooms				
indirect value coefficient δ	0.252***	0.247***	0.259***	0.210***
	(0.034)	(0.034)	(0.028)	(0.030)
decay rate in social distance γ	-1.067***	-1.514***	-1.328***	-0.876***
	(0.279)	(0.339)	(0.280)	(0.210)
Full Controls	Yes	Yes	Yes	Yes
Observations	2,227	2,237	2,791	2,810
R-squared	0.128	0.173	0.168	0.084
Panel B: four coed schools with separated-gender classrooms				
indirect value coefficient δ	0.211***	0.164***	0.173***	0.229***
	(0.040)	(0.031)	(0.034)	(0.037)
decay rate in social distance γ	-1.070***	-1.765***	-1.423***	-0.540***
	(0.365)	(0.749)	(0.559)	(0.144)
Full Controls	Yes	Yes	Yes	Yes
Observations	2,227	2,237	2,791	2,810
R-squared	0.128	0.173	0.168	0.084
Panel C: two male-only schools				
indirect value coefficient δ	0.208**	0.247***	0.285***	0.365***
	(0.080)	(0.058)	(0.052)	(0.058)
decay rate in social distance γ	-0.720**	-0.884***	-0.743***	-0.447***
	(0.353)	(0.387)	(0.223)	(0.125)
Full Controls	Yes	Yes	Yes	Yes
Observations	2,227	2,237	2,791	2,810
R-squared	0.128	0.173	0.168	0.084

Notes: Full controls include demographic and SES controls, cognitive and non-cognitive controls, and homophily controls. The observations of infinity social distance are excluded. Robust standard errors in parentheses are clustered at individual ID. School fixed effects are included.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

educational schools with mixed-gender classrooms. We select parameter vector θ so that the moment vector, $m_{sim}(\theta)$, consisting of the following moments predicted by the model best match the four actual moments vector m_{emp}, observed in the data:

(1) p_{00}: the ratio of links between nodes are neither the same gender nor classmates;

(2) p_{10}: the ratio of links between nodes that have the same gender but not classmates;

(3) p_{01}: the ratio of links between nodes that have the same gender but not classmates; and

(4) p_{11}: the ratio of links between nodes that have both the same gender and classmates.

For other datasets consisting of a different set of schools, we consider the corresponding moments previously explained.

For each possible choice of θ, we simulate the unilateral perturbation of networks. Fix a network $g_k \in \mathscr{G}$. Then, for each node i in the network, we calculate the following probabilistic perturbation of links from i simultaneously: for each node j in the network,

$$\text{if } ij \in g_k \text{ and } u_{ij}(g_k - ij) > u_{ij}(g_k), \text{ then delete link } ij \text{ from network } g_k,$$

if $ij \notin g_k$ and $u_{ij}(g_k + ij) > u_{ij}(g_k)$, then add link ij to network g_k.

Note that since the intrinsic value α_{ij} in equation (A.1) appears in both the left and right-hand sides of the above inequalities, it does not play any role in the above calculations. Let $\widetilde{g}_k$ be the resulting network after perturbation and calculate the four moments for each node. To reduce the variation generated by the randomness of the costs, we repeat the perturbation 100 times and take the average over all the nodes in all the networks and repetition.

The above process yields a vector of average simulated moments, and we denote it by $m_{sim}(\theta)$ as a function of θ. We let m_{emp} be the vector of empirical moments for the corresponding collection of networks. Then, we find $\widehat{\theta}$, which minimizes a quadratic criterion function as

$$\widehat{\theta} = \underset{\theta \in \overline{\Theta}}{\operatorname{argmin}} ||m_{sim}(\theta) - m_{emp}||^2.$$

When we numerically find $\widehat{\theta}$, we consider a compact set $\overline{\Theta} = [\varepsilon, 10 - \varepsilon]^4$ as the domain of $m_{sim}(\theta)$, where $\varepsilon = 2.2204 \times 10^{-16}$ is the smallest number in `MATLAB`. We search over the entire set of possible parameters by using `patternsearch`, a global optimization tool in `MATLAB`.[12]

We aim to measure treatment effects on the link formation costs between nodes. To this end, recall that there are four combinations of the network collections. Thus, for each parameter θ_{st}, we find its the difference-in-differences parameter $\widehat{\theta}^{did}_{st}$, which is defined as

$$\widehat{\theta}^{did}_{st} = (\theta_{st}(\text{treatment, after}) - \theta_{st}(\text{treatment, before})) - (\theta_{st}(\text{control, after}) - \theta_{st}(\text{control, before})),$$

where $\widehat{\theta}_{st}(\text{treatment, after})$ denotes the estimate from the networks in the treatment group after the intervention, and other parameters are similarly defined.

To establish the statistical inference of the causal impact of the PBL intervention on homophilic costs, we find the bootstrapping estimates of each parameter θ_{st}. We generate 500 empirical bootstrap samples drawn from networks. Specifically, for each school network g_k of size $|N_k|$, we draw a set random samples of $|N_k|$ nodes with replacement. Then, for each chosen node, we calculate perturbation of link formations from the node to all other nodes in the same network g_k. We repeat the perturbation 20 times. Consequently, we obtain a vector of average moments $m^b_{sim}(\theta)$. Let m^b_{emp} be the vector of

[12] We refer to an online documentation https://kr.mathworks.com/help/gads/patternsearch.html as a reference. As options for numerical calculations, we set `MeshTolerance` to be 10^{-4}, which specifies the minimum tolerance for mesh size in the optimization algorithm.

empirical moments for the chosen nodes. Then, we find a bootstrap estimator $\widehat{\theta}^b$ defined as

$$\widehat{\theta}^b = \underset{\theta \in \overline{\Theta}}{\operatorname{argmin}} ||m^b_{sim}(\theta) - m^b_{emp}||^2.$$

As a result, we obtain 500 bootstrapping estimates $\widehat{\theta}^b$. Finally, by using the bootstrapping estimates, we calculate the bootstrapping estimate of each $\widehat{\theta}^{did}_{st}$ and calculate its 95% confidence interval.

B Additional Tables and Figures

Table 8: Impacts of PBL on Homophilic Network Formation

Variables	$P(sc)$ (1)	$P(nomination)$ (2)
PBL_{st}	0.015*	0.099***
	(0.002)	(0.000)
Controls	Yes	Yes
Observations	661	668
R-squared	0.038	0.098

Notes: $P(sc)$ represents the proportion of friends in the same classroom. $P(nomination)$ is the ratio of outdegree to 10, the upper limit of the number of friendship nominations. Control variables include students' gender, height, and weight. Other controls used in previous analyses, such as demographic and SES information, are not included because they are available only for the students who participated in the dictator game experiment. School fixed effects are included. Robust standard errors in parentheses are clustered at the school level.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 9: Parameter Estimates and Bootstrap Confidence Intervals for Two Male Schools

	Comparison		Treatment		DID	Lower CI	Upper CI
	Before	After	Before	After			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
different classroom ($\theta(dc)$)	3.906	4.631	4.006	3.875	-0.856	-0.975	-0.706
same classroom ($\theta(sc)$)	1.936	2.355	2.567	2.213	-0.773	-0.931	-0.577

Notes: There was only one male-only school in each group of treatment and control schools. As such, there are only two homophily types depending on whether a pair of students are classmates. $\theta(dc)$ represents the parameter for the pairs of students in the same classroom, and $\theta(sc)$ denotes the parameter for the pairs of students in different classrooms.