Peer Effects in Active Learning*

Vinicius Lima[†] Vladimir Ponczek[‡] Priscilla Tavares[‡]

May 2026

Abstract

This paper studies peer effects in higher education using an active learning environment where peer interaction is central to instruction. We exploit random assignment of students to tutorial groups, which generates exogenous variation in two dimensions: the ability composition of each group and the number of groups that pairs of students share. We show that both dimensions predict social proximity among peers — pairs who shared more groups and pairs closer in predetermined ability are more likely to interact, a pattern we term social proximity. We then compare outcomes under different social proximity regimes induced by this variation. Among low-ability students, a shift from the 25th to the 75th percentile in the number of peers likely to become socially proximate raises exam scores by 6.2% per discipline, and increases the probability of an outstanding participation grade by 31 percentage points, which we interpret as evidence of increased effort. There are no detectable effects of social proximity on high-ability students. The conventional ability peer effects regression, which does not distinguish between peers likely and unlikely to become socially proximate, yields a near-zero aggregate estimate. Our results show that peer effects in higher education operate through distinct channels depending on students' ability level, and that policies targeting low-ability students must account for the social proximity that a given group composition actually generates.

JEL Codes: D62, D85, I21, I23, J24

*We are grateful to Joelson Sampaio and the Sao Paulo School of Economics for providing the data and for all the support. We thank Stephen Ross for comments on this paper. All remaining errors are on our own.

[†]Mackenzie Business School. E-mail: vinicius.lima@mackenzie.br

[‡]São Paulo School of Economics – FGV

IRB approval from the Research Ethics Committee led by Fundação Getulio Vargas (FGV) under the project P.437.2023.

1 Introduction

Policies targeting the performance of low-skilled students often rely on group ability composition as the main lever to foster spillovers from high-skilled peers. However, since peer dynamics are a key mechanism to generate these spillovers, changing a group's ability mix entails a trade-off. Adding high-skilled peers to a group increases the supply of high-performing learning partners for low-ability students, but it is documented that distance in ability makes students less likely to actually interact (McPherson et al., 2001; Currarini et al., 2009; Smirnov and Thurner, 2017). If peer teaching is the relevant mechanism justifying such a policy, then peer interaction is essential (Kimbrough et al., 2022) and, as Carrell et al. (2013) demonstrate, the policy might backfire if homophily in interaction prevails. Thus, in this paper we ask whether social proximity among peers generates the interaction through which peer effects on performance operate – and whether this mechanism differs across students of different ability levels.

Studies of peer effects on achievement in college have shown that a common channel is a transfer of general skills rather than specific subject knowledge — occurring, for instance, through interaction outside class (Griffith and Rask, 2014) or through the development of good learning habits fostered by peers with favorable personality traits (Golsteyn et al., 2021). Although some studies suggest that some mechanisms of peer effects improve the learning technology without necessarily raising effort, prior evidence does suggest that peer effort can affect one's own effort (Mehta et al., 2019). We provide evidence that peers who are more likely to be socially proximate improve low-ability students' exam scores and increase the probability that they receive a distinction associated with outstanding performance in group work – a pattern consistent with higher effort during group work. Because effort is likely complementary across students in group work, this mechanism has the potential to generate aggregate effects that exceed the sum of individual gains.

We study first-semester economics undergraduates at an institution whose active learning methodology places peer interaction at the core of instruction — students meet weekly in small, task-oriented groups, their individual contribution during these sessions is formally evaluated by a tutor as a participation grade, and their final score combines this grade with performance in written exams. Our identification strategy exploits a random assignment algorithm that allocated students to discipline-specific tutorial groups, generating exogenous variation in two dimensions: the ability mix across groups and the frequency with which pairs of students meet for group work. Variation in ability mix and meeting frequency shapes the likelihood that any two students become socially proximate — a concept we proxy by the mutual desire to meet again for future group work, elicited in a survey applied after the first academic quarter. Our main specifications compare students with more peers likely to be socially proximate against those with fewer, exploiting within-student variation across disciplines to control for unobserved heterogeneity. Because the exogenous variation we exploit — differences in meeting frequency and ability mix across disciplines — is itself generated by the assignment across multiple disciplines, the within-student estimator differences out any general performance benefit that a denser network of socially proximate peers produces uniformly across a student's courses. The estimated coefficient therefore identifies the average of a subject-specific average premium of social proximity: the additional performance gain in a discipline attributable to having one more peer likely to be socially proximate present in that discipline, net of any general benefits the relationship has already generated across the semester.

Our main results show that a conventional ability peer effects regression yields an estimate indistinguishable from zero. Specifications that account for social proximity reveal a richer picture. For low-ability students, having more peers likely to be socially proximate increases average exam scores by 2.8% — a average subject-specific premium of having a socially proximate peer. Among these students, a shift from the 25th to the 75th percentile in the number of peers likely to become socially proximate raises exam scores by 6.2% per discipline. We interpret this as a lower bound on the total effect of social proximity, to the extent that denser peer networks may also generate spillovers across a student's full course load — benefits that are absorbed by the student fixed effect. Also, the low-ability students are 31% more likely to receive a distinction associated with outstanding performance in group work, which is consistent with higher effort during tutorial sessions. We rule out a simple redistribution of distinctions within groups — whereby some students gain at the direct expense of others — as an explanation for this result: the total count of students with outstanding participation grades in a group increases with its network density There are no detectable effects on the performance or in the probability of outstanding performance of high-ability students, suggesting that peer effects operate through distinct mechanisms in different groups of students.

Our paper contributes to the literature on peer effects in college by providing evidence that social proximity is the mechanism through which spillovers emerge for low-ability students. Moreover, we show that peer effects operate through a distinct channel for high-ability students: it is the average ability of peers that shapes outstanding performance for them, suggesting that a single group composition policy is unlikely to activate both channels simultaneously. Our unique knowledge of the group formation process allows us to circumvent a key limitation of approaches that infer social proximity from records of time spent together (Martin et al., 2020; Presler, 2022), which may suffer from confounding due to unobserved factors that simultaneously drive social closeness and academic outcomes. While Coveney and Oosterveen (2021) employ a closely related strategy — exploiting random assignment to subgroups within tutorial groups to identify the role of social interaction in peer effects — their setting differs from ours in a key dimension: proximity is generated through informal social meetings unrelated to course material, and their evidence suggests effects operate through substitution of lecture attendance with self-study. We show that proximity arising from academic group work generates a distinct mechanism: increased effort during the sessions themselves.

In the rest of the paper, we present the institutional setting and data in Section 2, followed by a simple model of peer effects in Section 3. Section 4 describes the empirical strategy. Section 5 presents the main results, and Section 6 discusses robustness exercises. Section 7 concludes.

2 Organizational Framework

2.1 Background

The Sao Paulo School of Economics is a private higher education institution established in Brazil in 2003. In 2019, the annual course fee was 65,000 BRL, roughly 3.7 times the estimated Brazilian per capita income for that year. Students in our sample, therefore, rank at the top of the income distribution in Brazil. Between 2003 and 2016, the school used a highly selective admission exam taken by about 1,500 applicants to choose up to 60 students for its undergraduate program in economics, the only one offered by the school. The number of students admitted each year gradually increased to 120 between 2017 and 2021.

Since 2013, all admitted students have completed their coursework within the problem-based learning (PBL) framework. This active learning environment at the Sao Paulo School of Economics is structured in weekly tutorial sessions. There were, on average, 13 students per group, and each group had a tutor – either a professor or a Ph.D. student. Students attended two or three weekly tutorial sessions, contingent on the specific course (Math, Finance, etc, for instance). In sessions of the same discipline, the group of students remained constant throughout the academic semester, while students faced distinct peer groups across different disciplines. Over the period examined in the paper, the distribution of traditional lectures and tutorial sessions varied depending on the specific course. Nevertheless, tutorial sessions comprised at least 65% of the coursework in any course.

In each tutorial session, the students' main goal is solving a problem by applying concepts and techniques they learn through self or group study outside the class. The approach to a specific problem has three phases. First, students get to know the problem and identify the learning goals. Then, they study outside the class using the bibliographic references and go to the next meeting, where they effectively solve the problem. The tutor's main job is to ensure that all students put effort in the work developed during the sessions. Students are encouraged to participate in the discussions through questions directly posed by the tutor or through individual feedback, publicized at the end of the meeting. The tutor's end-of-session feedback is a mandatory task in which grades between zero and one based on individual performance are assigned to each student. A grade below one denotes that performance fell short of the standard that was expected for the session. Also, it is possible that a student receives a grade slightly above one to denote an outstanding performance. Absence in the session implies a participation grade equal to zero.

Summary In the active learning methodology, peer interaction is a meaningful learning mechanism, and in the environment that we analyze, this is reinforced by two characteristics. First, all students have real-world incentives to interact with peers because the evaluation of their effort in group work is part of their final assessment in each discipline they take, and correct incentives are important to foster peer interaction (Li et al., 2014). Besides, we observe students who meet in relatively small task-oriented groups. This proximity should benefit peer interaction (Hong and Lee, 2017; Lu and Anderson, 2015) and provide opportunities for collaboration between peers of different ability levels (Brady et al., 2017). These features make peers an important input of a student's achievement production function in our setting.

2.2 Students Assignment Method

We implemented the allocation of new students from the 2018 and 2019 cohorts in their first semester at the school. During this semester, they took six mandatory courses, and for every discipline, we assigned each student to a tutorial group with 12 students on average.

The algorithm to allocate students into tutorial groups ensured random variation in two dimensions. First, considering the number of low- and high-ability students in each group, the algorithm created considerable variation across groups. We classified students by their predetermined ability measured by their ranking in the admission exam, which was the information we had available at the time of the allocation. Besides, since students take six disciplines, some shared more than one group. Thus, some pairs of students met weekly in only one group while others, by chance, met in another group.

Now, we explain the relevant aspects of the assignment mechanism with a simple example. Suppose we have to allocate 18 students in three groups of a given discipline, and the admission ranking determined nine low-ability students and nine high-ability students. Then, we followed these steps:

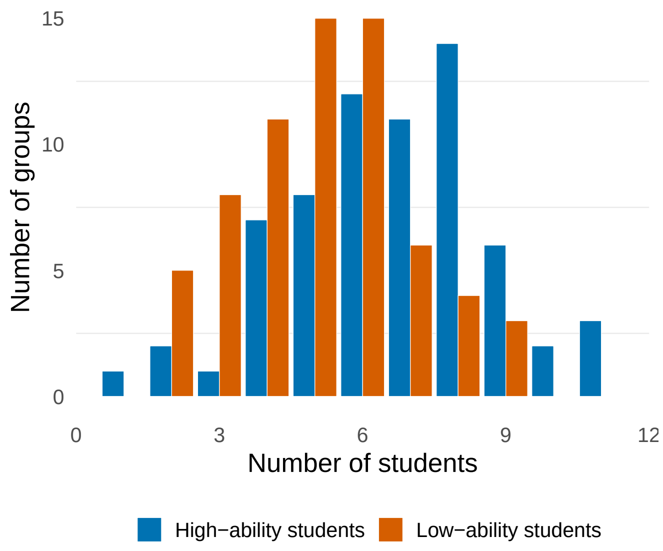
- Step 1) The algorithm randomly chooses how many students of each ability to place in each group. Example: Group A will have 5 low- and 1 high-ability students, group B will have 1 low- and 5 high-ability students and group C will have 3 low- and 3 high-ability students. We run this lottery without replacement (conditional on type) to minimize the chance that two groups have the same composition, and then we increase variation across groups. This step is constrained by group size and total students by ability level.

- Step 2) Given the composition of groups defined in step 1, this step defines, at random, which students will be in each group. That is, if there is a total of 9 low-ability students and the previous step defined that a group must have 5 of them, this step sets the identity of these 5 students. Then, among the 4 remaining low-ability students, the algorithm draws a subset for the next group, and so on.

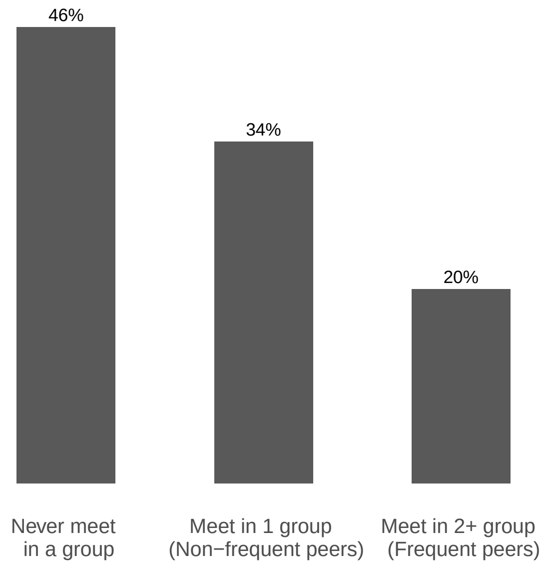
Random variation in group composition In step 1, the algorithm generated what we call *random variation in group composition* throughout the paper. It means random variation in the number of low- and high-ability students across groups. Besides, step 2 ensured that conditional on the ability level, being in a group with many or a few high-ability peers is a random outcome. The result of such random variation is in figure (1a), where we present the distribution of the number of students by ability level considering the 67 groups created. Although the algorithm used the admission ranking to allocate students, we present group composition with students classified by their predetermined math ability. This is the ability measure we will consider in the paper as it is the strongest predictor of GPA in previous cohorts. Apart from low- and high-ability students, there were few students with no ability classification and students redoing the discipline (see descriptive statistics). These students were randomly allocated after the low- and high-ability ones.

Random variation in the frequency of meetings Consider three students i, j, k randomly selected to be in the same group of some discipline. In a second discipline, a different draw might place the pair ij in the same group while k goes to another group. This implied that ij met twice for group work, while ik jk met only once. This is what we call *random variation in the frequency of meetings*. There were 4,729 possible pairs of students in the cohorts we analyzed. However, in the actual allocation, 46% of those potential pairs never happen to meet for group work, 35% meet in only one group, and 20% meet in more than one group. Thus, among all pairs allocated to some group, more than one-third meet at least once again in a different group.

After doing the above procedure, we provided the school staff with a list indicating the groups of each student in each discipline. After the school received this information, but before the beginning of the classes, some students withdrew from the enrollment before knowing their allocation (usually to accept offers from other institutions). Students admitted to replace the leaving students were assigned to available slots following the same original assignment rule.



(a) Group composition



(b) Frequency of meetings

Figure 1: Summary of the assigment mechanism – Figure (1a) displays the distribution of groups by the number of students classified by their math ability level, which is the strongest predictor of GPA in previous cohorts (correlation 0.52). Figure (1b) displays the distribution of pairs of students by number of meetings for group work. The total number of pairs in the sample is 4,729.

Then, considering the pairs student-discipline provided to the school staff, the compliance rate with the random allocation was 98% among those who effectively started the course. Finally, the school staff independently assigned tutors to each group in advance of the students' assignment. We show in subsection 3.1 that students' ability is unrelated to the tutors' characteristics.

3 Data

The school provided data on students' academic performance, their scores in the admission exam, and the files recording the students' allocation in each discipline (based on the assignment rule). From a survey conducted among the students, provided by the school, we gather data on the peers they indicated as a choice for groups of subsequent courses. In this section, we describe how we construct the variables based these data sets, present descriptive statistics, and perform an evaluation of the random assignment procedure.

Peer reporting During the period we analyzed, the school applied a survey to get information on students' perceptions about the PBL method. The survey was applied some weeks after the end of the first academic quarter. About 2/3 of the students answered the questionnaire, and participation was unrelated to ability.[1] At the end of the survey, students indicated up to seven peers they wished to meet again in future groups, independently of sharing some group at the

[1] See Table A1 in the appendix.

time.[2] We use this information to construct the variable *Match* at the dyad level. Considering the sample of students who answered the questionnaire, *Match* is equal to 1 when both i reports j and j reports i, and 0 otherwise.

We assume that by revealing the desire to meet a specific peer in future tutorial groups, students recognize that the specific peer contributes to their academic performance. Besides, when this choice is reciprocal, we assume that peer interaction on academic matters is likely to explain part of the social proximity between i and j. Thus, we use *Match* as a proxy to indicate social proximity between the pair ij.

Academic performance During their initial semester at the school, students must complete six courses. Three of these courses extend throughout the entire semester, while two conclude at the end of the first academic quarter. The sixth course starts in the second academic quarter. Within the period of one week between the first and second quarters, students take exams for the five courses of the first quarter, and at the end of the second quarter, they take exams for the four courses of this period.

We use three performance measures: *Exam*, *Participation*, and *Participation* > 1. *Exam* is the student's score on each discipline's exam, standardized by discipline and year. For the three courses that have two exams we define *Exam* as the average between the two exams. *Participation* is simply the average of participation grades received by the student in each tutorial session of a discipline. The variable *Participation* contains important qualitative information. A student can only achieve a participation grade above 1 if she has outstanding performance in the group work. Thus, we define *Participation* > 1 as an indicator for the case where the student's participation grade in a given discipline signals that she received this distinction at least once.

Peers variables The variables *low-ability peers* and *high-ability peers* count the number of peers in the group by each level of ability. *Frequent* and *non-frequent peers* count how many peers in a student's group appear in at least some other group of that same student. We also count peers classified according to the possible combinations of the two categories, such as *frequent high-ability peers*, *non-frequent high-ability peers*, and so on. There are three types of students we count separately to include in the regression as a control labeled as *other peers*: those doing a discipline for a second time, a few that did not use the school's admission exam and cannot be classified by the same ability measure, and students from other departments of the school.

[2] The constraint of a maximum of seven peers was due to space in the questionnaire.

3.1 Descriptive Statistics and Balance Checks

Table 1 provides descriptive statistics of the variables used in the analysis. We observe data from 132 students allocated in 67 groups of 6 disciplines per year. The average group has 6.5 high-ability and 5.1 low-ability students, comprising an average group size of 12.5 students. The average numbers of frequent low- and high-ability peers in a group are 2.7 and 3.4. The average performance in the exam is 7.1 (the school's passing grade is 6), and the participation grade is skewed towards 1 with low dispersion. There is much more variation in the indicator of outstanding performance as evidenced by 12% of participation grades above 1.

Table 1: Descriptive Statistics

	Mean	SD	Min	Max	N
Panel A. Student's ability					
Admission score	5.28	0.55	4.18	6.71	132
Math score	5.56	0.97	3.72	7.90	132
Language score	5.59	0.78	3.21	6.77	132
Panel B. Group variables					
No. of high-ability students	6.54	2.14	1.00	11.00	67
No. of low-ability students	5.13	1.78	2.00	9.00	67
Other types of students	0.72	0.87	0.00	3.00	67
Group size	12.48	2.31	8.00	16.00	67
Teacher's years of experience	7.34	7.15	0.00	21.00	67
Teacher is a woman	0.21	0.41	0.00	1.00	67
Panel C. Student x Discipline variables					
No. of frequent high-ability peers	3.41	1.71	0.00	10.00	782
No. of frequent low-ability peers	2.71	1.39	0.00	9.00	782
Final Score in the Discipline	6.53	2.04	0.00	10.00	782
Performance in the 1st Exam	7.12	1.83	0.00	10.00	652
Performance in the 2nd Exam	5.80	2.32	0.00	10.00	518
Participation grade	0.99	0.05	0.34	1.10	782
Participation > 1	0.12	0.33	0.00	1.00	782

Notes: Student's ability measures are their scores in the admission exam and ranges from 0 to 10. High-ability students are those with math score in the admission exam above the median, and low-ability ones those below it. *Frequent peers* are those peers that a student meets in more than one tutorial group. Final score is the participation grade times performance in the exams (an average of the exams when there are more two).

To obtain an unbiased estimate of peer effects, it is important that the students' average ability is uncorrelated with the number of low- and high-ability students as well as the number of frequent and non-frequent peers. We test this hypothesis using the variation of peer related variables across groups of the same discipline. Specifically, we regress the students' admission score on the peer variables, conditional on randomization controls – year, discipline, and whether the student is classified as either low- or high-ability. Table 2 shows results that do not reject this hypothesis. Column 1 displays point estimates for the coefficients on the number of low- and high-ability peers which are negligible. In column 2, low- and high-ability peers are decomposed into frequent and non-frequent peers, and again, coefficients are very close to zero, which is also true for the coefficient on *Other peers* in both cases.

However, one potential concern in our context is that we must assign students to a given discipline by drawing peers *wihtout replacement* from a relatively small pool of peers in each year: we observed two cohorts of 51 and 82 students. This could lead to some bias even with random assignment since the probability of having a high-ability peer might differ to low- and high-ability students (Guryan et al., 2009). Indeed, in the appendix we show that the distribution of the estimates displayed in Table 2 is biased away from zero.[3] We address this problem by leveraging the within-student rather than between-group variation of peer related variables. Different from what we just described, for a given student, peers across groups of different disciplines are drawn *with replacement*.[4]

To conclude, since teachers are allocated by the school staff in advance and independently of the students allocation into groups, their characteristics are unrelated to students' average ability across group as evidenced by column 3 of Table 2.

4 Empirical Strategy

4.1 Identification

Here, we present the empirical approach that we adopt to (i) test the hypothesis that two students meeting more frequently are more likely to be socially proximate and (ii) estimate peer effects on performance. Following the discussion in 3.1, leveraging between-group variation of the students performance measure and peer related variable would produce biased estimates. Thus, our main strategy is to estimate peer effects explores the within-student variation of the variables. Each regression is weighted by the number of meetings per discipline during the semester. After that, we discuss the interpretation of the coefficients. To conclude, we present

[3]We do so by constructing an empirical distribution for each coefficient displayed in Table 2 using 10,000 placebo samples generated from the same assignment mechanism. See Tables A2 and A3, and Figures F1 and F2 in the appendix.

[4]In the Appendix Table A4, we give further evidence that using the within-variation solves the problem by showing that running our main specifications while controlling for the expected values of the peer related variables (Guryan et al., 2009; Borusyak and Hull, 2023) does not alter the results.

Table 2: Randomization check – Regression of Admission Score on Group Composition

	Admission score		
	(1)	(2)	(3)
High-ability peers	-0.006		
	(0.008)		
Low-ability peers	0.001		
	(0.009)		
Frequent high-ability peers		0.000	
		(0.008)	
Frequent low-ability peers		-0.003	
		(0.009)	
Non-frequent low-ability peers		-0.014	
		(0.009)	
Non-frequent high-ability peers		0.004	
		(0.009)	
Other peers	0.005	0.004	
	(0.011)	(0.011)	
Teacher is a woman = 1			-0.039
			(0.025)
Years of experience			0.000
			(0.001)
Num.Obs.	782	782	782
Students	132	132	132
Groups	67	67	67

Notes: The dependent variable is the raw score used in the admission exam and ranges from 0 to 10. This was the ability measure used to classify students in the group assignment procedure. Peers variables count, for each student, the number of peers in each group categorized by their math ability and by the frequency of meetings with the student. Robust standard errors in parenthesis.

the inference procedure we adopted.

Understanding social proximity To test the hypothesis that peers meeting in more than one group are more likely to be socially proximate we perform a regression at the dyad level. Consider the pair n composed by student i and peer j. Using the sample restricted to students that answered the school's questionnaire, we estimate the equation

$$\text{Match}_n = \alpha + \beta \text{nonfrequent\_peer}_n + \gamma \text{frequent\_peer}_n + \delta \text{math\_distance}_n + \varepsilon_n \qquad (1)$$

where Match_n is a variable indicating a match in peer reporting: i reported the desire of having j in her groups of subsequent courses *and* j reported i for the same reason. This is our proxy variable for i and j being socially proximate. The pair n not necessarily met in some group – the question was unconditional on meeting – thus, never_meet_n indicates that the pair n never met for group work, $\text{nonfrequent\_peer}_n$ indicates they met in only one group, and frequent_peer_n indicates they met in more than one group. The variable math_distance_n is the absolute distance between the math ability measures of i and j.

In equation (1), the constant α represents the estimated probability of a match between pairs who never met in any tutorial group. The coefficients β and γ capture the differences in the probability of a match—relative to pairs who never met—for those who met in one tutorial group and for those who met in more than one tutorial group, respectively. Controlling for ability distance through math_distance_n allows us to assess whether the effect of frequent meetings on social proximity operates independently of peers' relative ability levels — that is, whether repeated meetings capture a dimension of social proximity beyond homophily in ability. We report results with and without this control.

Ability peer effects A basic regression to estimate peer effects on performance using within-student variation of a peer-related variables is

$$y_{i,d} = \beta \text{low\_peers}_{g(i,d)} + \mathbf{z}'_{g(i,d)}\pi + \eta_i + \theta_{c(i),d} + \gamma \text{peers\_ability}_{g(i,d)} + \varepsilon_{i,g(i,d)} \tag{2}$$

where $y_{i,d}$ is the outcome of student i in discipline d, $\text{low\_peers}_{g(i,d)}$ is the number of low-ability peers the student has in the group of that discipline, and $\mathbf{z}_{g(i,d)}$ controls for the quantity of peers either not classified by ability or redoing the course, and group size. The individual fixed-effect η_i captures the student unobserved heterogeneity and the randomization controls (student's cohort and student ability level). Also, we use a cohort-discipline fixed-effect $\theta_{c(i),d}$ to account for occasional discipline-specific changes regarding contents and methods. The variable $\text{peers\_ability}_{g(i,d)}$ is defined as the average score in math obtained in the school's admission exam by student i's peers in group $g(i, d)$. The term $\varepsilon_{i,g(i,d)}$ is an unobserved random shock of i in group $g(i, d)$. The ordinary least squares estimate of β then gives the average effect of moving i to a group with one more low-ability peer in replacement of a high-ability one, since group size and other types of students are constant.

Peer interaction The potential problem in estimating an equation like (2) is that changing the group ability distribution will likely affect peer interaction if peers' ability is important in determining who interacts with whom. However, in our setting we can distinguish subsets of peers that are more likely to be socially proximate and estimate the equation

$$y_{i,d} = \beta \text{frequent\_peers}_{g(i,d)} + \mathbf{z}'_{g(i,d)}\pi + \eta_i + \theta_{c(i),d} + \gamma \text{peers\_ability}_{g(i,d)} + \varepsilon_{i,g(i,d)} \tag{3}$$

where $\text{frequent\_peers}_{g(i,d)}$ is the number of peers that i meets in group $g(i,d)$ and also in some other group $g(i,d')$. The estimate for β gives the average effect of moving i to a group with one more frequent peer in replacement of a non-frequent peer, which increases the probability that i has a socially proximate peer in the group. Since group average ability and the frequency of meetings are independent by construction, we expect $\text{peers\_ability}_{g(i,d)}$ to have little bearing on the estimate of β. We report results with and without this control variable.

***Ex-ante* closeness** Although *frequent peers* captures an exogenous component of peer interaction, social proximity depends on both the frequency of meetings and on ability similarity — a well-documented pattern of homophily in social networks. Students are therefore most likely to be socially proximate to peers they meet repeatedly *and* who are similar in ability. We formalize this by defining *close peers* as those satisfying both conditions simultaneously. For low-ability students, this excludes non-frequent high-ability peers — the combination most likely to produce social distance given limited meeting opportunities and ability heterogeneity. For high-ability students, the symmetric argument excludes non-frequent low-ability peers. Thus, we estimate the equation

$$y_{i,d} = \beta \text{close\_peers}_{g(i,d)} + \mathbf{z}'_{g(i,d)}\pi + \eta_i + \theta_{c(i),d} + \gamma \text{peers\_ability}_{g(i,d)} + \varepsilon_{i,g(i,d)} \tag{4}$$

While in (3), replacing a non-frequent peer with a frequent one leaves the group average ability unchanged on average, this is no longer true when replacing a distant peer with a closer one. Thus, $\text{peers\_ability}_{g(i,d)}$ plays an important role as a control so that the estimate for β gives the average effect of moving i to a group with one more socially close peer in replacement of a distant one, holding group average ability constant.

Finally, for all equations, from (2) to (4), we present results for versions of these equations that display the effects separately for low- and high-ability students by interacting the peer variables with dummies for the student's own ability level.

Summary Given the definitions of the variables *frequent* and *close peers*, there are two possibilities for an effect of social proximity to arise. The first is a discipline-local channel: because a frequent peer in discipline d is by definition a student previously encountered in some other discipline d', the pair enters d with an established familiarity that facilitates interaction *within* that specific tutorial session — lowering the cost of asking questions, raising the incentive to contribute, and strengthening the complementarity in effort during group work, for instance. The second is a cross-discipline channel: the overlap across disciplines increases the total time the pair spends together over the semester, potentially generating joint study outside the classroom, informal exchanges of content, or accumulated social capital that manifests in performance gains across the full set of a student's courses. Because the within-student estimator differences out any component of performance that is constant across a student's disciplines — including any general boost that a denser network of socially proximate peers produces over the semester — the estimate for β captures only the discipline-specific component of the effect. The estimated $\hat{\beta}$ therefore identifies the subject-specific average premium of repeat pairing: the additional performance gain in discipline d attributable to having one more socially proximate peer present in that discipline, net of whatever general benefits the cross-discipline relationship has already generated.

4.2 Inference

For the dyadic regression (1) we compute standard errors following the estimator proposed by Aronow et al. (2015). For any equation derived from (2) or (4) we report standard errors clustered at both student and group levels (Cameron et al., 2011). However, since we know the data-generating process of our regressors of interest, we can report p-values calculated through a randomization inference approach. This is helpful since we do not have a large number of clusters and the correlation across error terms can be quite complex in a setting like ours.

The procedure consisted in replicating the assignment rule described in section 2.2 10,000 times to generate a set of allocations that could have been implemented instead of the actual one. Then, we use students' actual performance to run each regression in each of these placebo allocations. Thus, assuming that potential outcomes are unchanged, we compare the estimate obtained in the actual data with the distribution of estimates computed out of the placebo estimates. For each coefficient, we report the p-value associated with the test $H_0 : \beta = 0$ and $H_1 : \beta \neq 0$. The reported number is the frequency at which the placebo estimate is larger in absolute value than the actual estimate. In the appendix we provide summary statistics for each distribution used for inference (see section A.8). In the decomposition exercise reported in Section 5, we also report a Wald statistic for a joint hypothesis test constructed from the placebo distributions. Finally, to account for multiple hypothesis testing in this decomposition exercise, we apply the stepdown procedure of List et al. (2019), which controls the familywise error rate while retaining more power than classical corrections by exploiting the empirical correlation across outcomes. We define two families of hypotheses, one per ability subgroup, each comprising the coefficients on close peers across both outcomes. Test statistics are normalised by the standard deviation of their placebo distribution rather than by analytical standard errors, consistent with our randomization inference approach. Adjusted p-values are reported in appendix section A.3.

Table 3: The Effect of Frequent Meetings on Social Proximity

	Any pair	Any pair	Low-Low	Low-High	High-High
Never meet in a group (Constant)	0.023*	0.039*	−0.002	0.014	0.072*
	(0.007)	(0.010)	(0.014)	(0.012)	(0.028)
Non-frequent peers (meet in 1 group)	0.014	0.014	0.056*	0.011	−0.011
	(0.011)	(0.011)	(0.025)	(0.012)	(0.022)
Frequent peers (meet in 2+ groups)	0.071*	0.072*	0.095*	0.055*	0.079*
	(0.013)	(0.013)	(0.030)	(0.016)	(0.031)
Distance in math ability		−0.013*	0.004	−0.001	−0.019
		(0.005)	(0.034)	(0.006)	(0.021)
Dyads	1,921	1,921	336	960	625

Notes: The dependent variable is a binary variable indicating whether in a pair of students that answered the survey reported each other as a choice to be in some tutorial group in the future. Dyadic cluster-robust standard errors in parentheses calculated according to Aronow et al. (2015). $^{+}$ $p < 0.1$, * $p < 0.05$

5 Results

In this section, we present three sets of results. Firstly, we show that pairs of students meeting more frequently are more likely to be socially proximate and that social closeness also depend on the peers' ability level. Then, we show that social proximity is an important source of peer effects on the performance of low-ability students on the exams, and affects the effort they put into group work.

5.1 Understanding Social Proximity

The main takeaway from this section is that meeting a second time for group work substantially increases social proximity among peers relative to meeting only once — an effect that holds across all combinations of ability levels but is especially important for mixed-ability pairs.

Column 1 of Table 3 shows that among pairs of students who never meet in a group, there is a 2.3% chance of them reporting each other as a choice to share some group in the future. We do not reject the hypothesis that pairs who met in only one group have the same probability of a match as those who never met: the estimated difference is 1.4 percentage points and is not statistically significant at usual levels. However, pairs that interact in more than one group — the frequent peers — are four times more likely to report each other than pairs that never meet. The estimated coefficient for the difference in the probability of a match between frequent peers and those who never met is 0.071, significant at 5%, and we also reject the equality with the coefficient for non-frequent peers (p-value < 0.01).

Although the assignment rule renders meeting frequency orthogonal to peers' average ability, repeated meetings may still allow students to learn about one another's skills, potentially more so for frequently meeting pairs. Consequently, survey reports may capture perceived ability rather than true social proximity. To address this concern, in column 2 we control for the peers' distance in math ability. The estimated effect of frequent meetings on *Match* is unchanged, but the baseline probability of a match among pairs that never met for group work increases, revealing a homophily pattern: conditional on meeting frequency, peers of similar math ability levels are more likely to report each other. Because this pattern implies that the ability composition of the group shapes social proximity independently of how often peers meet, in columns 3 to 5 we examine separately the three ability-pair types relevant to our peer effects analysis.

Among pairs of low-ability students (column 3), the probability of a match is statistically zero when they never meet in a tutorial session. Among non-frequent peers, there is a 5.6% significant chance of reporting a match, and among frequent peers this chance is 70% greater, as indicated by the significant coefficient of 0.095, although we do not reject equality between these two estimates. Among mixed-ability pairs (column 4), both the never-meet and the non-frequent baseline are indistinguishable from zero, and only frequent meetings generate a meaningful and significant 5.5% chance of a match — an estimate that we reject as equal to either of the other two (p-value < 0.05). This pattern confirms that, for low- and high-ability students to become socially proximate, repeated group work is a necessary condition, not merely a facilitating one. Finally, among high-ability pairs (column 5), even students who never meet share a 7.2% baseline probability of a match — likely reflecting ties formed outside the tutorial groups. The differential for non-frequent peers is statistically zero, but frequent high-ability pairs are more than twice as likely to report a match relative to those who never met, as indicated by the significant coefficient of 0.079.

5.2 Peer Effects on Performance

Peer effects on the exam scores Columns 1 to 3 of table 4 present estimates of peer effects on students' exam scores, standardized by discipline and year. We begin with the ability peer effects specification in Panel A, then progressively sharpen the identification of the interaction channel in Panels B and C.

Column 1 of Panel A shows that the average effect of replacing a high-ability peer with a low-ability one on *Exam* is indistinguishable from zero. Column 2 reveals that this zero masks opposing effects by student type: the substitution implies a 5% standard deviation decrease in the exam scores of high-ability students and a 3.9% standard deviation increase for low-ability students, though the latter is not statistically significant. Both estimates are consistent with an ability-mix effect: replacing a high-ability peer lowers the group's average ability,

Table 4: Peer Effects on Performance

Dependent variable:	Exam			Participation > 1		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Ability peer effects						
Low-ability peers	-0.007			0.018*		
	(0.019)			(0.014)		
	[0.675]			[0.043]		
Low-ability student × Low-ability peers		0.039	0.044		0.037*	0.013
		(0.023)	(0.052)		(0.016)	(0.027)
		[0.136]	[0.270]		[0.002]	[0.492]
High-ability student × Low-ability peers		-0.050*	-0.045		0.000	-0.023
		(0.027)	(0.053)		(0.017)	(0.034)
		[0.032]	[0.238]		[0.970]	[0.226]
Peers' average math ability			0.031			-0.144
			(0.264)			(0.134)
			[0.865]			[0.128]
Panel B. Frequency of meetings						
Frequent peers	0.029+			0.009		
	(0.018)			(0.009)		
	[0.094]			[0.291]		
Low-ability student × Frequent peers		0.050*	0.050*		0.020+	0.020+
		(0.024)	(0.024)		(0.015)	(0.015)
		[0.049]	[0.049]		[0.098]	[0.097]
High-ability student × Frequent peers		0.015	0.015		0.002	0.002
		(0.023)	(0.023)		(0.011)	(0.011)
		[0.500]	[0.502]		[0.873]	[0.865]
Peers' average math ability			0.039			-0.118*
			(0.104)			(0.062)
			[0.679]			[0.017]
Panel C. Ex-ante Closeness						
Close peers	0.052*			0.022*		
	(0.017)			(0.011)		
	[0.009]			[0.026]		
Low-ability student × Close peers		0.074*	0.081*		0.039*	0.031*
		(0.023)	(0.026)		(0.015)	(0.014)
		[0.005]	[0.004]		[0.003]	[0.023]
High-ability student × Close peers		0.037	0.031		0.010	0.017
		(0.024)	(0.026)		(0.014)	(0.015)
		[0.143]	[0.243]		[0.450]	[0.215]
Peers' average math ability			0.092			-0.105+
			(0.118)			(0.065)
			[0.385]			[0.056]
Observations	782	782	782	782	782	782
Groups	67	67	67	67	67	67
Students	132	132	132	132	132	132

Notes: The dependent variables are students' standardized score in the disciplines' exams and an indicator of participation grade greater than 1. Standard errors clustered at both student and group levels displayed in parentheses. We show in brackets p-values calculated through the randomization inference procedure based on 10,000 placebo allocations of students.
$+ \, p < 0.1, * \, p < 0.05$

which would tend to harm high-ability students while potentially benefiting low-ability ones through increased interaction with similar peers. Column 3 adds peers' average math score as a control; the point estimates are of similar magnitude, though the negative estimate for high-ability students loses statistical significance.

The positive, although imprecise, estimate for low-ability students in Panel A suggests a channel not driven by ability peer effects. As shown in Table 3, low-ability students are 70% more likely to become socially proximate with another low-ability peer than with a high-ability one. Replacing a high-ability peer with a low-ability one therefore increases not only the number of low-ability classmates but also the expected social proximity of low-ability students within the group — which may be an important operative mechanism. To isolate this channel, Panel B estimates the effect of replacing a non-frequent peer with a frequent one, holding the group's ability composition unchanged on average.

Column 1 of Panel B shows a positive estimated effect on *Exam*, and column 2 shows that this is driven entirely by low-ability students: having one more frequent peer in the group raises their average exam score by 5% of a standard deviation. There are no detectable effects on high-ability students. Because group average ability and the frequency of meetings are independently assigned, adding peers' average ability as a control in column 3 leaves these estimates unchanged.

Replacing a non-frequent peer with a frequent one increases the likelihood of interaction, but still does not maximize social proximity. Table 3 shows that, among low-ability students, non-frequent high-ability peers are the least likely to be socially close. It follows that replacing such a peer with any other type — including a non-frequent low-ability peer — raises the expected social proximity of the group for a low-ability student. Panel C estimates this effect directly. Column 2 shows a 7.4% standard deviation increase in exam scores for low-ability students when a socially distant peer is replaced by a closer one. Since the distant peers for low-ability students are high-ability, this substitution also lowers the group's average ability. Controlling for peers' average math score in column 3, the estimate rises to 8.1% of a standard deviation. For high-ability students, we do not reject the hypothesis of no effect on exam scores. The point estimates in columns 2 and 3 are approximately twice the size of the corresponding Panel B estimates. Since peers' average ability is held constant, this pattern is consistent with the homophily documented in Table 3: replacing a socially distant peer increases the likelihood of interaction with a similar peer for high-ability students, but the effect remains imprecise.

Peer effects on participation in group work Social proximity among students may improve the exam performance of low-ability students through two channels: encouraging additional study outside the classroom, potentially with peer support, or increasing the effort students exert during tutorial sessions themselves. Although we lack data on out-of-classroom activities to test the first channel, participation grades provide evidence on the second. A student receives a participation grade above 1 only when the tutor judges her performance as outstanding during the session, making *Participation* > 1 a direct signal of within-session effort.

Columns 4 to 6 of Table 4 present estimates of peer effects on this indicator. Columns 4 and 5 of Panel A show that replacing a high-ability peer with a low-ability one increases the probability of receiving a participation grade above 1 by 1.8 percentage points in the aggregate, driven by a 3.7 percentage point increase for low-ability students. This effect disappears once peers' average ability is added as a control in column 6. This pattern is consistent with a relative ranking channel: if tutors assess participation relative to the group's ability distribution, a lower group average makes it easier for low-ability students to stand out, independently of any change in their effort. Panels B and C allow us to separate this channel from social proximity. The negative and significant coefficient on peers' average ability in column 6 of both panels confirms that the relative ranking channel remains operative: higher group average ability reduces the probability of outstanding performance. Crucially, however, the positive effect of frequent peers on *Participation* > 1 for low-ability students is unchanged by this control in Panel B, isolating social proximity as a distinct mechanism driving within-session effort.

Panel C, which maximizes the contrast in social proximity, produces the sharpest estimate. Column 6 shows a 3.1 percentage point increase in the probability of outstanding performance for low-ability students. Since this estimate holds conditional on peers' average ability, it isolates the effort response to social proximity from any ability-mix confound. For high-ability students, there is no detectable effect on outstanding performance across any specification, though the point estimates in Panel C are again larger than those in Panel B, consistent with the pattern observed for exam scores.

Heterogeneity by student and peers ability In Table 5, we decompose the effects of close high-ability students. In each case, the baseline category is the socially distant peer type: non-frequent high-ability peers for low-ability students, and non-frequent low-ability peers for high-ability students. This allows us to test whether the close peers effect in Table 4 is driven by a particular peer type or is broadly consistent across all peers with a higher chance of social proximity.

For low-ability students, Panel A shows a significant positive effect on exam scores for frequent high-ability peers ($\beta_2 = 0.078$, p-value $= 0.034$), with point estimates of similar magnitude for frequent low-ability peers ($\beta_1 = 0.081$) and smaller for non-frequent low-ability peers ($\gamma_1 = 0.058$) that are less precisely estimated. For the probability of outstanding performance, frequent low-ability peers produce a significant 5.7 percentage point increase ($\beta_1 = 0.057$,

Table 5: Peer Effects on Performance – Decomposing Social Proximity

	Low-ability students		High-ability students	
	Exam	Part.>1	Exam	Part.>1
Panel A. Close peers: Regression estimates				
Frequent low-ability peers (β_1)	0.081	0.057*	-0.011	-0.004
	[0.168]	[0.044]	[0.775]	[0.838]
Frequent high-ability peers (β_2)	0.078*	0.033+	0.035	0.032
	[0.034]	[0.061]	[0.505]	[0.255]
Non-frequent low-ability peers (γ_1)	0.058	0.040		
	[0.312]	[0.139]		
Non-frequent high-ability peers (γ_2)			0.007	0.032
			[0.889]	[0.250]
Peers' average math ability	-0.076	-0.039	0.173	-0.198
	[0.784]	[0.774]	[0.480]	[0.143]
Panel B. Tests for the equality of coefficients				
$\beta_1 = \beta_2 = \gamma_1$	0.314	0.926		
	[0.855]	[0.624]		
$\beta_1 = \beta_2 = \gamma_2$			1.051	1.698
			[0.592]	[0.434]
Observations	344	344	438	438
Groups	67	67	67	67
Students	132	132	132	132

Notes: The dependent variables are students' standardized score in the disciplines' exams and an indicator of participation grade greater than 1. We show in brackets p-values calculated through the randomization inference procedure based on 10,000 placebo allocations of students. For the joint equality test $H_0 : \beta_1 = \beta_2 = \gamma$, we report the Wald statistic and its p-value also obtained through the RI procedure.
$+ \, p < 0.1, * \, p < 0.05$

p-value $= 0.044$), with positive but less precise estimates for the other two types. The Wald tests reported in Panel B confirm that we cannot reject equality across coefficients for either outcome. The insignificant coefficient on peers' average ability across all outcomes suggests that the close peers effect operates through social proximity rather than through relative ranking within the group. This pattern is consistent with Table 3, which shows that low-ability students have a positive probability of social proximity with all three peer types.

For high-ability students, all point estimates on exam scores and participation outcomes are statistically indistinguishable from zero, and we cannot reject the equality of coefficients across peer types for any outcome. The absence of a social proximity effect is consistent with the aggregate results in Table 4. Although the positive point estimates for high-ability peers in Part.> 1 ($\beta_2 = \gamma_2 = 0.032$) and the negative coefficient on peers' average ability (-0.198) are suggestive of peer teaching and relative ranking channels respectively, neither reaches conventional significance levels in this specification. The evidence from Table 5 therefore does not allow us to attribute the null aggregate result for high-ability students to any specific mechanism — only to rule out that social proximity, as defined here, is the operative channel.

To address potential concerns about multiple hypothesis testing, we apply the stepdown procedure of List et al. (2019) to the coefficients reported in Table 5, defining two separate families of hypotheses — one for low-ability students (β_1, β_2, γ_1) and one for high-ability students (β_1, β_2, γ_2) — each tested across both outcomes simultaneously. The adjusted p-values, reported in Appendix Table A6, show that no individual hypothesis survives the correction at conventional significance levels. This is expected given the penalty imposed by joint testing over six hypotheses simultaneously, but does not affect the main conclusion of the paper: the primary evidence for the effect of social proximity rests on Table 4, where each hypothesis is tested individually and the effects for low-ability students are precisely estimated.

Ruling out participation grade redistribution One might wonder whether the gains in outstanding participation among low-ability students simply reflect a reranking within the group — tutors awarding more outstanding evaluations to some students at the expense of others, rather than a genuine increase in aggregate effort. Under a pure redistribution story, social proximity would shift grades across students but leave the total count unchanged; a denser frequent-peer network should therefore bear no relationship to the number of students exceeding the threshold. We test this null directly by regressing the group-level count of students with outstanding participation on a measure of within-group network density.[5] We collapse the

[5]Network density is computed at the group $\times$ discipline level as the ratio of observed frequent-peer links to the maximum number of possible links in the group. Formally, $\rho_g = \bar{d}_g/(n_g - 1)$, where $\bar{d}_g$ is the mean number of frequent peers across students in group g and n_g is group size. This expression follows from the handshaking lemma, which implies that the mean degree equals twice the number of edges divided by group size, so that dividing by $(n_g - 1)$ recovers the standard density measure (see Appendix A.5).

Table 6: Group-level test: network density and participation

Dep. variable:	All students		Low-ability		High-ability	
No. of Students with Part. > 1	(1)	(2)	(3)	(4)	(5)	(6)
Network density	5.259*	4.285*	3.461*	3.141*	1.799	1.145
	(1.842)	(1.776)	(1.221)	(1.193)	(1.433)	(1.425)
	[0.006]	[0.019]	[0.007]	[0.011]	[0.215]	[0.425]
Mean math ability	-0.662	-0.457	-1.246	-1.179	0.584	0.722
	(0.664)	(0.677)	(0.802)	(0.777)	(0.411)	(0.423)
	[0.323]	[0.503]	[0.126]	[0.135]	[0.162]	[0.093]
Group size	0.288		0.095		0.193	
	(0.213)		(0.095)		(0.162)	
	[0.181]		[0.323]		[0.238]	
Groups	67	67	67	67	67	67

Notes: The dependent variable is the number of students with participation grades above 1 in the group, computed for all students (columns 1–2), low-ability students (columns 3–4), and high-ability students (columns 5–6). Standard errors clustered at the group level displayed in parentheses. We show in brackets p-values calculated through the randomization inference procedure based on 10,000 placebo allocations of students.
$+\, p < 0.1$, $*\, p < 0.05$

data to the group $\times$ discipline level and regress the count of students with Participation > 1 on the within-group network density, controlling for mean math ability and group size. Inference is based on the same 10,000 placebo allocations used throughout the paper. Table 6 shows a positive significant coefficient on network density in all specifications. A denser network is associated with more students achieving outstanding participation, which is inconsistent with a zero-sum redistribution story and supports the interpretation that social proximity raises effort at the group level.

Columns (3)–(4) and (5)–(6) show the result for samples split by student ability. The effect is concentrated among low-ability students, where the coefficient on network density ranges from 3.141 to 3.461 and is significant at the 1% level in both specifications. For high-ability students, the coefficient is smaller and statistically indistinguishable from zero, mirroring the heterogeneity pattern documented in Table 5. To grasp some intuition about the magnitude of these effects, consider moving a group from the mean observed network density (0.532) to the maximum observed density (0.833) — a large shift but still within the support of the data. The point estimate from column (1) implies that such an increase in density would raise the expected number of students with outstanding participation by 1.58, relative to a mean of 1.43 per group — a 110.7% increase over the baseline. The effect is even more pronounced among low-ability students: the same shift in density predicts an increase of 1.04 additional low-ability students with outstanding participation, relative to a mean of 0.52 — a 199.4% increase. These magnitudes suggest that the aggregate effort response to a denser frequent-peer network is economically relevant and driven predominantly by students at the lower end of the ability distribution.

Summary Taken together, the results show that increasing the social proximity of low-ability students — by replacing peers likely to be socially distant with closer ones — raises their average exam score by 8.1% of a standard deviation and increases by 31% the probability of outstanding performance in group work. The decomposition in Table 5 shows that this effect is broadly consistent across all closer peer types, consistent with low-ability students having a positive probability of social proximity with any peer who is not a non-frequent high-ability one. The coefficients are best interpreted as a subject-specific average premium from interacting with a socially close peer. Because student fixed effects absorb semester-wide network benefits, this estimate is a lower bound on the total effect, provided social proximity does not reduce effort — a condition supported by the participation results. Scaling the effect on the exam by the interquartile range of close peers — a shift of 3 peers, from the 25th to the 75th percentile of the distribution — implies a total gain of 24.3% of a standard deviation (6.2% in raw score terms) per discipline. An algorithm redesigned to maximize repeated encounters would produce a comparable shift of 2.85 peers on average, implying a gain of 23.1% of a standard deviation (5.9% in raw score terms), though realizing such gains requires that students interact as intended — a concern documented by Carrell et al. (2013). For high-ability students, the aggregate results show no detectable effects, and the decomposition in Table 5 does not allow us to attribute this null to any specific mechanism. The point estimates suggest that ability composition — rather than social proximity — may shape their outcomes, but none reaches conventional significance levels. The contrast with low-ability students remains sharp: social proximity improves both exam performance and within-session effort for the latter, while the former appear unaffected by the density of socially close peers.

The group-level results in Table 6 rule out the most straightforward redistribution story: if social proximity merely reshuffled a fixed number of outstanding grades across students, network density should not predict the total count within a group, yet it does. The individual-level results further rule out redistribution across ability subgroups — the effect is positive for low-ability students and statistically indistinguishable from zero for high-ability students, rather than positive for one group at the expense of the other. A more subtle redistribution scenario remains possible, however: a tutor could award more outstanding grades to previously less-recognized low-ability students while reducing them for others in the same subgroup, generating a positive average effect while leaving aggregate effort unchanged. We cannot rule this out with the available data, as we only observe each student's end-of-semester average participation grade rather than the full sequence of tutorial evaluations. Nevertheless, even under this residual scenario, the positive average effect within the low-ability subgroup and the aggregate increase in outstanding participation are difficult to reconcile with a purely zero-sum account of the results.

5.3 Robustness Exercises

Table 7 presents three robustness exercises that assess the sensitivity of the main results to alternative measurement and sample choices. We first replicate the main specifications using peer shares rather than counts to verify that the linear-in-counts approximation does not affect the conclusions. We then restrict the sample to first quarter disciplines to address a potential concern with the definition of frequent peers for the discipline that begins in the second quarter. Finally, we redefine peers' ability using their writing admission score rather than their math score.

Share of peers If we use the share of different types of peers instead of counting them, the results are qualitatively unchanged. The coefficient of 0.711 in column 1 of Panel B of Table 7, for example, implies that replacing a non-frequent peer with a frequent one raises the exam performance of low-ability students by approximately 5.7% of a standard deviation, since one peer represents on average 8% of a group. This is close to the 5% standard deviation increase in Table 4.

Only first quarter disciplines Our main analysis uses students' achievement across all 6 disciplines in their first semester. One concern is whether the definitions of frequent and non-frequent peers are appropriate for *Probability*, which begins in the second quarter. A peer is defined as frequent if she shares more than one group with a given student, but once the first quarter ends, some pairs defined as frequent may share only a single group in the second quarter. To assess whether this affects the results, we replicate the main specifications from Table 4 excluding *Probability*. The outcome *Exam* is now restricted to scores from the first quarter only; participation grades are unavailable for this period and we use the same measure as in the main analysis. The estimates are qualitatively similar, suggesting this is not a major concern.

Peers classified by their writing ability Throughout the paper, ability is defined based on performance in the math admission exam, which has the highest correlation with subsequent academic performance. However, language skills may also matter for group work dynamics. Columns 5 and 6 of Table 7 present results redefining ability based on the writing admission exam. Panel A shows the same signs as Table 4 for all outcomes. The one notable difference is that replacing a high writing ability peer with a lower-skilled one significantly reduces the average participation grade of high-ability students by 0.4 percentage points (p-value $= 0.008$),

Table 7: Peer Effects on Performance – Robustness Exercises

	Share of peers		1st quarter disciplines		Peers by writing ability	
	Exam	Part.>1	Exam	Part.>1	Exam	Part.>1
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Ability peer effects						
Low-ability student × Low-ability peers	0.578	0.355	0.056	0.018	0.016	0.018
	(0.571)	(0.333)	(0.050)	(0.027)	(0.036)	(0.024)
	[0.257]	[0.158]	[0.188]	[0.385]	[0.607]	[0.205]
High-ability student × Low-ability peers	-0.387	-0.161	-0.045	-0.017	-0.032	-0.040*
	(0.627)	(0.396)	(0.052)	(0.033)	(0.035)	(0.024)
	[0.425]	[0.516]	[0.271]	[0.409]	[0.242]	[0.008]
Peers' average math ability	0.078	-0.080	0.103	-0.119	0.031	-0.130*
	(0.262)	(0.130)	(0.264)	(0.130)	(0.110)	(0.060)
	[0.683]	[0.408]	[0.613]	[0.245]	[0.761]	[0.012]
Panel B. Frequency of meetings						
Low-ability student × Frequent peers	0.711*	0.301+	0.058*	0.024+		
	(0.275)	(0.198)	(0.027)	(0.015)		
	[0.034]	[0.076]	[0.040]	[0.080]		
High-ability student × Frequent peers	0.146	0.012	0.005	-0.001		
	(0.289)	(0.133)	(0.025)	(0.012)		
	[0.609]	[0.944]	[0.834]	[0.925]		
Peers' average math ability	0.040	-0.117*	0.082	-0.118*		
	(0.105)	(0.062)	(0.105)	(0.065)		
	[0.673]	[0.019]	[0.422]	[0.027]		
Panel C. Ex-ante Closeness						
Low-ability student × Close peers	1.155*	0.519*	0.101*	0.036*	0.060*	0.019
	(0.293)	(0.185)	(0.028)	(0.015)	(0.029)	(0.018)
	[0.003]	[0.006]	[<0.001]	[0.014]	[0.028]	[0.152]
High-ability student × Close peers	0.199	0.159	0.026	0.015	0.026	0.020
	(0.326)	(0.189)	(0.029)	(0.016)	(0.027)	(0.017)
	[0.578]	[0.404]	[0.350]	[0.321]	[0.331]	[0.157]
Peers' average math ability	0.129	-0.086	0.163	-0.096	0.045	-0.119*
	(0.125)	(0.067)	(0.119)	(0.068)	(0.106)	(0.062)
	[0.242]	[0.141]	[0.152]	[0.107]	[0.642]	[0.017]
Observations	782	782	652	652	782	782
Groups	67	67	67	67	67	67
Students	132	132	132	132	132	132

Notes: The dependent variables are students' standardized score in the disciplines' exams and an indicator of participation grade greater than 1. Columns under "Share of peers" use each indicated variable as a proportion of the tutorial group. Columns under "1st quarter disciplines" restrict the sample to disciplines taken during the first quarter. Columns under "Peers by writing ability" classify peers by their writing score in the admission exam. Standard errors clustered at both student and group levels displayed in parentheses. We show in brackets p-values calculated through the randomization inference procedure based on 10,000 placebo allocations of students.
$+\, p < 0.1, *\, p < 0.05$

consistent with writing skills playing a role in the quality of within-session discussion. Effects on the probability of outstanding performance are in the same direction but do not reach conventional significance levels. Panel C shows that redefining close peers using writing ability produces qualitatively similar but weaker estimates than the math-based definition, suggesting that math ability is the more relevant dimension for peer interactions in this setting.

6 Discussion

In this section, we discuss how adjustments in the effort put into the work during the tutorial sessions in response to a potential denser network of peers could have produced different impacts on the performance of students of different ability levels. The arguments follow the simple peer effects model tailored to clarify the potential mechanisms at work under the structure of incentives in the environment we analyze (see the section A.6 in the appendix).

6.1 Interpreting the Results

A student's final score is the product of exam performance and average participation grade, so students have two levers for improving outcomes: doing better on exams and earning higher participation grades. Participation grades below one act as a penalty on exam scores, giving students a strong incentive to reach a grade of one.

The effects on the exam scores Tutorial sessions strengthen students' understanding of the material through active discussion: raising questions, engaging with peers' answers, and working through problems collectively. The key feature of this environment is that the returns to a student's own effort in discussion depend on how actively peers engage — effort is complementary across students. A student who attempts to raise a question benefits more when peers are willing to engage seriously with it, and vice versa. Social proximity amplifies this complementarity if students who know their peers well are more likely to ask questions without fear of signaling low ability in front of strangers (Bursztyn et al., 2019), reducing the social cost of participation. Thus, having close peers in the group can reduce the cost due to some type of peer pressure.

Why does this mechanism generate exam gains for low-ability students but not for high-ability ones? High-ability students tend to arrive at tutorial sessions with a strong prior grasp of the material. For them, incremental gains from more active discussion are small — they are already close to what exam performance they can achieve through group work alone. Low-ability students, by contrast, enter sessions with more room to improve their understanding of the core material, so the same increase in discussion effort translates into a more meaningful gain in exam performance. The asymmetry is not about willingness to exert effort, but about where each group sits relative to what additional engagement can realistically deliver.

The effects on participation Participation grades cluster near one for most students — not because effort is uniformly high, but because tutors might calibrate their grading to the ability composition of each group. A tutor who wants to incentivize effort without discouraging weaker students will set a lower effort bar for low-ability students to reach a grade of one, and a higher bar for high-ability students. This compression in the grading scale limits the variance of participation grades even when underlying effort levels differ substantially across students.

This compression also explains why we do not find effects of social proximity on the continuous participation grade (see Table A5 in the appendix), despite finding effects on exam performance. If most students are already grading near one, a social-proximity-induced increase in effort has little room to move the continuous measure substantially. What it can do is push a student just above one — the threshold for outstanding performance. This is precisely what we find: social proximity raises the probability that low-ability students receive an outstanding participation evaluation by 31 percent relative to the baseline. The discrete threshold captures the effort response that the continuous grade cannot. We conclude with a conjecture that public recognition of outstanding performance may itself reinforce the exam gains we document. Moreira (2019) shows that award receipt increases subsequent effort among winners, suggesting that the participation channel could amplify exam performance beyond the direct effect of social proximity — though we cannot separately identify this reinforcement mechanism in our data.

7 Concluding Remarks

Peer effects estimates might be a valuable resource for policymakers seeking to design groups that maximize aggregate performance. Much of the recent literature on ability peer effects recognize that the pattern of interaction among students matters to produce spillovers on performance. In our paper, we highlighted the importance of social proximity and the incentives for group work in explaining these effects. Our results provided further understanding about how to foster positive spillovers on the achievement of low-skilled individuals.

References

Aronow, P. M., Samii, C., and Assenova, V. A. (2015). Cluster-robust variance estimation for dyadic data. *Political Analysis*, 23(4):564–577.

Borusyak, K. and Hull, P. (2023). Nonrandom exposure to exogenous shocks. *Econometrica*, 91(6):2155–2185.

Brady, R. R., Insler, M. A., and Rahman, A. S. (2017). Bad company: Understanding negative peer effects in college achievement. *European Economic Review*, 98:144–168.

Bursztyn, L., Egorov, G., and Jensen, R. (2019). Cool to be smart or smart to be cool? understanding peer pressure in education. *The Review of Economic Studies*, 86(4):1487–1526.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business and Economic Statistics*, 29(2):238–249.

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

Coveney, M. and Oosterveen, M. (2021). What drives ability peer effects? *European Economic Review*, 136:103763.

Currarini, S., Jackson, M. O., and Pin, P. (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4):1003–1045.

Golsteyn, B. H., Non, A., and Zölitz, U. (2021). The impact of peer personality on academic achievement. *Journal of Political Economy*, 129(4):1052–1099.

Griffith, A. L. and Rask, K. N. (2014). Peer effects in higher education: A look at heterogeneous impacts. *Economics of Education Review*, 39:65–77.

Guryan, J., Kroft, K., and Notowidigdo, M. J. (2009). Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics*, 1(4):34–68.

Hong, S. C. and Lee, J. (2017). Who is sitting next to you? peer effects inside the classroom. *Quantitative Economics*, 8(1):239–275.

Jackson, M. O. and Zenou, Y. (2015). Games on networks. In *Handbook of game theory with economic applications*, volume 4, pages 95–163. Elsevier.

Kimbrough, E. O., McGee, A. D., and Shigeoka, H. (2022). How do peers impact learning? an experimental investigation of peer-to-peer teaching and ability tracking. *Journal of Human Resources*, 57(1):304–339.

Li, T., Han, L., Zhang, L., and Rozelle, S. (2014). Encouraging classroom peer interactions: Evidence from chinese migrant schools. *Journal of Public Economics*, 111:29–45.

List, J. A., Shaikh, A. M., and Xu, Y. (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics*, 22(4):773–793.

Lu, F. and Anderson, M. L. (2015). Peer effects in microenvironments: The benefits of homogeneous classroom groups. *Journal of Labor Economics*, 33(1):91–122.

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

Martin, D. D., Wright, A. C., and Krieg, J. M. (2020). Social networks and college performance: Evidence from dining data. *Economics of Education Review*, 79:102063.

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(Volume 27, 2001):415–444.

Mehta, N., Stinebrickner, R., and Stinebrickner, T. (2019). Time-use and academic peer effects in college. *Economic Inquiry*, 57(1):162–171.

Moreira, D. (2019). Success spills over: How awards affect winners' and peers' performance in brazil. Unpublished.

Presler, J. L. (2022). You are who you eat with: Academic peer effects from school lunch lines. *Journal of Economic Behavior & Organization*, 203:43–58.

Smirnov, I. and Thurner, S. (2017). Formation of homophily in academic performance: Students change their friends rather than performance. *PloS one*, 12(8):e0183473.

A Appendix for Online Publication

A.1 Survey Response Analysis

Table A1 shows there is no statistical difference in the rate of response to the questionnaire applied by the school by student ability level. In the regression, the dependent variable is an indicator for response in the school survey about the students' perceptions about the PBL method.

Table A1: Correlation between ability and response to the survey

	Survey respondent
High-ability student	0.623*
	(0.076)
Student's math ability	0.090
	(0.076)
Student's writing ability	-0.010
	(0.045)
Low-ability student	0.073
	(0.147)
Observations	135

+ $p < 0.1$, * $p < 0.05$
Robust standard errors in parentheses.

A.2 Balance Test Analysis

In Section 3 of the main text, we discussed that, within a given discipline, the average ability of low-ability students' peers tends to be higher than that of high-ability students' peers, due to the small pool of students in each cohort. This pattern can bias the parameter estimates presented in Table 2 away from zero. Indeed, this bias is confirmed by Figures F1 and F2, as well as Tables A2 and A3, which display the distributions of coefficients obtained from regressions run in each of the 10,000 replications of the assignment rule actually used to place students into tutorial sessions.

If we were to rely on between-group variation in peer-related variables to identify peer effects, one potential solution would be to control for the average expected peer quality, as summarized in Tables A2 and A3. The idea would be to compare students who, in expectation, faced the same average peer ability but ended up in groups with different peer composition. However, a simpler and more robust solution was available: leveraging the within-student variation in peer-related variables. This variation remains random, and crucially, the expected average peer quality for a given student across different groups is the same, as peers were drawn with replacement within each discipline.

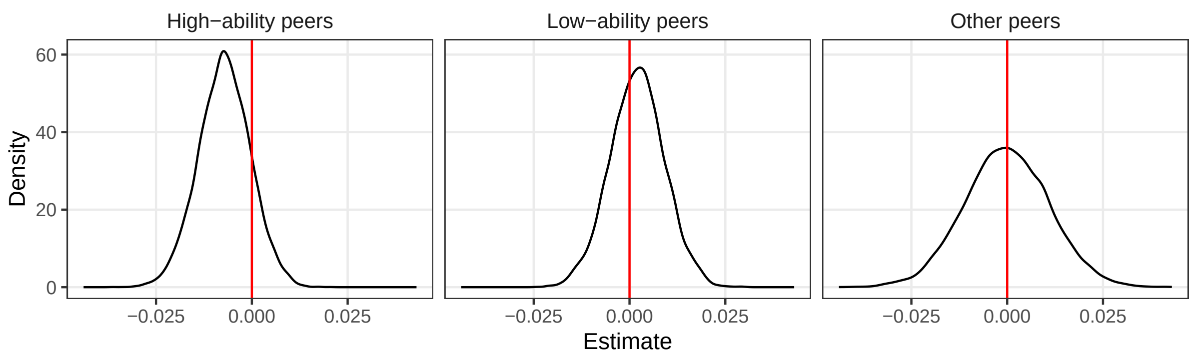
Table A2: Summary statistics for the estimates in Column 1 of Table 2

Coefficient	Mean	SD
High-ability peers	-0.007	0.007
Low-ability peers	0.002	0.007
Other peers	0.000	0.011

Table A3: Summary statistics for the estimates in Column 2 of Table 2

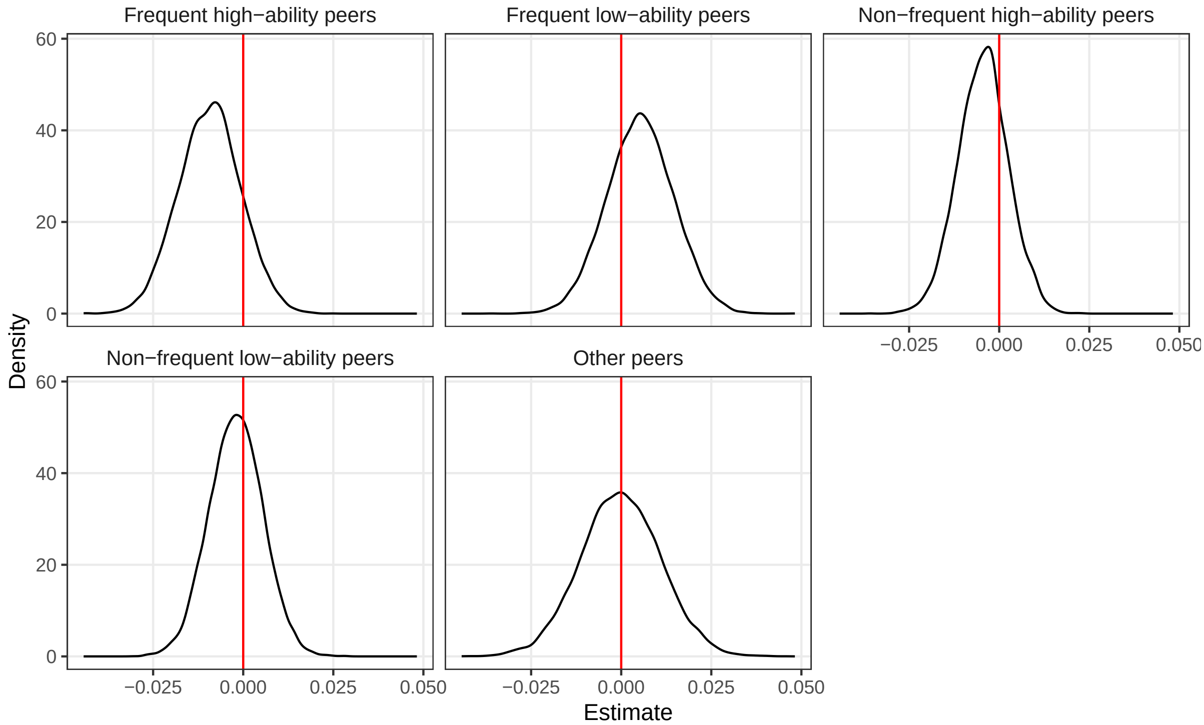
Coefficient	Mean	SD
Frequent high-ability peers	-0.009	0.009
Non-frequent high-ability peers	-0.004	0.007
Frequent low-ability peers	0.005	0.009
Non-frequent low-ability peers	-0.002	0.007
Other peers	0.000	0.011

Figure F1: Placebo distributions – Column 1 of Table 2



Notes: Each panel displays the density of placebo estimates obtained from 10,000 replications of the assignment rule. The dependent variable is the admission score. The vertical red line marks zero.

Figure F2: Placebo distributions – Column 2 of Table 2



Notes: Each panel displays the density of placebo estimates obtained from 10,000 replications of the assignment rule. The dependent variable is the admission score. The vertical red line marks zero.

We will show in Table A4 that augmenting the specification used in Columns 3 and 6 of Table 4 by controlling for the student's expected peer quality in each group leaves the results unchanged. Peer quality is measured as the average number of peers of each ability type computed across the 10,000 placebo allocations.

A.3 Multiple Testing Correction

When testing multiple hypotheses simultaneously, the probability of at least one false rejection—the familywise error rate (FWER)—exceeds the nominal significance level of any individual test. We address this concern by applying the stepdown bootstrap procedure of List et al. (2019), which controls the FWER while retaining more power than classical corrections such as Bonferroni.

We define two separate families of hypotheses, one for low-ability students and one for high-ability students, comprising the coefficients on what we define as close peers across the three outcomes in Table 5. Within each family, the procedure works as follows. Let $T_1 \geq T_2 \geq \cdots \geq T_m$ denote the observed test statistics ordered by absolute value. In the first step, the procedure computes the distribution of $\max_j |T_j|$ under the joint null hypothesis using the 10,000 placebo allocations described in Section 4, and rejects any hypothesis whose test statistic exceeds the 95th percentile of this distribution. In each subsequent step, previously rejected hypotheses are removed from the family and the critical threshold is recomputed using the maximum of the remaining test statistics under the null.

The key insight behind this stepdown structure is that the FWER is only violated when a true null hypothesis is falsely rejected. Once a hypothesis is rejected in an earlier step, it is treated as false and therefore can no longer contribute to a FWER violation. The set of hypotheses requiring protection shrinks with each step, and controlling the FWER over a smaller set requires a less stringent threshold. As a result, the critical value declines monotonically across steps, giving hypotheses with smaller test statistics a progressively better chance of rejection than they would receive under a single-step procedure — while the probability of any false rejection across the entire sequence remains bounded by $alpha$.

Relative to Bonferroni, which applies a uniform penalty of α/m regardless of the dependence structure among test statistics, the List et al. (2019) procedure exploits the empirical correlation across outcomes within each family. When outcomes are positively correlated – as exam scores and participation outcomes plausibly are – the maximum test statistic under the null is smaller in distribution than it would be under independence, and the critical threshold adjusts downward accordingly. The adjusted p-value reported for each hypothesis is the smallest α at which that hypothesis would be rejected by the procedure, and incorporates the full information about the ordering and dependence structure of the family.

Since our inference throughout relies on randomization inference rather than analytical standard errors – motivated by the small number of clusters and the potentially complex dependence structure of the error terms – we construct the test statistics for the stepdown procedure in a manner consistent with that approach. Specifically, rather than studentizing each placebo coefficient by its analytical standard error, we divide it by the standard deviation of its own placebo distribution:

$$t^*_{k,b} = \frac{\hat{\beta}_k^{(b)}}{\mathrm{sd}\left(\hat{\beta}_k^{(\cdot)}\right)},$$

where $\hat{\beta}_k^{(b)}$ is the estimate of coefficient k in placebo draw b and $\mathrm{sd}(\hat{\beta}_k^{(\cdot)})$ is the standard deviation of its placebo distribution across all $B = 10{,}000$ draws. This normalization ensures that test statistics are comparable in scale across hypotheses within the same family – a requirement of the stepdown procedure – without introducing analytical standard errors into the construction of the null distribution. The observed test statistic for hypothesis k is defined analogously as $t^*_k = \hat{\beta}_k \,/\, \mathrm{sd}(\hat{\beta}_k^{(\cdot)})$.

A.4 Additional Results

Table A4: Peer Effects on Performance – Controlling for Expected Exposure

	Baseline		Avg. Exposure control	
	Exam (1)	Part.>1 (2)	Exam (3)	Part.>1 (4)
Panel A. Ability peer effects				
Low-ability student × Low-ability peers	0.044	0.013	0.047	0.013
	(0.052)	(0.027)	(0.051)	(0.028)
	[0.270]	[0.492]	[0.230]	[0.490]
High-ability student × Low-ability peers	-0.045	-0.023	-0.039	-0.024
	(0.053)	(0.034)	(0.052)	(0.034)
	[0.238]	[0.226]	[0.287]	[0.214]
Peers' average math ability	0.031	-0.144	0.013	-0.159+
	(0.264)	(0.134)	(0.261)	(0.134)
	[0.865]	[0.128]	[0.948]	[0.094]
Panel B. Frequency of meetings				
Low-ability student × Frequent peers	0.050*	0.020+	0.051*	0.021+
	(0.024)	(0.015)	(0.024)	(0.015)
	[0.049]	[0.097]	[0.047]	[0.093]
High-ability student × Frequent peers	0.015	0.002	0.016	0.002
	(0.023)	(0.011)	(0.023)	(0.011)
	[0.502]	[0.865]	[0.497]	[0.847]
Peers' average math ability	0.039	-0.118*	0.057	-0.118*
	(0.104)	(0.062)	(0.106)	(0.061)
	[0.679]	[0.017]	[0.549]	[0.018]
Panel C. Ex-ante Closeness				
Low-ability student × Close peers	0.081*	0.031*	0.079*	0.031*
	(0.026)	(0.014)	(0.026)	(0.015)
	[0.004]	[0.023]	[0.005]	[0.022]
High-ability student × Close peers	0.031	0.017	0.026	0.017
	(0.026)	(0.015)	(0.025)	(0.015)
	[0.243]	[0.215]	[0.319]	[0.223]
Peers' average math ability	0.092	-0.105+	0.095	-0.110*
	(0.118)	(0.065)	(0.116)	(0.064)
	[0.385]	[0.056]	[0.370]	[0.047]
Observations	782	782	782	782
Groups	67	67	67	67
Students	132	132	132	132

Notes: The dependent variables are students' standardized score in the disciplines' exams and an indicator of participation grade greater than 1. Columns (1)–(2) replicate the specification with peers' average ability control from Table 4. Columns (3)–(4) add a control for the student's expected exposure to each peer type across disciplines. Standard errors clustered at both student and group levels displayed in parentheses. We show in brackets p-values calculated through the randomization inference procedure based on 10,000 placebo allocations of students.
$+ p < 0.1, * p < 0.05$

Table A5: Peer Effects on Performance – Continuous Participation Measure

Dependent variable:	Participation		
	(1)	(2)	(3)
Panel A. Ability peer effects			
Low-ability peers	-0.001 (0.001) [0.592]		
Low-ability student × Low-ability peers		-0.000 (0.001) [0.875]	-0.001 (0.002) [0.592]
High-ability student × Low-ability peers		-0.001 (0.000) [0.470]	-0.002 (0.001) [0.379]
Peers' average math ability			-0.006 (0.007) [0.564]
Panel B. Frequency of meetings			
Frequent peers	0.001 (0.001) [0.291]		
Low-ability student × Frequent peers		0.000 (0.001) [0.982]	0.000 (0.001) [0.980]
High-ability student × Frequent peers		0.002 (0.001) [0.103]	0.002 (0.001) [0.104]
Peers' average math ability			0.001 (0.004) [0.895]
Panel C. Ex-ante Closeness			
Close peers	0.000 (0.001) [0.722]		
Low-ability student × Close peers		-0.000 (0.002) [0.790]	-0.001 (0.002) [0.752]
High-ability student × Close peers		0.001 (0.001) [0.374]	0.001 (0.001) [0.396]
Peers' average math ability			-0.001 (0.005) [0.825]
Observations	782	782	782
Groups	67	67	67
Students	132	132	132

Notes: The dependent variable is the student's average participation grade in the discipline. Standard errors clustered at both student and group levels displayed in parentheses. We show in brackets p-values calculated through the randomization inference procedure based on 10,000 placebo allocations of students.
$+\ p < 0.1, *\ p < 0.05$

Table A6: Multiple Hypothesis Testing Correction (List, Shaikh & Xu, 2019)

Outcome	Coefficient	t^*	p-value	
			RI unadjusted	MHT adjusted
Panel A. Low-ability students				
Exam	β_1	1.375	0.168	0.318
Exam	β_2	2.137	0.034	0.161
Exam	γ_1	0.993	0.312	0.318
Part.>1	β_1	2.037	0.044	0.167
Part.>1	β_2	1.879	0.061	0.203
Part.>1	γ_1	1.485	0.139	0.318
Panel B. High-ability students				
Exam	β_1	-0.288	0.775	0.988
Exam	β_2	0.663	0.505	0.912
Exam	γ_2	0.136	0.889	0.988
Part.>1	β_1	-0.202	0.838	0.988
Part.>1	β_2	1.141	0.255	0.746
Part.>1	γ_2	1.156	0.250	0.746

Notes: Each panel defines a separate family of hypotheses. t^* is the observed test statistic normalized by the standard deviation of its placebo distribution. The p-value RI unadjusted is the unadjusted randomization inference p-value already reported in Table 5. The p-value MHT adjusted is the familywise-error-rate-adjusted p-value from the stepdown procedure of **?**, using the 10,000 placebo allocations described in Section 4.

A.5 Measuring Network Density

In Section 5 we present a group-level exercise relating network density to the count of students with outstanding participation grades. Here we derive the density measure used in that exercise and show its algebraic equivalence to the standard graph-theoretic definition. Throughout, we proxy the student interaction network by frequent-peer links, on the grounds that pairs meeting repeatedly across disciplines are more likely to develop meaningful social ties.

Let $\mathcal{G}_g = (\mathcal{V}_g, \mathcal{E}_g)$ denote the undirected graph of frequent-peer links within tutorial group g, where $n_g = |\mathcal{V}_g|$ is group size and $|\mathcal{E}_g|$ is the number of frequent pairs.

Step 1: node degree. The degree of student i equals the number of frequent peers she has in the group:

$$d_i = \text{frequent\_peers}_i = \sum_{j \neq i} \mathbf{1}\big\{(i,j) \in \mathcal{E}_g\big\}. \tag{5}$$

Step 2: mean degree via the Handshaking Lemma. Since every edge contributes one to each of its two endpoints, $\sum_{i \in \mathcal{V}_g} d_i = 2\,|\mathcal{E}_g|$, so the group mean is

$$\bar{d}_g \;=\; \text{mean\_freq\_peers}_g \;=\; \frac{1}{n_g} \sum_{i \in \mathcal{V}_g} d_i \;=\; \frac{2\,|\mathcal{E}_g|}{n_g}. \tag{6}$$

Step 3: network density. Network density is the ratio of observed edges to the maximum possible number of edges in a group of size n_g:

$$\rho_g \;=\; \frac{|\mathcal{E}_g|}{\binom{n_g}{2}} \;=\; \frac{|\mathcal{E}_g|}{\frac{n_g(n_g-1)}{2}} \;=\; \frac{2\,|\mathcal{E}_g|}{n_g(n_g-1)}. \tag{7}$$

Step 4: equivalence. Substituting $2\,|\mathcal{E}_g| = \bar{d}_g \cdot n_g$ from (6) into (7) yields

$$\rho_g \;=\; \frac{\bar{d}_g \cdot n_g}{n_g(n_g-1)} \;=\; \frac{\bar{d}_g}{n_g - 1} \;=\; \frac{\text{mean\_freq\_peers}_g}{\text{size}_g - 1}, \tag{8}$$

which is the expression computed as *Network density* in the group-level exercise. Equation (8) shows that dividing the mean individual degree by $(n_g - 1)$ is algebraically identical to dividing the number of observed pairs by the maximum number of pairs – the standard definition of network density. Because the two measures differ only by the scalar $(n_g - 1)$, any regression that controls for group size yields numerically equivalent estimates whether ρ_g or $\bar{d}_g$ is used as the treatment variable.

A.6 A Simple Model of Peer Effects

The basics Any active action within the group, such as raising a question, answering someone else's questions (including the tutor's), or volunteering to explain some concept, entails some level of effort. This is the type of effort the tutor rewards in each tutorial session. Of course, a student who is only sitting in the class and paying attention to the discussion is making some cognitive effort and can learn from peers. But we are interested in the "active" effort, which we simply call "effort" from now on. For simplicity, we refer to "friends" as a pair of students who have some level of interaction during group work.

Here, we discuss four elements needed to set up and solve the student's problem of optimal effort choice in group work: the production function for student's achievement in the exam, the cost of effort, a social interaction component, and the tutor's participation reward scheme. In the model, a student i is characterized by the ability level θ_i. The vector containing the ability of i's peers is $\boldsymbol{\theta}_{-i}$. The effort in group work is e_i and $\mathbf{e}_{-i}$ for the student and peers, respectively, and there is an interaction matrix $\mathbf{G}$ describing the group's network. Students will maximize performance – achievement times participation – net of their effort cost plus a social component that affects the incentive to make effort.

Student's achievement in the exam We assume the student's production function has two components:

$$A(e_i, \theta_i; \mathbf{e}_{-i}, \boldsymbol{\theta}_{-i}) = z(\mathbf{e}_{-i}, \boldsymbol{\theta}_{-i}) + f(e_i, \theta_i)$$

In our environment, the most common form of participation usually involves some form of peer teaching in which a student exposes and develops ideas on the blackboard. This is why both peers' effort and ability impact i's achievement through a continuous and concave function z that we assume to be increasing in both arguments. This makes feasible that students present in the session can, in principle, learn from peers independent of their own effort.

Besides, we assume that a student internalizes the output of their own participation through f, a continuous, strictly increasing, and concave function of effort ($f_e > 0$, $f_{ee} < 0$). We assume that more skilled students perform better for a given level of effort ($f_\theta > 0$). Analogous to the effect of peer teaching in z, the function f is able to capture effects apart from peer interaction, such as developing an exercise on the blackboard without any feedback, which could serve as some practice that improves achievement. The separability in A allows for these types of independent effects. But notice that A can still account for the output from student interactions, where questions and answers increase f or z depending on the role played by the student. We will introduce a premium for interaction below.

The cost of effort The cost of exerting effort is given by $c(e_i, \theta_i)$, which we assume to be an increasing and convex function of effort ($c_e > 0, c_{ee} > 0$), decreasing in θ_i ($c_\theta < 0$), and with the marginal cost of effort also decreasing in θ_i ($c_{e\theta} < 0$). The decreasing effects of ability characterize students' heterogeneity in terms of the cognitive requirements for each student to exert a certain level of effort and, potentially, a peer pressure channel that makes a low-ability student less likely to raise questions (Bursztyn et al., 2019).

The social interaction component A social interaction component

$$s(e_i; \mathbf{e}_{-i}, \mathbf{G}) = \left(\sum_{j \neq i} g_{ij} e_j \right) e_i$$

can be interpreted as either a reduction of cost or an increase in performance due to an increased incentive to exert effort once the student has some level of interaction in the group as indicated by $g_{ij} \in (0, 1]$ for at least some j (there is complementarity in effort). Social proximity is important in our context for two reasons. The ability heterogeneity in cost due to peer pressure can decrease with friends in the group. Besides, even a minimal level of interaction produces a complementarity in effort that raises performance. Importantly, g_{ij} could depend on effort in the sense that helping someone during classwork increases the likelihood of either establishing or strengthening a friendship. However, we make a simplifying assumption that g_{ij} represents only the exogenous portion of such a friendship (the frequent meetings in our environment, for instance), which turns $\mathbf{G}$ into an exogenous interaction matrix.

The participation reward scheme We assume the tutor evaluates participation through an increasing and strictly concave function $p(e_i - \theta_i)$ bounded above by 1 (the maximum participation grade). Remember that to internalize their achievement in the exam, students must be as close to having participation in grade 1 as possible. Thus, the function p implies that getting close to participation grade 1 requires more effort the greater the θ_i. [6]

Student's optimal effort Each student must solve

$$\max_{e_i} \quad p(e_i - \theta_i)A(e_i, \theta_i; \mathbf{e}_{-i}, \boldsymbol{\theta}_{-i}) - c(e_i, \theta_i) + s(e_i; \mathbf{e}_{-i}, \mathbf{G}) \tag{9}$$

Given that $\mathbf{G}$ is exogenous, the concavity and continuity of the objective function plus the assumption that students have a finite set of choices for e_i ensure the existence of a Nash equilibrium $\mathbf{e}^*$. Thus, each student takes $\mathbf{e}^*_{-i}$ as given, and the problem's first-order condition for each student i

$$\underbrace{p'(e_i^* - \theta_i)}_{(i)}\left[z(\mathbf{e}^*_{-i}, \boldsymbol{\theta}_{-i}) + f(e_i^*, \theta_i)\right] + p(e_i^* - \theta_i)\underbrace{f_e(e_i^*, \theta_i)}_{(ii)} = c_e(e_i^*, \theta_i) - \sum_{j\neq i} g_{ij}e_j^* \tag{10}$$

is necessary and sufficient for optimization.

Equation 10 makes it clear that a student adjusts the equilibrium effort in response to exogenous changes in social proximity with peers ($\mathbf{g}_i$) and that this adjustment impact (i) the student's participation grade and (ii) the achievement in the exam. It is worth noting that the effects we can analyze through 10 are still reduced-form: changing g_{ij} implies an adjustment

[6] This assumption can be justified by a setting in which the tutor knows each θ_i and acts to maximize aggregate effort in the group, subject to the constraint that the cost of effort is below a personal threshold for each student so that they do not quit:

$$\begin{aligned} \max_{\mathbf{e}} \quad & \sum_i e_i \\ \text{s.t.} \quad & c(e_i, \theta_i) \leq \bar{c}_i \text{ for all } i \end{aligned}$$

If possible, the tutor would like to induce $\mathbf{e}^*$ so that $c(e_i^*, \theta_i) = \bar{c}_i$ for all i. If the student's cost threshold is not "too much" decreasing in ability, the tutor can induce high-ability students to exert more effort than low-ability ones to get the same participation grade. However, the tutor does not know c exactly and the best he can do is designing a reward scheme with this property.

through 10 for i but also for j, which would induce further adjustment for i, and so on. This is just an example of the *reflection problem* (Manski, 1993). This means that an exogenous impact on peer interaction make the student's own achievement to change through f, but also through z.

A.7 Explaining our findings in terms of the model

The assumption on the participation reward scheme implies that low-ability students get close to 1 at lower effort levels than high-ability students. It means that the marginal contribution of effort to the participation grade $p'(e_i^* - \theta_i)$ approaches zero earlier for the low-ability students (low θ_i). However, for these students, the equilibrium effort level should not be too low so that it jeopardizes achievement through f or if it is insufficient to internalize the achievement through p. Thus, the equilibrium choice e_L^* of a low-ability student θ_L can be such that $p'(e_L^* - \theta_L)$ is very small and then the adjustment in effort from a change in g_L would have, at most, a negligible effect on the participation grade while still having substantial effect on achievement through $A(e_L^*, \theta_L; \mathbf{e}_{-L}^*, \boldsymbol{\theta}_{-L})$.

To explain the results for high-ability students, we need to add assumptions on the behavior of f that do not change the above interpretation for low-ability students. Since f is intended to describe how the student's own effort translates into achievement, we assume that f has an upper limit common to all students (the maximum score in the exam, for instance). Also, we assume that at low effort levels, f increases much faster for high-ability students than for low-ability ones.[7]In a limiting case, a highly skilled student could reach the achievement's upper limit even without attending the tutorial sessions (f would be constant at the upper limit). On the other hand, low-ability students would need to exert more effort to grasp the core of the learning goals to achieve an appropriate exam score.[8] Since the tutor requires more effort from a high-ability student θ_H, our estimates are consistent with the equilibrium effort level being large enough so that both $f_e(e_H^*, \theta_H)$ and $p'(e_H^* - \theta_H)$ are very small.

A.8 Summary Statistics for the Placebo Estimates

Here we provide tables with summary statistics for every distribution of the estimates obtained from the placebo allocations that we used for inference in the text.

[7]Formally, $f_e(e; \theta_H) > f_e(e; \theta_L)$ up to some $\bar{e}$.

[8]The function $f(e, \theta) = 1 - \exp\big(-\theta e\big)$ satisfies the assumptions, for instance.

Table A7: Placebo Distribution – Table 4, Panel A

Coefficient	Min	p5	Mean	Median	p95	Max	Std.Error
Panel A. Ability peer effects							
Outcome: Exam							
Column (1)							
Low-ability peers	-0.066	-0.031	-0.001	-0.001	0.028	0.079	0.018
Column (2)							
Low-ability student × Low-ability peers	-0.090	-0.043	-0.001	-0.001	0.042	0.105	0.026
High-ability student × Low-ability peers	-0.087	-0.040	-0.002	-0.002	0.036	0.083	0.023
Column (3)							
Low-ability student × Low-ability peers	-0.174	-0.066	0.000	0.000	0.065	0.150	0.040
High-ability student × Low-ability peers	-0.152	-0.064	-0.001	-0.001	0.062	0.140	0.038
Peers' average math ability	-0.709	-0.304	0.006	0.007	0.314	0.739	0.188
Outcome: Participation > 1							
Column (4)							
Low-ability peers	-0.036	-0.015	-0.000	-0.000	0.014	0.042	0.009
Column (5)							
Low-ability student × Low-ability peers	-0.052	-0.020	-0.001	-0.001	0.019	0.051	0.012
High-ability student × Low-ability peers	-0.051	-0.021	-0.000	-0.000	0.020	0.054	0.012
Column (6)							
Low-ability student × Low-ability peers	-0.081	-0.031	0.000	0.001	0.031	0.086	0.019
High-ability student × Low-ability peers	-0.083	-0.031	0.000	0.000	0.033	0.075	0.019
Peers' average math ability	-0.381	-0.149	0.006	0.007	0.162	0.364	0.094

Notes: This table presents descriptive statistics for the coefficients on peer variables based on all the regressions performed with the 10,000 placebo allocations generated according to the students assignment rule.

Table A8: Placebo Distribution – Table 4, Panel B

Coefficient	Min	p5	Mean	Median	p95	Max	Std.Error
Panel B. Frequency of meetings							
Outcome: Exam							
Column (1)							
Frequent peers	-0.067	-0.029	0.000	0.000	0.029	0.066	0.018
Column (2)							
Low-ability student × Frequent peers	-0.103	-0.042	-0.001	-0.001	0.041	0.094	0.026
High-ability student × Frequent peers	-0.088	-0.036	0.001	0.001	0.038	0.083	0.023
Column (3)							
Low-ability student × Frequent peers	-0.104	-0.042	-0.001	-0.001	0.041	0.093	0.026
High-ability student × Frequent peers	-0.089	-0.036	0.001	0.001	0.038	0.084	0.023
Peers' average math ability	-0.436	-0.150	0.008	0.008	0.165	0.378	0.095
Outcome: Participation > 1							
Column (4)							
Frequent peers	-0.029	-0.014	0.000	0.001	0.015	0.044	0.009
Column (5)							
Low-ability student × Frequent peers	-0.042	-0.020	-0.000	0.000	0.020	0.045	0.012
High-ability student × Frequent peers	-0.042	-0.019	0.001	0.001	0.020	0.043	0.012
Column (6)							
Low-ability student × Frequent peers	-0.041	-0.020	-0.000	-0.000	0.020	0.045	0.012
High-ability student × Frequent peers	-0.042	-0.019	0.001	0.001	0.020	0.046	0.012
Peers' average math ability	-0.220	-0.075	0.004	0.003	0.086	0.204	0.049

Notes: This table presents descriptive statistics for the coefficients on peer variables based on all the regressions performed with the 10,000 placebo allocations generated according to the students assignment rule.

Table A9: Placebo Distribution – Table 4, Panel C

Coefficient	Min	p5	Mean	Median	p95	Max	Std.Error
Panel C. Ex-ante Closeness							
Outcome: Exam							
Column (1)							
Close peers	-0.074	-0.033	0.000	0.000	0.033	0.071	0.020
Column (2)							
Low-ability student × Close peers	-0.112	-0.044	-0.001	-0.001	0.042	0.101	0.026
High-ability student × Close peers	-0.106	-0.039	0.001	0.002	0.042	0.097	0.025
Column (3)							
Low-ability student × Close peers	-0.112	-0.046	-0.001	-0.001	0.044	0.100	0.027
High-ability student × Close peers	-0.086	-0.043	0.001	0.001	0.044	0.107	0.026
Peers' average math ability	-0.368	-0.165	0.006	0.006	0.181	0.390	0.105
Outcome: Participation > 1							
Column (4)							
Close peers	-0.036	-0.016	0.001	0.001	0.017	0.041	0.010
Column (5)							
Low-ability student × Close peers	-0.049	-0.021	-0.000	-0.000	0.021	0.049	0.013
High-ability student × Close peers	-0.048	-0.021	0.001	0.001	0.022	0.051	0.013
Column (6)							
Low-ability student × Close peers	-0.064	-0.022	-0.000	0.000	0.022	0.051	0.013
High-ability student × Close peers	-0.053	-0.022	0.001	0.001	0.023	0.054	0.014
Peers' average math ability	-0.228	-0.087	0.003	0.002	0.094	0.211	0.055

Notes: This table presents descriptive statistics for the coefficients on peer variables based on all the regressions performed with the 10,000 placebo allocations generated according to the students assignment rule.

Table A10: Placebo Distribution – Table 5

Coefficient	Min	p5	Mean	Median	p95	Max	Std.Error
Low-ability students							
Panel A. Close peers: Regression estimates							
Outcome: Exam							
Frequent low-ability peers (β_1)	-0.254	-0.099	-0.002	-0.002	0.094	0.244	0.059
Frequent high-ability peers (β_2)	-0.146	-0.062	-0.002	-0.003	0.060	0.127	0.037
Non-frequent low-ability peers (γ_1)	-0.281	-0.097	-0.002	-0.002	0.095	0.232	0.059
Peers' average math ability	-1.370	-0.468	-0.002	-0.001	0.455	1.149	0.282
Outcome: Participation $>$ 1							
Frequent low-ability peers (β_1)	-0.107	-0.046	-0.000	0.000	0.045	0.100	0.028
Frequent high-ability peers (β_2)	-0.067	-0.029	0.000	0.000	0.029	0.070	0.017
Non-frequent low-ability peers (γ_1)	-0.127	-0.045	-0.000	-0.000	0.044	0.097	0.027
Peers' average math ability	-0.583	-0.227	0.002	0.003	0.228	0.496	0.138
Panel B. Tests for equality of coefficients							
Outcome: Exam							
$\beta_1 - \beta_2$	-0.232	-0.100	-0.000	0.000	0.100	0.231	0.061
$\beta_1 - \gamma_1$	-0.181	-0.072	-0.000	-0.000	0.071	0.177	0.043
$\beta_2 - \gamma_1$	-0.219	-0.096	-0.000	-0.001	0.097	0.278	0.059
Outcome: Participation $>$ 1							
$\beta_1 - \beta_2$	-0.149	-0.049	-0.000	0.000	0.048	0.108	0.029
$\beta_1 - \gamma_1$	-0.068	-0.033	0.000	-0.000	0.034	0.075	0.020
$\beta_2 - \gamma_1$	-0.098	-0.045	0.000	-0.000	0.046	0.137	0.028
High-ability students							
Panel A. Close peers: Regression estimates							
Outcome: Exam							
Frequent low-ability peers (β_1)	-0.129	-0.061	0.001	0.001	0.062	0.143	0.037
Frequent high-ability peers (β_2)	-0.198	-0.084	0.001	0.001	0.089	0.237	0.053
Non-frequent high-ability peers (γ_2)	-0.211	-0.086	-0.000	-0.001	0.086	0.190	0.052
Peers' average math ability	-0.995	-0.406	0.011	0.011	0.418	0.996	0.250
Outcome: Participation $>$ 1							
Frequent low-ability peers (β_1)	-0.069	-0.032	0.001	0.001	0.034	0.079	0.020
Frequent high-ability peers (β_2)	-0.110	-0.047	-0.000	-0.000	0.045	0.105	0.028
Non-frequent high-ability peers (γ_2)	-0.108	-0.046	-0.001	-0.001	0.044	0.098	0.028
Peers' average math ability	-0.484	-0.212	0.010	0.010	0.236	0.503	0.136
Panel B. Tests for equality of coefficients							
Outcome: Exam							
$\beta_1 - \beta_2$	-0.195	-0.091	-0.001	0.000	0.090	0.199	0.055
$\beta_1 - \gamma_2$	-0.206	-0.085	0.001	0.001	0.087	0.204	0.053
$\beta_2 - \gamma_2$	-0.125	-0.051	0.002	0.002	0.055	0.128	0.032
Outcome: Participation $>$ 1							
$\beta_1 - \beta_2$	-0.107	-0.047	0.001	0.001	0.050	0.117	0.030
$\beta_1 - \gamma_2$	-0.120	-0.045	0.002	0.001	0.049	0.097	0.028
$\beta_2 - \gamma_2$	-0.059	-0.028	0.000	0.000	0.028	0.073	0.017

Notes: This table presents descriptive statistics for the coefficients on peer variables based on all the regressions performed with the 10,000 placebo allocations generated according to the students assignment rule. Panel B reports statistics for linear combinations computed within each placebo draw before summarising.

Table A11: Placebo Distribution – Table 7, Panel A. Ability peer effects

Coefficient	Min	p5	Mean	Median	p95	Max	Std.Error
Panel A. Ability peer effects							
Share of peers							
Outcome: Exam							
Low-ability student × Low-ability peers	-2.067	-0.833	0.009	0.012	0.849	1.985	0.512
High-ability student × Low-ability peers	-2.343	-0.804	-0.000	-0.004	0.800	1.844	0.486
Peers' average math ability	-0.830	-0.313	0.010	0.011	0.332	0.761	0.197
Outcome: Participation > 1							
Low-ability student × Low-ability peers	-1.104	-0.403	0.010	0.009	0.433	0.967	0.253
High-ability student × Low-ability peers	-0.946	-0.394	0.011	0.014	0.416	1.019	0.249
Peers' average math ability	-0.366	-0.155	0.008	0.008	0.166	0.366	0.098
1st quarter disciplines							
Outcome: Exam							
Low-ability student × Low-ability peers	-0.169	-0.070	0.000	0.000	0.073	0.164	0.043
High-ability student × Low-ability peers	-0.159	-0.069	-0.001	-0.001	0.067	0.152	0.041
Peers' average math ability	-0.761	-0.329	0.007	0.007	0.343	0.784	0.204
Outcome: Participation > 1							
Low-ability student × Low-ability peers	-0.086	-0.033	0.001	0.001	0.034	0.089	0.021
High-ability student × Low-ability peers	-0.086	-0.034	0.001	0.001	0.036	0.077	0.021
Peers' average math ability	-0.400	-0.162	0.007	0.007	0.176	0.398	0.103
Peers by writing ability							
Outcome: Exam							
Low-ability student × Low-ability peers	-0.122	-0.051	0.000	0.001	0.052	0.124	0.031
High-ability student × Low-ability peers	-0.114	-0.046	-0.000	0.000	0.045	0.130	0.028
Peers' average math ability	-0.396	-0.154	0.009	0.008	0.172	0.363	0.099
Outcome: Participation > 1							
Low-ability student × Low-ability peers	-0.066	-0.024	-0.001	-0.000	0.023	0.054	0.014
High-ability student × Low-ability peers	-0.062	-0.025	-0.000	-0.000	0.024	0.059	0.015
Peers' average math ability	-0.222	-0.079	0.004	0.003	0.087	0.222	0.051

Notes: This table presents descriptive statistics for the coefficients on peer variables based on all the regressions performed with the 10,000 placebo allocations generated according to the students assignment rule.

Table A12: Placebo Distribution – Table 7, Panel B. Frequency of meetings

Coefficient	Min	p5	Mean	Median	p95	Max	Std.Error
Panel B. Frequency of meetings							
Share of peers							
Outcome: Exam							
Low-ability student × Frequent peers	-1.397	-0.560	-0.021	-0.022	0.533	1.285	0.333
High-ability student × Frequent peers	-1.026	-0.451	0.018	0.020	0.479	1.048	0.284
Peers' average math ability	-0.440	-0.151	0.008	0.007	0.165	0.379	0.095
Outcome: Participation > 1							
Low-ability student × Frequent peers	-0.581	-0.283	-0.001	-0.000	0.277	0.596	0.168
High-ability student × Frequent peers	-0.554	-0.245	0.011	0.013	0.268	0.635	0.155
Peers' average math ability	-0.222	-0.076	0.004	0.003	0.085	0.204	0.049
1st quarter disciplines							
Outcome: Exam							
Low-ability student × Frequent peers	-0.100	-0.046	-0.001	-0.001	0.046	0.108	0.028
High-ability student × Frequent peers	-0.103	-0.039	0.001	0.001	0.041	0.091	0.024
Peers' average math ability	-0.459	-0.160	0.007	0.008	0.177	0.424	0.102
Outcome: Participation > 1							
Low-ability student × Frequent peers	-0.046	-0.023	0.000	0.000	0.022	0.049	0.014
High-ability student × Frequent peers	-0.047	-0.021	0.001	0.001	0.022	0.048	0.013
Peers' average math ability	-0.240	-0.083	0.004	0.004	0.092	0.220	0.053

Notes: This table presents descriptive statistics for the coefficients on peer variables based on all the regressions performed with the 10,000 placebo allocations generated according to the students assignment rule.

Table A13: Placebo Distribution – Table 7, Panel C. Ex-ante Closeness

Coefficient	Min	p5	Mean	Median	p95	Max	Std.Error
Panel C. Ex-ante Closeness							
Share of peers							
Outcome: Exam							
Low-ability student × Close peers	-1.514	-0.638	-0.022	-0.020	0.610	1.375	0.378
High-ability student × Close peers	-1.227	-0.572	0.017	0.019	0.605	1.440	0.358
Peers' average math ability	-0.389	-0.178	0.004	0.004	0.188	0.426	0.111
Outcome: Participation > 1							
Low-ability student × Close peers	-0.862	-0.322	-0.004	-0.005	0.309	0.635	0.192
High-ability student × Close peers	-0.740	-0.309	0.013	0.015	0.330	0.799	0.194
Peers' average math ability	-0.258	-0.092	0.002	0.001	0.099	0.224	0.058
1st quarter disciplines							
Outcome: Exam							
Low-ability student × Close peers	-0.112	-0.049	0.000	0.000	0.049	0.117	0.030
High-ability student × Close peers	-0.100	-0.047	0.000	0.000	0.048	0.122	0.028
Peers' average math ability	-0.377	-0.176	0.008	0.008	0.196	0.466	0.113
Outcome: Participation > 1							
Low-ability student × Close peers	-0.068	-0.024	-0.000	-0.000	0.024	0.057	0.015
High-ability student × Close peers	-0.055	-0.024	0.001	0.001	0.025	0.057	0.015
Peers' average math ability	-0.248	-0.094	0.003	0.002	0.102	0.241	0.060
Peers by writing ability							
Outcome: Exam							
Low-ability student × Close peers	-0.102	-0.046	-0.000	-0.001	0.045	0.092	0.027
High-ability student × Close peers	-0.114	-0.044	0.001	0.001	0.045	0.100	0.027
Peers' average math ability	-0.386	-0.150	0.008	0.008	0.166	0.372	0.096
Outcome: Participation > 1							
Low-ability student × Close peers	-0.048	-0.021	-0.000	-0.000	0.021	0.062	0.013
High-ability student × Close peers	-0.050	-0.023	0.001	0.001	0.024	0.056	0.014
Peers' average math ability	-0.222	-0.076	0.004	0.003	0.086	0.207	0.049

Notes: This table presents descriptive statistics for the coefficients on peer variables based on all the regressions performed with the 10,000 placebo allocations generated according to the students assignment rule.

Table A14: Placebo Distribution – Table A4, Columns (3)–(4)

Coefficient	Min	p5	Mean	Median	p95	Max	Std.Error
Panel A. Ability peer effects							
Outcome: Exam							
Low-ability student × Low-ability peers	-0.173	-0.064	0.001	0.001	0.065	0.142	0.039
High-ability student × Low-ability peers	-0.147	-0.062	0.001	0.001	0.063	0.138	0.037
Peers' average math ability	-0.673	-0.302	0.003	0.003	0.309	0.703	0.185
Outcome: Participation > 1							
Low-ability student × Low-ability peers	-0.082	-0.030	0.001	0.001	0.032	0.087	0.019
High-ability student × Low-ability peers	-0.082	-0.031	0.001	0.001	0.033	0.078	0.019
Peers' average math ability	-0.394	-0.151	0.005	0.005	0.162	0.379	0.095
Panel B. Frequency of meetings							
Outcome: Exam							
Low-ability student × Frequent peers	-0.103	-0.042	0.000	-0.000	0.043	0.097	0.026
High-ability student × Frequent peers	-0.092	-0.038	-0.000	0.000	0.037	0.084	0.023
Peers' average math ability	-0.426	-0.151	0.008	0.007	0.166	0.372	0.096
Outcome: Participation > 1							
Low-ability student × Frequent peers	-0.042	-0.020	-0.000	0.000	0.020	0.046	0.012
High-ability student × Frequent peers	-0.043	-0.020	-0.000	0.000	0.020	0.045	0.012
Peers' average math ability	-0.227	-0.077	0.003	0.003	0.085	0.198	0.049
Panel C. Ex-ante Closeness							
Outcome: Exam							
Low-ability student × Close peers	-0.111	-0.045	0.000	0.000	0.046	0.100	0.028
High-ability student × Close peers	-0.089	-0.044	-0.001	-0.001	0.043	0.106	0.026
Peers' average math ability	-0.361	-0.165	0.007	0.006	0.182	0.390	0.105
Outcome: Participation > 1							
Low-ability student × Close peers	-0.062	-0.022	0.000	0.000	0.022	0.051	0.013
High-ability student × Close peers	-0.054	-0.023	-0.000	-0.000	0.022	0.053	0.014
Peers' average math ability	-0.225	-0.087	0.002	0.001	0.093	0.214	0.055

Notes: This table presents descriptive statistics for the coefficients on peer variables based on all the regressions performed with the 10,000 placebo allocations generated according to the students assignment rule. Results correspond to columns (3)–(4) of Table A4, which add a control for the student's average exposure to each peer type across disciplines.

Table A15: Placebo Distribution – Table A5

Coefficient	Min	p5	Mean	Median	p95	Max	Std.Error
Panel A. Ability peer effects							
Column (1) No heterogeneity							
Low-ability peers	-0.004	-0.002	-0.000	-0.000	0.002	0.004	0.001
Column (2) Heterogeneity							
Low-ability student × Low-ability peers	-0.006	-0.003	-0.000	-0.000	0.003	0.007	0.002
High-ability student × Low-ability peers	-0.004	-0.002	-0.000	-0.000	0.002	0.004	0.001
Column (3) + Peers' avg. ability							
Low-ability student × Low-ability peers	-0.010	-0.004	-0.000	-0.000	0.004	0.010	0.002
High-ability student × Low-ability peers	-0.007	-0.003	0.000	0.000	0.003	0.008	0.002
Peers' average math ability	-0.040	-0.016	0.000	0.000	0.016	0.040	0.010
Panel B. Frequency of meetings							
Column (1) No heterogeneity							
Frequent peers	-0.004	-0.002	-0.000	-0.000	0.002	0.004	0.001
Column (2) Heterogeneity							
Low-ability student × Frequent peers	-0.007	-0.003	-0.000	-0.000	0.003	0.007	0.002
High-ability student × Frequent peers	-0.004	-0.002	0.000	0.000	0.002	0.004	0.001
Column (3) + Peers' avg. ability							
Low-ability student × Frequent peers	-0.007	-0.003	-0.000	-0.000	0.003	0.007	0.002
High-ability student × Frequent peers	-0.004	-0.002	0.000	0.000	0.002	0.004	0.001
Peers' average math ability	-0.021	-0.008	0.000	0.000	0.009	0.018	0.005
Panel C. Ex-ante Closeness							
Column (1) No heterogeneity							
Close peers	-0.005	-0.002	-0.000	-0.000	0.002	0.004	0.001
Column (2) Heterogeneity							
Low-ability student × Close peers	-0.007	-0.003	-0.000	-0.000	0.003	0.006	0.002
High-ability student × Close peers	-0.005	-0.002	0.000	0.000	0.002	0.005	0.001
Column (3) + Peers' avg. ability							
Low-ability student × Close peers	-0.007	-0.003	-0.000	-0.000	0.003	0.006	0.002
High-ability student × Close peers	-0.005	-0.002	-0.000	-0.000	0.002	0.005	0.001
Peers' average math ability	-0.022	-0.009	0.000	0.000	0.010	0.021	0.006

Notes: This table presents descriptive statistics for the coefficients on peer variables based on all the regressions performed with the 10,000 placebo allocations generated according to the students assignment rule. The dependent variable is the student's average participation grade in the discipline.