Social and Cognitive Peer Effects: Experimental Evidence from Selective High Schools in Peru

Román Andrés Zárate*

November 14, 2018

Job Market Paper

Please find the latest version here.

Abstract

A growing literature emphasizes the importance of social skills in the labor market. However, to date, no study addresses the role of peer characteristics in the formation of social skills. This paper reports estimates of cognitive and *social* peer effects from a large-scale field experiment at selective boarding schools in Peru. My experimental design overcomes some methodological challenges in the peer effects literature. I randomly varied the characteristics of roommates with two treatments: (a) less or more sociable peers (identified by their position in the school's friendship network before the intervention) and (b) lower- or higher-achieving peers (identified by admission test scores). While more sociable peers enhance the formation of social skills, higher-achieving peers do not improve academic achievement; in fact, they further reduce the academic performance of lower-achieving students. These results appear to be driven by students' self-confidence. I interpret these findings in the context of a simple self-confidence model where students infer their skills by interacting with their peers.

*Department of Economics, MIT. Email: razarate@mit.edu. I am deeply grateful to my advisors Joshua Angrist, Esther Duflo, and Parag Pathak for their guidance throughout this project. I also thank Daron Acemoglu, Abhijit Banerjee, and Frank Schilbach for their helpful suggestions. This paper would not have been possible without the support of the staff of the Ministry Education of Peru; I especially thank María Antonieta Alva and Cinthia Wicht for their collaboration and support, and Paola Encalada for her dedicated work at different stages of this project. I also thank Andrei Bartra, Pía Basurto, Antonio Campos, José Carlos Chávez, Yessenia Collahua, Gerald Eliers, Fiorella Guevara, Diego Jara, Ricardo Montero, Raul Panduro, Barbara Sparrow, Guillermo Trefolgi, and Paola Ubillus for their help. I am also grateful to all the directors of the COAR Network for their work during the implementation. Heather McCurdy provided generous help with MIT Institutional Review Board (IRB) documentation. I am particularly grateful to Sydnee Caldwell, Joaquín Klot, Matt Lowe, and Francine Loza for their feedback. Claudia Allende, Zach Brown, Alonso Bucarey, Adriana Camacho, Lorena Caro, Nicolás de Roux, Josh Dean, Juan Dubra, Leopoldo Fergusson, Chishio Furukawa, Isabel Hincapie, Nicolás Idrobo, Juán Galán, Arda Gitmez, Gabriel Kreindler, Stephanie Majerowicz, Santiago Melo, Mateo Montenegro, Andrés Moya, Alan Olivi, Gautam Rao, Cory Smith, Vira Semenova, Román David Zárate, David Zarruk, and the participants at the Development, Labor, Organizational lunches at MIT and NEUDC 2018 also provided useful comments and suggestions. Jorge Alva and María Antonieta Luperdi were very generous during my stay in Peru. This project is part of the MineduLab. Funding for this project was generously provided by the Weiss Family Fund and the Spencer/NAEd dissertation fellowship. The experiment was approved by the MIT IRB (ID 1702862092), and is registered at the AEA RCT Registry (ID 0002600).

1 Introduction

Social skills have a deep influence on individuals' well-being and labor market success. For instance, social skills are important for communication within organizations and team productivity (Woolley et al., 2010). They facilitate interactions between people, and thus cannot be easily substituted by automation (Autor, 2014). Recent empirical evidence shows that the labor market increasingly rewards social skills (Deming, 2017), and that social skills are complementary to cognitive skills (Weinberger, 2014). Yet there is little evidence on how social skills are formed.

Policy makers may be able to use peer effects to influence the formation of social skills. Intuitively, students could develop these skills by interacting with highly sociable peers. Sociable students may also affect the formation of their peers' cognitive skills. For example, it may be easier for students to befriend sociable peers, and having more friends improves academic achievement (Lavy and Sand, 2012). Likewise, some evidence suggests that students only benefit from being in school with higher-achieving peers when they are studying together (Carrell et al., 2013). Therefore, peers' social and cognitive skills could have complementary effects on academic achievement. While social scientists have extensively studied peer effects on academic achievement, behaviors, and racial attitudes (Epple and Romano, 2011; Sacerdote, 2011; Boisjoly et al., 2006), to my knowledge there is no evidence of the impact of peers' sociability on the formation of cognitive or social skills.

This paper reports estimates of social and cognitive peer effects from a large-scale field experiment at selective boarding schools in Peru. While other studies have exploited random assignment to dormitories and classrooms, I use a novel experimental design to generate large variation in peer skills. Specifically, I assign students to two cross-randomized treatments in the allocation to beds in a dormitory: (1) less or more sociable peers, and (2) lower- or higher-achieving peers. This design surmounts many of the challenges with traditional approaches to study peer effects (Manski, 1993; Angrist, 2014; Caeyers and Fafchamps, 2016).

To classify students as less vs. more sociable, I use the eigenvector centrality in the social network the year before the intervention.[1] Eigenvector centrality is a good indicator of sociability. It measures a student's influence in the network, accounting for the fact that influential individuals are connected to other influential individuals as well. In the context of my study, this indicator is highly correlated with other metrics introduced by psychologists to measure social skills. To classify students as higher vs. lower achieving, I use students' scores from the schools' admissions tests, which include math and reading comprehension scores.

[1] I evaluate the school's aggregate social network, which includes dorm preferences, friendships, study, and social partnerships.

I show that the allocation influenced the social network formation in the schools. Students befriend, study, and play more with peers that were assigned to nearby beds; the closer the peer, the stronger the interaction. Being neighbors in a dormitory increases the likelihood of social interactions by 18 percentage points. While the proximity effect is no different for students assigned to higher- and lower-achieving peers, it is slightly higher for students assigned to more sociable peers.

I estimate the impact of each treatment on the formation of social and cognitive skills. To measure social skills, the primary outcome is a social skills index that includes psychological tests and the number of peers that perceive the student as a leader, or a popular, friendly, or shy person. By using peers' perceptions to measure a student's social skills, I account for biases in self-reported psychological tests. I also present results for the two "Big Five" personality traits that are related to social skills: (i) extraversion, characterized by positive affect and sociability and (ii) agreeableness, the tendency to act in a cooperative and selfless manner (McCrae and John, 1992; John and Srivastava, 1999; Almlund et al., 2011). This paper includes several social skills measures to assess the robustness of the results. I use grades and test scores in math and reading comprehension to measure cognitive outcomes.

I find that sociable peers have a positive effect on a student's social skills. Students that were randomly assigned to dormitories with more sociable peers have a higher social skills index—0.067 standard deviations—after the intervention. This effect is mainly driven by the impact on students that were less sociable at baseline. These results are consistent with the impacts on the Big Five personality traits, and on measures that account for biases in self-reported tests—students assigned to more sociable peers are perceived as more friendly and popular. I do not find that having more sociable peers affects a student's cognitive skills.

By contrast, I find that higher-achieving peers have no impact on the average student's social or cognitive skills. Furthermore, my results suggest that higher-achieving peers decrease the academic achievement of lower-achieving students. These effects are similar for grades and test scores, and for both math and reading comprehension.

I exploit the experimental variation in a two-stage least-squares (2SLS) model that jointly estimates the impact of roommates' sociability and academic achievement on students' outcomes. This model accounts for imperfect compliance between the assignment to treatments and actual roommates. The results from the 2SLS model are consistent with the treatment effects. For the average student, a one-standard-deviation increase in their roommates' sociability has a positive effect on social skills but no effect on test scores in math or reading. For students that were less sociable at baseline, the impact of peers' sociability on social skills is 117% greater with an effect of 0.237 standard deviations. Roommates' academic achievement does not affect social or cognitive outcomes, on average. However, for lower-achieving students, a one-standard-deviation increase in their roommates' academic scores leads to a reduction in math and reading scores of 0.082 and 0.122 standard deviations, respectively.

My results are twofold. First, I provide evidence that, for less sociable students, exposure to more sociable peers has a positive effect on social skills. Second, I find that exposure to higher-achieving peers does not positively influence academic achievement; there is evidence suggesting that they further decrease the academic achievement of low achievers. Therefore, my main conclusion is that *while sociable peers make you more sociable, higher-achieving peers do not improve your academic learning*.

I then explore the potential mechanisms that drive my results. I examine whether the change in students' self-confidence in their skills is a valid mechanism. First, I show that my results are consistent with a simple model of self-confidence based on Compte and Postlewaite (2004). Second, I provide suggestive empirical evidence that changes in self-confidence explain my findings.

The idea behind the model is that, while more sociable peers make less sociable students feel better about their social skills, higher-achieving peers make lower achieving students feel worse about their cognitive skills. Under this framework, success in social or cognitive activities is a function of self-confidence, and self-confidence depends on past successes. It is easier for less sociable students to engage in social activities with more —rather than less— sociable peers, and these interactions make students more successful. Success translates into more self-confidence in social skills, making students more likely to do well in future interactions with both roommates and other people. By contrast, lower-achieving students feel less accomplished when assigned to higher-achieving peers. If students think they are doing worse, they are less confident in their academic skills, driving down their investment in cognitive activities.

I find empirical evidence consistent with self-confidence driving my results. First, less sociable students report better social interactions with their roommates when assigned to more sociable peers. In particular, they are happier with their dormitory assignments, and indicate that their roommates are more empathetic towards them. Likewise, there is evidence suggesting that less sociable students gain self-confidence in their social skills when assigned to more sociable peers. By contrast, lower-achieving students report less self-confidence in their cognitive abilities when assigned to higher-achieving peers.

I also rule out the possibility that the number of social interactions between students and their roommates is driving the empirical findings. I explore two questions to reach this conclusion: (1) do less sociable students have more social interactions with their peers when assigned to more sociable peers?, and (2) do lower-achieving students interact less with their peers when assigned to higher-achieving peers? I find evidence against both hypotheses. Less sociable students have a similar number of social interactions with their peers, regardless of their random assignment to the more sociable peers treatment. Likewise, although lower-achieving students are studying with higher-achieving peers, they still experience declines in academic achievement.

This paper builds on and contributes to three strands of the literature: (i) the formation of social skills, (ii) social networks, and (iii) the identification and consequences of peer effects.

My results explore how peer characteristics affect the development of social skills in school, extending the literature on the formation of social skills. While a substantial body of evidence documents positive and increasing returns to social skills in the labor market (Deming, 2017), little is known about how social skills are formed. There are a few exceptions: Rao (2013) shows that rich students are more altruistic and discriminate less when they are exposed to poor peers. Similarly, Falk et al. (2018) find that a mentorship program in Germany increased children's pro-sociality. Finally, there is some evidence of how income (Akee et al., 2018) and incentives (Donato et al., 2017) might affect the Big Five personality traits.[2]

This paper also builds on the literature on social networks. While some evidence highlights the role of eigenvector centrality for the diffusion of microfinance (Banerjee et al., 2013) and the monitoring of savings decisions (Breza and Chandrasekhar, 2018), this paper confirms that eigenvector centrality is correlated with social skills as measured by psychological tests. Moreover, my results show that having highly central peers has a positive effect on a student's social skills.

My experimental design and results reconcile some of the evidence in the peer effects literature. Most previous empirical studies of peer effects focus on baseline test scores of peers, and employ one of two methodologies for identification: they exploit either the random formation of groups or quasi-experimental variation in the skills of peers. Studies that use the former method find small positive peer effects in small groups such as dormitories (Epple and Romano, 2011; Sacerdote, 2001), and sizeable significant effects in large groups such as classrooms (Duflo et al., 2011), squadrons (Carrell et al., 2009), or large-size dormitories (Garlick, 2018). Studies that employ the second approach use exogenous variation in peer characteristics. For example, Abdulkadiroğlu et al. (2014) use school-specific admission cutoffs to estimate peer effects in exam schools in Boston and New York, and Duflo et al. (2011) use cutoffs from a tracking system to estimate peer effects in Kenya. In contrast to studies that use random allocation to groups, quasi-experimental studies have found zero peer effects.[3]

One potential explanation for the broad range of estimates is the methodological problems associated with studies that exploit random allocation to groups (Manski, 1993; Angrist, 2014). In general, when students are randomized into groups, all groups are very

[2] The Big Five personality traits are: openness to experience, conscientiousness , extraversion, agreeableness, and emotional stability.

[3] Garlick (2018) finds a negative impact of tracking for low-scoring students in a university in South Africa. He argues that this result is attributable to peer effects. However, as the author points out, the research design can't rule out that assignment to low-track dormitories has negative psychological effects on students.

similar by design. Therefore, these studies rely on weak variation in peer characteristics, which causes them to overestimate the magnitude of peer effects. My experimental design directly addresses this concern by guaranteeing systematic variation in group composition. Consistent with the quasi-experimental evidence, my results show that the impact of peers' academic achievement is either a precise zero or negative, ruling out positive peer effects.

Finally, I find that in my setting, academic peer effects operate not through the number of social interactions but through changes in students' self-confidence. The fact that lower-achieving students are interacting with higher-achieving peers and yet have lower academic achievement contradicts previous hypotheses in the literature (Carrell et al., 2013). The lower self-confidence of the lower-achieving students is consistent with the "big-fish-little-pond" effect from psychology (Marsh and Parker, 1984).[4] The lower self-confidence of the lower-achieving students is also aligned with recent empirical evidence in economics showing that students have lower self-concept in higher-achieving schools (Fabregas, 2017) and that a lower perceived ranking can affect later life outcomes (Ribas et al., 2018). Similarly, this indicates that peers not only affect students' behaviors because they want to send signals to avoid peer group rejection (Fryer and Austen-Smith, 2005; Bursztyn et al., 2018). It also illustrates how students extract information on their skill level from their interactions with peers.

The rest of the paper is organized as follows. Section 2 describes the setting of exam schools in Peru. Section 3 presents the experimental design. Section 4 shows the balance of the randomization and the impact of treatments on peer and friends characteristics. Section 5 describes the outcomes and outlines the empirical strategy. Section 6 documents the results on skill formation. Section 7 discusses potential mechanisms for the results. Section 8 concludes.

2 Setting: Exam Schools in Peru

The Peruvian Government operates a series of exam schools, *Colegios de Alto Rendimiento* (the COAR Network) to provide a high-quality education for talented low-income students during the last three years of secondary school. The first exam school opened near Lima, the capital of Peru, in 2010. As of 2017, there is now a COAR school in each of the country's 25 regions. There are 100 slots per cohort in each school, except for the school in Lima, which has 300.

The COAR Network meets the standards of elite private high schools in Latin America, where students have access to all the required inputs for a high-quality education. COAR are boarding schools, deliberately located close to the capital city of each region to

[4] Marsh and Parker (1984) described the big-fish–little-pond effect, whereby equally able students have lower academic self-concepts in high-ability schools than in low-ability schools.

reduce daily transportation costs for both families and the government. Upon admission, students receive school materials, uniforms, and a personal laptop for school use. All of the schools have a high-quality infrastructure, including a library and excellent scientific laboratories. Students also have the option of obtaining a world-renowned International Baccalaureate (IB) degree. Teachers are hired outside the public school system and receive higher salaries. The government covers all the necessary operating expenses, including laundry service and food.

Applicants are eligible for admission into COAR if they ranked in the top 10 of their public school cohort in the previous academic year. The admissions process consists of two rounds. In the first round, applicants take a written test in reading comprehension and mathematics. The highest-scoring applicants move onto a second round, during which psychologists rate them based on two activities: a one-to-one interview, and the observation of peer interactions during a set of tasks. I refer to these scores as the *interview* and *social fit* scores, respectively. Admissions decisions are determined by a composite score of all three tests, by the region of origin, and by the applicant's school preferences.

Before the experiment, school directors implemented their own individual systems to allocate students to dormitories and classrooms. Most schools allegedly fostered multicultural diversity by mixing students from different regions within the same dormitory. There was also variation across schools in how they allocated first-year students to classrooms. Classroom assignment for students in the upper cohorts depends on whether students apply for the IB degree and the track they choose for this program.

3 Experimental Design

This section presents the experimental design. The objective of the experiment is to estimate the impact of peers' sociability and academic achievement on students' outcomes. To do this, it is necessary to ensure systematic variation in peer characteristics across treatments. I do so by classifying students into types according to sociability and academic achievement, and by randomizing students into groups with systematic variation in the type of peer. There is substantial variation in peer characteristics across these groups, surmounting the weak variation problem pointed out by Angrist (2014) in other peer effects studies.

This section is divided as follows. First, I describe the data that was available before the intervention. Second, I illustrate how I used this data to classify students according to sociability and academic achievement. Third, I explain how students were randomized to groups with different types of peers, and describe how I used this assignment to allocate students to dormitories in the schools. Figure 1 illustrates the project's timeline.

3.1 Data

3.1.1 Administrative Data

Administrative data on student demographics and baseline scores was collected as part of the admissions process or from existing government databases. For all students enrolled in the COAR Network in 2017, I have data on admissions test scores in three categories: (i) the written test in math and reading comprehension, (ii) the admissions interview, and (iii) the social fit score determined by a team of psychologists.

In addition, I exploit existing government data to describe students' socio-demographic characteristics. The socio-demographic data I use is employed by the Government of Peru to determine households' eligibility for national social programs, and is available for 85% of students. It includes whether a student comes from a household classified as poor or extremely poor, and whether they come from a rural area.

Column 1 of Table 1 reports descriptive statistics for students in the COAR Network. Although these schools target students from the public school system, admitted students have diverse social and economic backgrounds. For example, 35-37% of the students come from poor households, and 20% from extremely poor households. Likewise, 26% of students come from rural households.

For the 2015-16 cohorts, the Ministry of Education also administered psychological tests. Some of these tests incorporate measures of social skills, including emotional intelligence (Law et al., 2004) and the "Reading the Mind in the Eyes" test (Declerck and Bogaert, 2008). Appendix C describes these tests in detail.

3.1.2 Surveys

With the Ministry of Education, we administered an online survey to measure social interactions and non-cognitive skills for students in the 2015 and 2016 cohorts. A team of psychologists monitored the survey. The compliance rate was above 95% for each school.

The survey asked students to list the names of their peers in four distinct categories of social interactions: (i) roommate preferences (students were told that their answers to this question could affect their dormitory assignment), (ii) friends , (iii) study mates, and (iv) people with whom they socialized or engaged in social activities. Appendix Table A.1 shows three statistics for each category of the network: total degree, mutual degree, and eigenvector centrality. The average mutual degree is half of (or lower than) the average total degree. For example, when we consider a broad social network that aggregates all four questions about social interactions, students report having 7.02 connections on average, of which only 3.24 are mutual.

The survey also included questions on students' perception by their peers. Students were asked to rank up to five peers in the categories of leadership, friendliness, popularity, and shyness. Table A.1 shows descriptive statistics for these variables. On average, a student was named by 3.48 of her peers as the best leader, by 3.59 as the most friendly, by 3.21 as the most popular, and by 2.59 as the shyest.

3.2 Classifying Students According to Academic and Social Skills

I use data from the admissions process and baseline social networks survey to identify more sociable and higher-achieving students.

I use the test score in the first round of the admissions process—that evaluates students in math and reading comprehension—to characterize students as lower- or higher-achieving at baseline. For each school-by-grade-by-gender cell, students above the cell-specific median are classified as higher achieving, and those below the median as lower achieving.

To identify more and less sociable students, I rely on the social networks baseline survey described in the previous section. I use the eigenvector centrality of an aggregate undirected social network that groups the four categories of social interactions described above; Banerjee et al. (2013) and Banerjee et al. (2014) perform a similar aggregation. Other studies have used eigenvector centrality to measure sociability, predict the diffusion of information in other contexts (Banerjee et al., 2014; Beaman and Dillon, 2018), and show how more central individuals can do a better job at monitoring savings decisions (Breza and Chandrasekhar, 2018). I use the same strategy to classify students as lower or higher achieving. That is, students with an eigenvector centrality above the cell-specific median are classified as more sociable, and those below the cell-specific median as less sociable. Appendix Figure A.2 shows that in the context of exam schools in Peru, eigenvector centrality and admissions test scores are positively correlated.

Since first-year students did not complete the baseline survey in 2016, eigenvector centrality at baseline is not available for this cohort. Therefore, I cannot study the impact of more sociable peers for this cohort; I only analyze the higher-achieving peers treatment.

Table A.1 (columns 2 to 5) presents descriptive statistics of the baseline social networks by student type. More sociable students have a better position in the schools' social networks, with a larger average degree, mutual degree, and eigenvector centrality for the four social networks reported (roommate preferences, friends, study mates, and social partnerships). For example, in the general network more sociable students have, on average, 4 more connections and 1.4 more mutual connections than less sociable students. More sociable students are also perceived as friendly by 4.6 peers on average, while only 2.5 peers perceive less sociable students as friendly.

More interestingly, I also find a large statistically significant correlation of eigenvector centrality and my set of indicators of social skills. Appendix Table A.2 reports standardized coefficients of an OLS regression of social skills measures[5] on the three admissions

[5]Some of these variables were collected before or after the intervention. They are described in detail in section 5.1 and Appendix C.

test scores, and on the eigenvector centrality of the baseline social network controlling for school×grade×gender fixed effects. For most of my social skills indicators, eigenvector centrality has a stronger correlation than admissions test scores. These results confirm that individuals who are assessed as very central in the schools' social networks at baseline also have highly developed social skills.

3.3 Randomization

To estimate the impact of peers' sociability and academic achievement on students' outcomes, I randomized students to two treatments: (1) more sociable peers, and (2) higher-achieving peers. In the previous section, I explained how students were classified into more sociable and higher-achieving students. Here I explain the details of the randomization.

3.3.1 Groups of Peers

By randomizing the type of peer that students have, instead of the simple randomization to groups, I assure that students in my study are exposed to peers with different levels of skills. This is a novel approach and is central to my study. It differs from the more traditional approach that exploits random assignments to groups; where, by virtue of the randomization, peer characteristics are the same in expectation —although there will be small variation across groups in the realized sample.

The experimental design accounts for the fact that a student, not only receives a treatment, but is also a treatment for her peers. Students were allocated to *groups of peers* in which they were matched with peers of their respective treatments. In each *group of peers*, half of the peers are of the same type as the student and the other half of the peers are of the type of her assigned treatment.

For exposition, consider the simple case of two types of students: high and low. The researcher is interested in identifying the Average Treatment Effect (ATE) of having high type peers. With two types of students there are three *groups of peers*: two homogenous groups, composed of individuals of a single type, and a heterogeneous group composed of individuals of both types. The following matrix shows the composition of *groups of peers*:

	High	Low
High	Group A	Group B
Low	Group B	Group C

In this case, there are three potential *groups of peers*:

a) Group A: a group composed of the high type only.

b) Group B: a mixed group, in which half are high-type students and the other half are low-type students.

c) Group C: a group composed of the low type only.

Notice that, conditional on a student's type, she can be assigned to a homogenous group (Groups A and C), with individuals of her own type, or to a mixed group (Group B), with individuals of both types.

To illustrate how this generates systematic variation across treatments, compare a high-type student in Group A versus a high-type student in Group B. In Group A, all peers are high types, while in Group B half of the peers are high types and the other half are low types. Hence, the difference in the proportion of high-type peers in Group A versus B is equal to 0.5. The Conditional Average Treatment Effect (CATE) of having high-type peers conditional on being a high-type student ($\tau_i = H$)can be identified by the difference between high-type students in group A and high-type students in Group B.

$$CATE_H = \mathbf{E}\left[Y_i | \tau_i = H, A\right] - \mathbf{E}\left[Y_i | \tau_i = H, B\right] \tag{1}$$

Similarly, consider a low-type student in Group B versus a low-type student in Group C. In Group B, half the peers are high types and the other half are low types, while in Group C all peers are low-type students. Hence, the difference in the proportion of high-type peers in Group B versus C is equal to 0.5. The Conditional Average Treatment Effect (CATE) of having high-type peers conditional on being a low-type student ($\tau_i = L$) can be identified by the difference between low-type students in group B and low-type students in Group C.

$$CATE_L = \mathbf{E}\left[Y_i | \tau_i = L, B\right] - \mathbf{E}\left[Y_i | \tau_i = L, C\right] \tag{2}$$

Considering the above, the average treatment effect of high-type peers is a weighted average of the CATE in equations 1 and 2, where weights capture the proportion of high type and low type students in the data, respectively. Since I am using the cell-specific median to classify students, the weights are equal.

$$ATE = 0.5 * CATE_H + 0.5 * CATE_L \tag{3}$$

Notice that the statistical power to estimate this average effect is maximized when all groups of peers —Groups A, B and C— are of the same size. The number of students who are treated (high-type students in Group A and low-type students in Group B) and the number of students who are not treated (high-type students in Group B and low-type students in Group C) will be the same.

The fact that all three groups are the same size implies that students are twice as likely to be assigned to peers of their same type. Hence, high-type students are twice as likely to receive the treatment (high-type peers) than low-type students. Given that the propensity score of receiving the treatment will vary by student type, we need to account for this in the empirical analysis.

The randomization in my field experiment is analogous to this example, with just one caveat. In my randomization I use two treatments instead of one, so rather than two types of students, I have four types: (i) more sociable and higher achieving, (ii) more sociable and lower achieving, (iii) less sociable and higher achieving, and (iv) less sociable and lower achieving. This implies that instead of the three groups of peers A, B and C from my previous example, there are ten potential *groups of peers* in my experimental strategy.[6]

Figure 2 shows the ten possible combinations of types of peers and student types. Each row corresponds to the student type, each column to the type of peer to whom she was assigned, and each cell to the combination of a student type-type of peer or *group of peers*.[7] Each group takes a different cell color in the symmetrical matrix of Figure 2.

I performed the randomization stratifying at the school-by-grade-by-gender level and student type level. The first stratification (school-by-grade-by-gender) is performed because the allocation to dormitories is specific to these strata. The second stratification (student type) is necessary because students were assigned to *groups of peers* based on their type as identified in the classification discussed above.

3.3.2 Assigning Students to Dormitories

This subsection describes how I implement my experimental strategy. I randomized students into *groups of peers*, and use these groups to allocate students to the dormitories in the COAR Network.

There is vast heterogeneity in the structure of dormitories across the COAR Network. For example, while the school in Lima has dormitories of three to five students, its counterpart in Cusco has a total of four dormitories, with approximately 80 students per dormitory.[8] To reconcile my *groups of peers* with the widely varying number of roommates across schools, I sorted the names of the students on a list based on the 10 *groups of peers* mentioned in the previous subsection. This list was later used to allocate students to dormitories. The *groups of peers* were randomly ordered on the list[9], and for mixed groups composed of more than one student type, the names of students of different types were alternated. Appendix B describes in detail how the lists determined the allocation to dormitories and classrooms.

The order on the list is directly linked to the physical distance between two students in a dormitory. Students who are adjacent on the list are more likely to be near each other in the dormitories. In small dorms, the assigned peers will likely share the same room. In bigger

[6]With 4 types of students there would be 16 possible combinations, but 6 of them are redundant.

[7]Group 1, for example, is composed of only more sociable and higher-achieving students. Group 3 is composed of less sociable and higher-achieving with more sociable and lower-achieving students.

[8]Appendix Figure A.1 shows a picture of the dormitories in the schools in Lima, Piura, and Cusco.

[9]The order was specific to each school$\times$grade$\times$gender.

dorms, students and assigned peers will be either placed in the same bunk bed or in beds next to each other. I used information from the school directors about the types of dorms available when creating my lists. Most of the schools (23 out of 25) in the COAR network used my lists to allocate students to dormitories. In some cases, the school directors sent the allocation they used, and I checked whether it was done based on my lists. There was not perfect compliance between the order of students on the list and the actual assignment to dormitories. For example, in some schools students were assigned to other beds for health reasons. Likewise, since there is a natural mismatch between the size of dormitories and the size of the groups of peers from my randomization, some students did not have their assigned peers as roommates.

4 Balance and First Stage

This section shows that the randomization is balanced in characteristics at baseline and that the experiment ensures substantial variation of peer characteristics across treatments. This variation translates into roommates with different academic skills and sociability at baseline. Furthermore, I also show that the intervention led to the formation of new friendships, influencing the social networks in the schools.

4.1 Balance of Baseline Characteristics

I use the following equation to estimate the correlation of the higher-achieving peers' treatment and the more sociable peers' treatment on students' outcomes and baseline characteristics:

$$y_{i\tau} = \alpha + \lambda_s s_{i\tau} + \lambda_c c_{i\tau} + \gamma_\tau + \nu_{i\tau} \tag{4}$$

Equation 4 explores how the treatment of more sociable peers, $s_{i\tau}$, and the treatment of higher-achieving peers, $c_{i\tau}$, correlate with the characteristic of individual i of type τ, $y_{i\tau}$. We include student type fixed effects, denoted by γ_τ since the propensity score of receiving the treatment varies by the student type. The parameters of interest are λ_s and λ_c, which represent the correlation of more sociable and higher-achieving peers, respectively.

In addition to the type fixed effect, all of my estimations control for the stratification variables of my randomization: the strata corresponds to cells by school-by-grade-by-gender-by-student type. Moreover, I control for the dependent variable at baseline to improve the efficiency of my estimates. All of my results are robust to an alternative specification in which baseline covariates are chosen based on the "post-double-selection" Lasso method developed by Belloni et al. (2014a,b). The standard errors are clustered at the student type$\times$*group of peer* level, since all the students within this unit share the same treatment peers.

For the 2017 cohort, I used a similar procedure to the one described in section 3.3.2 to assign students to classrooms. To exploit the same type of variation as with dorm assignments, I include a strata-by-classroom fixed effect for students in their first year when I estimate equation 4. The magnitude of peer effects from roommates could be different to the magnitude of peer effects from classmates. For example, evidence in the literature suggests that teachers change their behavior based on the composition of the classroom (Duflo et al., 2011). Hence, I make sure that the variation in peer characteristics is only in the sociability and academic achievement of roommates in the dormitories.

I estimate equation 4 on social and cognitive skills at baseline for all students, and for all subgroups of sociability and academic achievement, and present the balance tests for these variables in Table 2. The estimates reported in Table 2 show that the treatments are not correlated with social and cognitive skills at baseline. Furthermore, Tables A.3 and A.4 present balance tests on all other variables available at baseline. Overall, and as expected from an RCT, I do not reject a zero correlation of the treatments with baseline characteristics. The table also reports the F-statistic of multivariate regressions, evidencing how for both treatments and across all subgroups of students, treatments are not correlated with baseline characteristics.

4.2 First Stage

Next, I explore the impact of the randomization on the number of peers of each type and peer characteristics. First, I estimate equation 4 on the number of more sociable and higher-achieving peers in each group. Second, I estimate the impact of the treatments on average peer characteristics; this corresponds to the first stage and is depicted in equations 5a and 5b.

$$\bar{s}_{p_{i\tau}} = \theta_s + \delta_s s_{i\tau} + \phi_s c_{i\tau} + \gamma_\tau + \xi_{i\tau}, \tag{5a}$$

$$\bar{c}_{p_{i\tau}} = \theta_c + \delta_c s_{i\tau} + \phi_c c_{i\tau} + \gamma_\tau + \nu_{i\tau}, \tag{5b}$$

where δ_s and δ_c are the effects of the more sociable peers treatment on the average sociability and academic achievement of peers, respectively. Likewise, ϕ_s and ϕ_c represent the effects of the higher-achieving peers treatment on the same variables.

As expected from the randomization, the assignment to treatments leads to differences in the type of assigned peers. Table 3 reports the impact of the treatments on the type of peers that students have and on the average characteristics of these peers. Columns 1 and 2 of Table 3 show how each treatment changed the number of assigned more sociable and higher-achieving peers. In general, being assigned to more sociable peers increases the number of more sociable peers in a student's group by three, and equally as much with higher-achieving peers. That is, students have three additional peers associated with the type of treatment.

The design ensures substantial variation in the average characteristics of assigned peers. Columns 3 and 4 of Table 3 show the impact of the treatments on the average characteristics of the assigned peers. The more sociable peers treatment increases the average sociability of the assigned peers by 0.88 standard deviations. Likewise, the higher-achieving peers treatment increases the average academic achievement of the assigned peers by 0.92 standard deviations. The results also show that, because sociability and academic achievement are positively correlated at baseline, having higher-achieving peers increases a student's peers average sociability, and that having more sociable peers increases a student's peers average academic achievement.

In some cases there is no perfect compliance between the randomly assigned peers and actual roommates. Hence, I estimate the same set of equations on actual roommates rather than assigned peers in the groups. For small dormitories (less than 5 students), I defined roommates as peers in the same dormitory. For larger dormitories (more than 5 students), roommates are students in the same or in the adjacent bunk bed.

The data shows that the treatments predict roommates characteristics confirming that the schools followed the implementation procedures described in the previous section. The impact on the treatments on Columns 5 to 8 of Table 3 show the effect of each treatment on students' actual roommates. Columns 5 and 6 show the estimation of equation 4 on more sociable and higher-achieving roommates. Overall, each treatment, more sociable and higher-achieving peers, increases the number of roommates of their respective type by 1.6. Columns 7 and 8 show the effect on average roommate characteristics. Being assigned to more sociable peers increases the average sociability of roommates by 0.557 standard deviations. Likewise, the higher-achieving peers treatment increases the average academic achievement of roommates by 0.59 standard deviations. As expected, due to the non-compliance reasons mentioned above, these effects are weaker than those in columns 1 to 4 of Table 3 on assigned peers.

4.3 Social Interactions

After reviewing my intevention's impact on the assigned peers, I now analyze whether students became friends with their roommates. I show that the intervention had an influence on the social networks in the schools. I do this by showing that the intervention changed the average characteristics of friends, and it did because students formed new friendhsips with the peers near them in the dormitories.

I use social network data to show that the intervention affected the formation of new friendships. I administered two surveys with questions that measured social interactions after the intervention, as shown in the timeline in Figure 1. The first survey took place four months after the intervention, in August 2017. In this survey, students answered questions identifying their friends, study partners, and people with whom they engaged in social activities such as playing games or dancing. The second survey took place in

December 2017, using the same set of questions.[10] I then constructed a general network that aggregates the answers from both surveys.

To test whether the intervention had an impact on the social interactions within the schools, I estimate equation 4 on the number of friends of each type, and equations 5a and 5b on average friends' characteristics. Columns 9-12 of Table 3 present these results. Being assigned to more sociable peers increases in 0.483 the number of more sociable friends, and being assigned to higher-achieving peers in 0.414 the number of higher-achieving friends. These translate on an increase of 0.063 and 0.058 standard deviations the average sociability and academic achievement of friends, respectively. All of these effects are statistically significant at the 1% level.

In addition, I also study how the order on the lists affected the social interactions among students. I do this by estimating the following equation:

$$l_{ij} = \gamma_0 + \sum_{k=1}^{9} \gamma_k \mathbf{1}_{d=k_{ij}} + \nu_{ij}. \tag{6}$$

Equation 6 describes how dummy variables ($\mathbf{1}_{d=k_{ij}}$), denoting the distance between students i and j on the list change the likelihood of a link (l_{ij}) between i and j. The equation includes nine dummy variables, each of which represents a distance of 1–9 on the list.

The distance between students on the list predicts the formation of social links. Panel A of Figures 3 and 4 shows how the distance on the list predicts the likelihood that two students will form a social connection out of all the students. Figure 3 shows the estimation of equation 6 by the more sociable treatment status, and Figure 4 by the higher-achieving peers treatment status. The plots show the estimates of γ_k with the respective 95% confidence interval. Being the neighbor of a student on the list increases the likelihood of becoming friends, engaging in social activities together, or studying together by approximately 18 percentage points. Furthermore, there is a decreasing pattern in the distance on the list, showing that the physical distance in the allocation to dormitories has the power to predict social interactions.

I find no evidence of heterogeneous effects of distance dummies on social interactions by the more sociable peers or higher-achieving peers treatments. The plots in Panel A of Figures 3 and 4 show this clearly since the blue and purple bars—which denote the control and treatment groups, respectively—look very similar.

I also test this formally by estimating whether there are heterogeneous treatment effects of being neighbors on the list by treatments. Students i and j are neighbors on the list if their names are adjacent, which is equivalent to a distance d_{ij} equal to 1. I estimate

[10] In addition to the questions in the first survey, students also answered questions related to collaboration between them and diffusion of information in the second round. The questions on cooperation asked from whom did they receive help (and who did they help) with their studies and personal problems. Students also named up to five peers who the ministry should contact to diffuse academic or cultural information during the vacation period.

equation 7, which captures how the likelihood that individuals i and j will have a social interaction l_{ij} is predicted by being neighbors ($neighbor_{ij}$), and neighbors' heterogeneity by the more sociable treatment, $s_{i\tau}$, and the higher-achieving peers treatment, $c_{i\tau}$.

$$\begin{aligned} l_{ij} = \; & \gamma_0 + \gamma_1 neighbor_{ij} + \gamma_2 s_{i\tau} \times neighbor_{ij} + \gamma_3 c_{i\tau} \times neighbor_{ij} \\ & + \textstyle\sum_{k=2}^{9} \gamma_k \mathbf{1}_{d=k_{ij}} + \alpha_\tau + \mu_{ij\tau}, \end{aligned} \tag{7}$$

The parameters of interest in equation 7 are the impact of being neighbors, γ_1, and the differential effect of being neighbors by the more sociable peers treatment, γ_2, and the higher-achieving peers treatment, γ_3.

I find that being neighbors on the list has a substantial effect on the likelihood of forming social interactions. The estimates of equation 7 are reported in column 1 of Appendix Table A.5. If two students are neighbors on the list, this increases the likelihood that they will become friends, study together or engage in social activities together by 16.1 percentage points (p-value 0.000). I also find that the impact of being neighbors is greater when the student is assigned to more sociable peers. In particular, students are 3.7 percentage points (p-value 0.014) more likely to form a link with their neighbor when that peer is more sociable. In contrast, I do not find that there is a differentiated effect of being neighbors on social interactions when the student is assigned to higher-achieving peers.

For students assessed as less sociable at baseline, distance has the same effect on social interactions for those assigned to more and less sociable peers. Panel B of Figure 3 presents the estimates of equation 6 for less sociable students, and Panel C presents the results for more sociable students. Both plots show a similar pattern: distance on the list has a decreasing effect on the likelihood of forming social interactions. Furthermore, for less sociable students this pattern is similar regardless of whether they are assigned to less sociable or more sociable peers. By contrast, the plot in Panel C suggests that more sociable students are more likely to form social interactions with their neighbors when assigned to more sociable peers.

I derive the same conclusion by estimating equation 7 by subgroups of sociability at baseline. Columns 2 and 3 of Table A.5 present the estimation of equation 7 for less and more sociable students at baseline, respectively. The results are consistent with the discussion above. Less sociable students are 16.5 percentage points (p-value 0.000) more likely to form connections with their neighbors. There is no evidence of differentiated impacts of neighbors in different treatments. By contrast, more sociable students are 14.1 percentage points (p-value 0.000) more likely to form connections with their neighbors, 4.8 additional percentage points more likely when those neighbors are more sociable (p-value 0.022), and 3.5 additional percentage points more likely when those neighbors are higher achieving (p-value 0.106).

Next, I explore differences in the effect of distance by academic achievement at baseline. The evidence suggests that the level of interactions between lower-achieving students and higher-achieving peers is similar to the level with lower-achieving peers. Panel B of Figure 4 presents the estimates of equation 6 for lower-achieving students, and Panel C for higher-achieving students. Both plots show a decreasing effect of distance on the list on the likelihood of forming social interactions. For both subgroups, this pattern is similar regardless of being assigned to lower- or higher-achieving peers.

Similar evidence comes from the estimation of equation 7 by academic achievement at baseline. Columns 4 and 5 of Table A.5 present the estimation of equation 7 for lower- and higher-achieving students, respectively. Lower-achieving students are 17 percentage points (p-value 0.000) more likely to form connections with their neighbors, and there is no evidence of differentiated impacts of neighbors for either the more sociable or higher-achieving peers treatment. Higher-achieving students are 15.1 percentage points (p-value 0.000) more likely to form connections with their neighbors, and 7 additional percentage points (p-value 0.001) more likely when those neighbors are more sociable.

The fact that for lower-achieving students there are no differences in social interactions by peer assignment suggests that there is no sorting by students' academic achievement in the network formation. This opposed to previous evidence in the literature of peer effects. Carrell et al. (2013) argue that peer effects are negative because lower-achieving students interact among themselves, instead of connecting with higher-achieving peers (endogenous social networks). This is not true in my study.

A potential concern with this analysis is that the estimation has assumed link independence in the network formation. The most recent developments in the econometrics of networks address the correlation between linking decisions.[11] In Appendix E, I show that these results are robust to link dependencies.

In summary, this section concludes that the intervention had an influence on the network formation in schools. Students became friends with their peers near them in the dormitories and there is no evidence of of heterogenous effects of proximity on social interactions.

5 Outcomes and Empirical Strategy

5.1 Outcomes

In this section I describe the outcomes of my study. These are grouped into two broad categories according to the type of skill affected: social or cognitive. Social skills outcomes are measured using self-reported instruments and peers' perception of students according to different characteristics. Cognitive skills outcomes are measured by school grades and test scores collected by the Ministry of Education.

[11]See Chandrasekhar and Jackson (2016), Chandrasekhar (2016), de Paula Aureo et al. (2018), Graham (2017), Mele (2017a), Mele (2017b) for examples.

5.1.1 Social Skills Outcomes

The first set of outcomes corresponds to measures of social skills. Finding reliable measures of social skills is a big challenge. I use two categories of social skills outcomes: psychological self-reported tests and peers' perception measures.

My main outcome is expressed as a social skills index. It is constructed using the first component of a Principal Component Analysis (PCA) on the entire set of tests in this paper that measure social skills, including peers' perceptions. The psychological tests used for this index are described in Appendix C. The peers' perception measures capture the number of peers who report that the student is in the top five of four school-grade categories: leadership, friendliness, popularity, and shyness.

I reproduce my social skills index, with the available social skills measures at baseline. Appendix Figure A.3 displays a scatterplot between the two measures of students' social skills at baseline and after the intervention. There is a large, positive correlation between the two measures. An OLS regression shows that one standard deviation in the social skills index at baseline is correlated with a 0.43-standard-deviation increase in the social skills index after the intervention.

The most widely accepted taxonomy of psychological traits, both in the literature and in my data, is the Big Five (McCrae and John, 1992; John and Srivastava, 1999).[12] The *American Psychology Association Dictionary* defines the Big Five personality traits as follows (Table 1.1 in Almlund et al. (2011)):

1. Conscientiousness: the tendency to be organized, responsible, and hardworking.

2. Openness to Experience: the tendency to be open to new aesthetic, cultural, or intellectual experiences.

3. **Extraversion**: an orientation of one's interests and energies toward the outer world of people and things rather than the inner world of subjective experience; characterized by positive affect and sociability.

4. **Agreeableness**: the tendency to act in a cooperative, unselfish manner.

5. Neuroticism or Emotional Stability: Emotional Stability is "predictability and consistency in emotional reactions, with absence of rapid mood changes." Neuroticism is a chronic level of emotional instability and proneness to psychological distress.

Only two traits from the Big Five are associated with social skills: extraversion[13] and

[12] Almlund et al. (2011) summarizes the Big Five personality traits and their application to economics. Likewise, Akee et al. (2018); Donato et al. (2017); Kranton and Sanders (2017) provide recent evidence of the Big Five in economics research.

[13] The facets of extraversion correspond to: warmth (friendly), gregariousness (sociable), assertiveness (self-confident), activity (energetic), excitement seeking (adventurous), and positive emotions (enthusiastic).

agreeableness[14]. Empirical evidence shows that extraversion is associated with good labor market outcomes (Fletcher, 2013), and that agreeableness influences occupational decisions (Almlund et al., 2011; Cobb-Clark and Tan, 2011). These results are consistent with a study by Deming (2017) that concludes that the labor market increasingly rewards social skills.

In addition to the Big Five[15], the Ministry of Education collects other self-reported measures of social skills: altruism, empathy, emotional intelligence, intercultural sensitivity, and leadership. As part of the endline survey for this study, we also collected the *Reading the Mind in the Eyes* test. Other details for all of these tests are described in Appendix C.

While self-reported psychological tests are frequently used to measure social skills, they are subject to social desirability bias and can be manipulated by the respondent. Since social skills take full shape through interactions with peers, we also included questions of how peers perceive students.[16] This is additionally supported by empirical evidence which shows that relying on the perceptions of members of the same community relaxes information asymmetries (Hussam et al., 2017).

Students were asked to rank up to five of their peers in the four dimensions of leadership, friendliness, popularity, and shyness. I constructed a measure of peers' perception by adding the number of peers who name a given student in each of the four dimensions. The perception measure for individual i corresponds to the number of peers who believe that subject i is in the top five of characteristic c in their school-grade. There is a positive correlation between the number of peers who rank the student in the top five on leadership, friendliness, and popularity, and a negative correlation with shyness.

5.1.2 Cognitive Outcomes

Teachers assign grades to students for each subject based on their homework and test scores during the first three quarters of the year. Although these variables are available for the three cohorts, the 2015 cohort reports discrete grades with limited variation, so my empirical analysis focuses on the grades of the 2016–17 cohorts. While for the 2016–17 cohorts grades ranged between 1 and 20, for the 2015 cohort students received IB grades between 1 and 7; 83% of the students in the latter cohort obtained a grade between 3 and 5 for math, and 96% achieved a grade between 4 and 6 for Spanish. Due to this small variation, the analysis of grades focuses on the 2016–17 cohorts only.

Students in the 2016–17 cohorts were also assessed via standardized tests designed by the Ministry of Education. These tests determine the students' grades for their final

[14]The facets of agreeableness are: trust (forgiving), straight-forwardness (not demanding), altruism (warm), compliance (not stubborn), modesty (not show-off), tender-mindedness (sympathetic).

[15]The ministry implements the translation of the questionnaire developed by Goldberg (1999) into Spanish found in Cupani (2009). The English version is available at the following link: https://ipip.ori.org/New_IPIP-50-item-scale.htm. The Spanish version of this test is available upon request.

[16]This was also the case in the baseline survey, as described in Appendix A.

quarter at school. For the 2015 cohort, these test scores are not available; the Ministry used the IB grades instead. Hence, the study of cognitive outcomes focuses on the 2016–17 cohorts exclusively.

As described in section 3, the more sociable peers treatment is only available for the 2015–16 cohorts. Likewise, test scores and grades are only available for the 2016–17 cohorts. Appendix Table A.8 reconciles both sets of information, and indicates which cohorts were used for each treatment–outcome combination.

5.2 Empirical Strategy

I begin by estimating the effect of my two treatments—more sociable and higher-achieving peers—on the social and cognitive skills outcomes described in section 5.1. The following equation estimates the impact of each treatment:

$$y_{i\tau} = \alpha + \lambda_s s_{i\tau} + \lambda_c c_{i\tau} + \gamma_\tau + \varepsilon_{i\tau}. \tag{8}$$

Equation 8 shows how the more sociable peers treatment, $s_{i\tau}$, and the higher-achieving peers treatment, $c_{i\tau}$, affect the outcome, $y_{i\tau}$, of individual i of student type τ. I include student type fixed effects, γ_τ, because the likelihood of receiving the treatments varies by student type. The parameters of interest in this equation, λ_s and λ_c, the causal impact of the more sociable and higher-achieving peers treatments, respectively. I include the same set of controls as the ones in the balance tests of section 4. The standard errors are clustered at the student type$\times$*group of peer* level, since all the students within this unit share the same treatment peers. I also report the randomization inference p-values for my main results (Athey and Imbens, 2017; Young, 2017).

Estimates of equation 8 and equations 5a and 5b are of independent interest. They also are the Reduced Form and the First Stage of an IV estimate of the effect of peers' abilities. I estimate the effect of a one-standard-deviation in peers' average characteristics (i.e. roommates' sociability and academic achievement) on students' outcomes. I use the experimental variation in my study in a two-endogenous model, and jointly estimate the effect of peers' characteristics on the cognitive and social outcomes of students. The following equation introduces my two-endogenous model:

$$y_{i\tau} = \theta + \beta_s \overline{s}_{r_{i\tau}} + \beta_c \overline{c}_{r_{i\tau}} + \gamma_\tau + \varepsilon_{i\tau}, \tag{9}$$

where $\overline{s}_{r_{i\tau}}$ and $\overline{c}_{r_{i\tau}}$ denote the average baseline sociability and academic achievement $r$of student i of type τ. For small dormitories (less than 5 students), I defined roommates as peers in the same room. For larger dormitories (more than 5 students), roommates are defined as having the same or adjacent bunk bed. The parameters of interest are β_s and β_c; the effect of a one standard deviation in the average sociability and academic achievement of roommates on students' outcomes. The first stage of this model is depicted in equations 5a and 5b. It represents the impact of the assignment to treatment on peer characteristics.

As described in section 4, Columns 7 and 8 of Table 3 display the estimates of equations 5a and 5b. Being assigned to live with more sociable peers increases the average sociability of roommates by 0.55 standard deviations, and the higher-achieving peers treatment increases the average academic achievement of roommates by 0.59 standard deviations.

6 Main Results

This section describes the impact of the intervention on students' cognitive and social skills outcomes. I first present the reduced-form estimates on social skills, then the reduced-form estimates on cognitive skills. Finally, I present the 2SLS estimates using the experimental variation as an instrument for roommates' sociability and academic achievement at baseline.

6.1 Social Skills Outcomes

My description of the results starts by reporting the impact of my two treatments—the more sociable peers treatment and the higher-achieving peers treatment—on personality traits, peers' perception, and my social skills index. Panel A of Table 4 reports the reduced-form estimates of equation 8 for all students on all of my social skills indicators.

The results reveal that having more sociable peers has positive effects on openness, extraversion, and agreeableness of the Big Five personality traits. Columns 1 to 5 show the impact of both treatments on the Big Five: openness, conscientiousness, emotional stability, extraversion, and agreeableness. I focus on the last two traits, which are directly related to social skills. I find that the effect of more sociable peers on extraversion and agreeableness are 0.067 (p-value 0.029) and 0.066 (p-value 0.035) standard deviations, respectively. The higher-achieving peers treatment has no effect on the Big Five.

I find no evidence that either the more sociable or the higher-achieving peers treatment affects how peers perceive the students. Columns 6 to 9 show the treatment effects on peers' perception measure for social skills outcomes. The dependent variable in each column is the number of peers who think a student is in the top five of leadership, friendliness, popularity, and shyness at the school-by-grade level. Overall, I cannot reject that either the more sociable peers treatment or the higher-achieving peers treatment do not affect the peers' perceptions of a student.

The more sociable peers treatment has a similar positive impact on other measures of social skills and the social skills index. Column 10 displays the regression results for an index composed of other measures of social skills, which are described in Appendix C. By contrast, higher-achieving peers do not affect other measures of social skills. Column 11 of Table 4 shows the impact of my treatments on my main social skills outcome—the social skills index. Overall, the more sociable peers treatment has a positive impact on the social skills index. In particular, there is a treatment effect of 0.067 standard deviations (p-value 0.016), and it is possible to reject that this effect is equal to zero. Also consistent with my regressions for other social skills outcomes, I cannot rule out the possibility that the higher-achieving peers treatment does not affect the social skills index.

Table 4 additionally reports the Randomization Inference (RI) p-values of my regressions. The null hypothesis for this test is that the outcomes for all units in the sample would have been the same, regardless of whether the units received the treatment or control status. Thus, the null hypothesis implies that all students would have the same degree of social skills, regardless of the sociability of their peers. To compute the RI p-values, the treatments of more sociable peers and higher-achieving peers were randomly reassigned, 1,000 times, using the same stratification criteria as the original assignment. I estimate equation 8 for each of these 1,000 permutations. Then I compare the distribution of the coefficients that were induced by reassignment with the corresponding coefficients, $\hat{\lambda}_s$ and $\hat{\lambda}_c$, of the real assignment, and produce my RI p-values.

The RI p-values are generally consistent with the sampling inference p-values. Overall, I reject the null hypothesis for the more sociable peers treatment on the extraversion and agreeableness traits. I arrive at the same conclusion for my general social skills index in column 11. Hence, the results in Panel A of Table 4 suggest that more sociable peers have a positive impact on the formation of students' social skills. By contrast, there is no evidence that higher-achieving peers affect cognitive skills.

Next, I explore whether there are heterogeneous treatment effects according to students' degree of sociability at baseline. To do this I estimate equation 8 by subgroups: less vs. more sociable students at baseline (Panels B and C, respectively). I then compare the results for my subgroups to the estimates of equation 8 for all students, presented in Panel A.

The positive effects of the more sociable peers on the social skills of their roommates are mostly driven by the effect on the less sociable students. Comparing the results in Panels A and B of Table 4 shows that most of the positive effects of more sociable peers on social skills are driven by the impact on the students assessed as less sociable at baseline. In particular, more sociable peers have a positive impact on extraversion of 0.084 (p-value 0.067) standard deviations (column 4), and on agreeableness of 0.123 (p-value 0.008) standard deviations (column 5).

I also find that less sociable students assigned to more sociable peers are perceived to be more friendly and popular. Less sociable students are perceived to be friendlier by 0.062 (p-value 0.024) standard deviations (column 7) and more popular by 0.042 (p-value 0.024) standard deviations (column 8). Both effects are statistically significant at the 95% level. These effects amount to an increase of 0.25 and 0.30, accordingly, in the number of peers who perceive students as friendly and popular, which represents a respective 13% and 20% increase over the average at baseline. By contrast, higher-achieving peers decrease the number of peers who perceive the student as friendly and popular.

Likewise, the more sociable peers treatment has a larger effect on other social skills and the social skills index for less sociable students than for all students. Column 10 shows that the more sociable peers treatment increases other social skills measures for less sociable students by 0.085 standard deviations (p-value 0.029), compared to 0.048 standard deviations for the average student. The same pattern is observed for the social skills index (column 11)—an increase of 0.114 standard deviations (p-value 0.004) for less sociable students versus 0.066 standard deviations for average students. Additionally, the RI p-values (all < 0.1) support the general conclusion that more sociable peers have a particularly positive effect on the social skills of less sociable students. The effect of the higher-achieving peers treatment on both variables is a precise zero.

The more sociable peers treatment does not affect the formation of social skills for students assessed as more sociable at baseline. Panel C supports this general conclusion by showing the reverse side of the story. I cannot reject the possibility of zero treatment effects for most of the outcomes in this table.

Thus, more sociable peers have a positive impact on the formation of social skills. These impacts are driven by the effects on less sociable students. This reveals how less sociable students benefit from being assigned to more sociable peers.

6.2 Cognitive Skills Outcomes

Table 5 reports the cognitive skills outcomes based on the estimation of equation 8. Columns 1 and 2 report the treatment effects for the grades in math and reading comprehension. Analogously, columns 3 and 4 show the impact of each treatment on math and reading test scores.

Consistent with the peer effects estimates reported by quasi-experimental studies (Angrist and Lang, 2004; Duflo et al., 2011; Abdulkadiroğlu et al., 2014) that generate large variation in peers' skills, I find that the impact of higher-achieving peers on students' academic achievement is a precise zero. Panel A of Table 5 presents the cognitive skills outcomes for all students in my sample. This is a narrowly measured estimate in the context of my study. The 95% confidence interval for math test scores ranges between -0.06 and 0.01 standard deviations. For reading, it ranges between -0.07 and 0.02 standard deviations. Likewise, I do not find evidence that having more sociable peers affects the formation of cognitive skills.

Next, I examine treatment effect heterogeneity for the cognitive skills outcomes. I estimate equation 8 for two subgroups of academic achievement: lower- and higher-achieving students. Panels B and C of Table 5 report the reduced-form estimates for lower- and higher-achieving students at baseline.

I find that higher-achieving peers have heterogeneous treatment effects on the formation of cognitive skills. Columns 1 and 2 in Panel B of Table 5 show that the higher-achieving peers treatment has a negative effect on grades for both math and reading comprehension. I first analyze the heterogeneous effects on grades, and find that higher-achieving peers decrease students' math grades by 0.081 standard deviations (p-value 0.019), and reading grades by 0.049 standard deviations (p-value 0.201). I also explore whether there are heterogeneous effects on test scores. Columns 4 and 5 of Table 5 show that the effects of higher-achieving peers on lower-achieving students are negative and significant for both math (-0.049 with a p-value of 0.054) and reading comprehension (-0.068 with a p-value of 0.045). For the more sociable peers treatment, I arrive at the same conclusion for both subgroups: I find no evidence that more sociable peers affect the formation of cognitive skills. The heterogeneous effects that I find for both grades and test scores are consistent with the RI p-values.

Finally, I show that the effects of higher-achieving peers are indeed statistically different for higher- and lower- achieving students at baseline.

In summary, higher-achieving peers have, on average a zero effect on students' overall cognitive outcomes, but they are detrimental to the academic achievement of lower-achieving students. I find heterogeneous effects of the higher-achieving peers treatment when I divide my sample into lower- and higher-achieving students. I show that while the impact of this treatment on higher-achieving students is a precise zero, there is suggestive evidence of negative impacts on lower-achieving students.

6.3 2SLS Estimates

Table 6 presents the results of the 2SLS two-endogenous model described by equation 9 on social and cognitive skills. The table reports the estimates of parameters β_s and β_c, the impact of roommates' average sociability and academic achievement on students' outcomes. There are two endogenous variables: roommates' sociability and roommates' academic achievement (both calculated at baseline). I instrument for these variables using indicators for whether the student was assigned to the more sociable or the higher-achieving peers treatment. The table shows the estimation for five dependent variables: social skills index (column 1), math grades (column 2), reading comprehension grades (column 3), math test scores (column 4), and reading test scores (column 5).

I find that roommates' sociability has a positive impact on social skills, but no impact on cognitive outcomes. Panel A of Table 6 shows the results for all students. A one-standard-deviation increase in roommates' sociability has a 0.132-standard-deviation impact (p-value 0.009) on the social skills index for the average student (column 1). I cannot reject an effect equal to zero on grades or test scores (columns 2-5).

Nor can I reject that the academic achievement of roommates at baseline has a zero impact on social and cognitive skills outcomes. The 95% confidence interval of a one standard deviation in roommates' baseline academic achievement ranges between -0.11 and

0.02 for math test scores (Table 6 Panel A, column 4), and between -0.12 and 0.03 for reading test scores (Table 6 Panel A, column 5). This rules out the positive peer effects estimates of exploiting random allocation to dorms (between 0.06 and 0.12 standard deviations), and the large peer effects estimates of exploiting random allocation to large groups such as classrooms or squadrons (0.35-0.53 standard deviations).

The positive impact of roommates' sociability on social skills is driven by the effect on students assessed as less sociable at baseline. I explore heterogeneity in baseline sociability by reporting the estimates of equation 9 for less sociable students in Panel B of Table 6 and the estimates for more sociable students in Panel C. Consistent with the results in Table 4, a one-standard-deviation increase in roommates' average sociability at baseline increases the social skills index for less sociable students by 0.237 standard deviations (p-value 0.002). By contrast, the impact on more sociable students has a small point estimate, 0.031 , and it is not possible to reject that it is equal to zero (p-value 0.632). The estimates also show that for less sociable students, higher-achieving peers have a negative impact on math grades (Table 6 Panel B, column 2, p-value: 0.098) and more sociable peers have a negative impact on math test scores (Table 6 Panel B, column 4, p-value: 0.045). However, these effects are not consistent across the other outcomes. In a similar fashion, more sociable peers have a positive impact on reading comprehension grades for more sociable students (p-value 0.094).

The results also show that roommates' sociability has a positive impact on lower-achieving students' social skills. Panels D and E of Table 6 show the estimation of equation 9 by academic achievement at baseline. Panel D reports the estimates of β_s and β_c for lower-achieving students, and Panel E for higher-achieving students. A one-standard-deviation increase in roommates' sociability produces a 0.254-standard-deviation increase in the social skills index for lower-achieving students (p-value 0.001).

By contrast, there is suggestive evidence that roommates' academic achievement has a negative impact on test scores for lower-achieving students. A one standard deviation in roommates' academic achievement at baseline decreases math grades by 0.150 (p-value 0.018), reading grades by 0.093 (p-value 0.219), math test scores by 0.082 (p-value 0.082), and reading test scores by 0.122 standard deviations (p-value 0.040) for lower-achieving students. For higher-achieving students, there is no evidence that either roommates' sociability or academic achievement affects social or cognitive skills outcomes.

In summary, the conclusions of the 2SLS model are the same as those using the reduced-form estimates. Having more sociable roommates has a positive impact on social skills for less sociable students. Likewise, the evidence suggests that having higher-achieving roommates has a detrimental effect on the academic achievement of lower-achieving students.

7 Mechanism: Self-Confidence

Next, I explore potential mechanisms that might explain my results.

Recall that peer effects were negative for lower-achieving students. According to Carrell et al. (2013), peer effects are negative because lower-achieving students interact amongst themselves, instead of connecting with higher-achieving peers (endogenous social networks). I test this hypothesis as a mechanism that can explain the results in my study and explore other potential mechanisms that might describe my findings. Whatever mechanism is operating in the backend of my intervention, it must be able to explain two main outcomes from my study:

1. More sociable peers have a positive effect on the formation of social skills, particularly for less sociable students.

2. Higher-achieving peers have a negative effect on academic achievement for lower-achieving students, and no effect on average.

Moving back to subsection 5.1.1, I reviewed how the intervention impacted the network formation and the social interactions that students develop. However, some of the evidence presented there also rules out that the number of social interactions between students and their peers explains my findings. More precisely, I find no evidence that less sociable students have more interactions with more —rather than less— sociable peers , or that lower-achieving students are interacting or studying less with higher-achieving peers. The evidence indicating that lower-achieving students befriend higher-achieving peers and still have lower academic achievement contradicts the hypothesis introduced by Carrell et al. (2013). In my study, despite lower-achieving students interact with higher-achieving peers, they still experience declines in academic achievement.

In this section, I study whether changes in the self-confidence of students in their skills can explain my results. The main idea behind this mechanism is that students update their beliefs about their social and cognitive skills by interacting with their peers. First, I address the self-confidence story through the lens of the confidence-enhanced performance model from Compte and Postlewaite (2004). I show that my empirical findings are consistent with students: (1) succeeding in social activities with peers when at least one of them is sociable, and (2) failing in cognitive activities when they perform worse than their peers. Furthermore, I provide empirical evidence suggesting that the intervention changed students' self-confidence in their skills.

7.1 Theoretical Framework

7.1.1 General Model

The model that I describe next is based on Compte and Postlewaite (2004). It is the simplest unified framework that can yield the empirical findings from this study. The model addresses how peer effects can shift students' beliefs about their skills—defined as self-confidence— and eventually transform students' outcomes. All proofs are provided in Appendix D.

In the model, the agent faces a sequence of decisions on whether or not to invest time in an activity. There are two types of activities: r and o. Undertaking an activity entails a cost c. This cost is stochastic with support in $[0, 1]$. The random variables $\{c_t\}_{t=1}^{\infty}$ are independent, and when the agent makes a decision at t, she knows the realization of c_t and the type of the activity r or o.

In this theoretical framework, success in both types of activities depends on the self-confidence of the agent, and her self-confidence depends on her perception of past successes. Let ρ_a denote the probability of success in an activity of type a, for $a = \{r, o\}$, and $\kappa \in (0, 1]$ as a measure of the agent's self-confidence. The model assumes that:

$$\rho_a = \kappa\mu_a,$$

where μ_a is an exogenous parameter—the probability of success in an activity of type a, that does not depend on the agent's self-confidence.

The level of self-confidence κ is a function of the empirical frequency of past successes ϕ in activities r and o:

$$\kappa = \kappa(\phi),$$

$$\phi = \gamma\phi_r + (1 - \gamma)\phi_o,$$

where $\phi_a = \frac{s_a}{s_a+f_a}$ is the frequency of success in activities of type a, s_a denotes the number of successes and f_a the number of failures. The parameter γ is the weight the agent assigns to the r activities in her past successes. Intuitively, γ is the proportion of the r activities that the agent faces. I will assume that $\gamma \in (0, 1)$, so that the agent decides on both types of activities.

The model assumes that κ is a smooth and increasing function of ϕ: $\kappa = \kappa(\phi)$, where $\kappa'(\cdot) > 0$, $\kappa(0) > 0$, and, without loss of generality $\kappa(1) = 1$. Combining the two functions above, the probability of success in an activity is a function of the perception of success ϕ, where:

$$\rho_a(\phi) = \kappa(\phi)\mu_a, \text{for } a = \{r, o\}.$$

Finally, the model assumes that the effect of self-confindence on performance cannot be too strong, that is, $\rho_a'(\phi) < 1$.[17]

An equilibrium in this model implies that for each type of activity, the probability of success coincides with the agent's perception of the frequency of success. Among the possible perceptions ϕ that an agent may have about the frequency of success, the one

[17]This assumption guarantees an interior solution for the level of self-confidence in equilibrium.

that plays a focal role is the unique perception ϕ^*, such that:

$$\phi_r^* = \kappa(\phi^*)\mu_r, \tag{10a}$$

$$\phi_o^* = \kappa(\phi^*)\mu_o, \tag{10b}$$

where $\phi^* = \gamma\phi_r^* + (1-\gamma)\phi_o^*$.

The agent assesses whether it is worth investing in an activity of type a and forms a belief p_a on whether or not she will succeed. With a belief p_{a_t} about her chance of success at time t, the agent compares the expected payoff from undertaking the activity to its cost. The agent only undertakes the activity a if and only if:

$$p_{a_t} \geq c_t.$$

The belief of success is a function of the data that she recollects: $p_a = \beta(s_a, f_a)$. This function can be interpreted as an agent who is initially unsure of the probability of success in activities a and updates her initial beliefs as she acumulates experience. The restrictions on the function β are the following:

(i) $\forall s, f \geq 0$, $0 < \beta(s_a, f_a) < 1$.

(ii) There exists $A > 0$ such that $\forall s_a, f_a > 0$, $|\beta(s_a, f_a) - s_a(s_a + f_a)| \leq \frac{A}{s_a+f_a}$.

The first assumption states that beliefs must lie between 0 and 1. The second assumption is for "asymptotic consistency", ruling out belief formation processes for which there is a permanent divergence between the agent's perceived successes and failures and her beliefs.

Under these conditions, beliefs are correct in the long run, and $\lim_{t\to\infty} p_{a_t} = \rho_a(\phi^*)$. The agent invests in an activity in the long-run when $\rho_a(\phi^*) \geq 0$, and this probability is given by:

$$Pr(c \geq \rho_a(\phi^*)) = \int_0^{\rho_a(\phi^*)} g(c)dc = G(\rho_a(\phi^*)),$$

where $G(\cdot)$ corresponds to the c.d.f. of the cost parameter c. I assume that $G(\cdot)$ has no flat regions.

To understand the impact of the intervention on students' self-confidence and their subsequent changes in skills, I will argue that the intervention changed the parameter μ_r: the probability of success that doesn't depend on the self-confidence of the individual in activities r. This parameter affects students' self-confidence and the likelihood of investing in activities r through two different channels: (1) a direct channel, whereby the agent exogenously becomes more likely to succeed, and (2) an indirect channel, whereby she gains self-confidence in her skills and this translates into a larger probability of success in equilibrium. Hence, through self-confidence, an increase in μ_r amplifies the probability of success for the agent in activities r. We can observe these effects by calculating the derivative of equation 10a with respect to μ_r:

$$\frac{d\phi_r^*}{d\mu_r} = \underbrace{\kappa(\cdot)}_{\text{direct channel}} + \underbrace{\kappa'(\cdot)\mu_r \frac{d\phi^*}{d\mu_r}}_{\text{indirect channel}} \tag{11}$$

Equation 11 reveals the the two channels through which changes in μ_r can affect the probability of success in equilibrium ϕ_r^*: the direct channel is given by changes in μ_r, and the indirect channel—where I choose to focus my attention—is determined by the increase in self-confidence $\kappa'(\cdot)$ of an individual in equilibrium. Both channels jointly explain how increasing the parameter μ_r leads to an agent's higher success in activities r in equlibrium.

The changes in μ_r also affect the probability of success in other activities through self-confidence. More specifically, they affect the probability of success of the second type of activity in this model, activities of type o. We can observe this by applying the implicit funtion theorem to 10b:

$$\frac{d\phi_o^*}{d\mu_r} = \mu_o \underbrace{\kappa'(\cdot)\frac{d\phi^*}{d\mu_r}}_{\text{self-confidence}} \tag{12}$$

The overall changes in the probability of success in both types of activities, accounting for the fact that $\frac{d\phi^*}{d\mu_r}$ is also determined in equilibrium, is summarized as follows:

Proposition 1. Under the assumptions that $\kappa'(\cdot) > 0$, $\kappa(0) > 0$, $\rho_a'(\phi) < 1$, and $\gamma \in (0,1)$, then:

1. The probability of success in activities r in equilibrium ϕ_r^* increases with μ_r.
2. The probability of success in activities o in equilibrium ϕ_o^* increases with μ_r.
3. The level of self-confidence in equilibrium $\kappa(\phi^*)$ increases with μ_r.

Corollary 1. *The probability that the agent invests in both types of activities in equilibrium is an increasing function of the parameter μ_r.*

Once the agent becomes more likely to succeed, she is willing to invest more frequently in equilibrium. There is a direct channel for activities r, and an indirect channel through which self-confidence influences an individual's success, both in activities of type r and o.

7.1.2 The Intervention

To understand this model in the context of my intervention, I will assume that activities r involve interactions with roommates, and activities o do not. The randomization changed the peer characteristics and hence generated variation in the parameter μ_r. In particular, a student's success in roommate activities depends on the characteristics of the student and her roommate. More precisely, the parameter μ_r is determined by the types of the student and her roommate. I consider only two types of individuals $i \in \{l, h\}$ and assume that:

$$\mu_r = \mu_r(\tau_i, \tau_j) \quad \text{for} \quad \tau_i,\ \tau_j = \{h, l\}.$$

where τ_i and τ_j are the types of individual i (the student) and j (the student's roommate), respectively.

Furthermore, I consider two types of activities: social and cognitive activities.

(a) Social activities form social skills. These are described as activities where students interact with peers by having a conversation or playing a game. Roommate activities involve interaction with roommates and other activities involve interaction with other people. Success in a social activity can be intuitively described by a student who enjoyed the interaction.

(b) Cognitive activities form cognitive skills. Cognitive activities involve the likes of studying for a test, attending class or receiving support from a tutor. Success can take different shapes for roommate and other activities. For example, success in other activities can occur when a student learns or when she receives a good grade. Alternatively, success in roommate activities is defined as a situation where students do not perform worse than their roommates. That is, I will assume that students will have a lower utility if they underperfom or learn less than their roommates: cognitive success in roommate activities is determined relatively, not absolutely.

I impose two additional assumptions on the likelihood of a student's success. The probability of success in roommate activities depends on student and roommate types as follows:

Assumption 1. ***Social success***: *The probability of success in a social roommate activity is larger when either the student or her roommate is high in social skills:* $\mu_r(l, l) < \mu_r(\cdot, h) = \mu_r(h, \cdot)$.

The intuition here is that a student is prone to feel good about participating in social activities with roommates when at least one of them is more sociable.

Assumption 2. ***Cognitive success***: *The probability of success in a cognitive roommate activity is larger when the student and her peer are of the same type, or the student is a high-type and her peer a low-type:* $\mu_r(l, h) < \mu_r(l, l) = \mu_r(h, h) = \mu_r(l, h)$.

The intuition behind this is that a student is more likely perform worse than her roommate when she is lower-achieving and her roommate is higher-achieving.

Proposition 2. Under the assumption of social success:

1) When assigned to more sociable peers, in equilibrium less sociable students:

   (a) Invest more in social activities with roommates.

(b) Invest more in other social activities.

(c) Increase their self-confidence in social skills.

2) More sociable students do not change their self-confidence or investment decisions on social skills.

The intuition behind this statement is that less sociable students benefit from interacting with more sociable peers. It is a direct result of proposition 1. Less-sociable students can more easily engage in conversations or other activities with more sociable —rather than less sociable— roommates. As a result, students feel more successful in social activities. The higher the success, the more the increase in students' self-confidence, which loops back into more successful activities with roommates as well as with other people. By contrast, the intervention did not change the probability of success of more sociable students in social roommate activities. Hence, the sociability of roommates does not affect the self-confidence of more sociable students. This rationalizes why more sociable students do not change their social skills under the lens of this model.

Proposition 3. Under the assumption of cognitive success:

1) When assigned to higher-achieving peers, in equilibrium lower-achieving students:

   (a) Invest less in cognitive activities with roommates.

   (b) Invest less in other cognitive activities.

   (c) Decrease their self-confidence in cognitive skills.

2) Higher-achieving students do not change their self-confidence or investment decisions on cognitive skills.

The intuition behind this statement is that lower-achieving students feel less accomplished when they compare their performance to their higher-achieving roomates. The lower the perceived success, the greater the frustration, and the lower the self-confidence, which brings down their investment in cognitive skills. The lower self-confidence loops back into less successful study interactions with their roommates as well as other learning activities. By contrast, the intervention did not change the probability of success for higher-achieving students in cognitive roommate activities. Hence, higher-achieving students did not experience changes in their academic achievement.

7.2 Empirical Evidence

In this section I present empirical evidence suggesting that self-confidence is a mechanism driving my results.

First, I study whether less sociable students feel they have better social interactions when assigned to more sociable peers. Through the lens of the model of the previous section, this can be interpreted as having more success in social activities. Table 7 presents these results on the following set of variables: whether the student feels supported by her friends (column 1), the level of satisfaction with her dormitory (column 2), whether the student thinks she is friends with her roommates (column 3), whether the student reports being empathetic to her roommates (column 4), whether the student reports that her roommates are empathetic to her (column 5)[18], whether the student reports that their roommates are worried for her problems (column 6), and a roommates' quality index that aggregates all these variables (column 7).

There is evidence suggesting that less sociable students perceive a higher quality of interactions with roommates when assigned to more sociable peers. Panels A and B of Table 7 present the results on social interactions quality for less sociable and more sociable students, respectively. The results show that less sociable students perceive a higher quality of social interactions when assigned to more sociable peers. They feel more supported by their friends, a higher level of satisfaction with their dormitory, and that their roommates are more empathetic to them. The results in column 6 show that the more sociable peers treatment has an impact of 0.345 standard deviations (p-value 0.010) on the quality index for less sociable students. By contrast, while more sociable students report a higher quality of interactions (Panel B, column 7), this difference is not statistically significant.

Second, I find some evidence that less sociable students changed their beliefs about their social skills due to the intervention. To support this claim, I estimate equation 8 based on whether the student named herself (weighted by the ranking) in the top five of leadership, friendliness, popularity, and shyness in her school-by-grade. Table 8 presents these results. Panels A and B show the results for students assessed as less sociable and more sociable at baseline. The estimates in Panel A show that except for leadership (column 1), the student is more likely to consider herself friendly (column 2), popular (column 3), and not shy (column 4) if she is assigned to more sociable peers. While some of these impacts are not significant, the effect on an index constructed by aggregating these three variables (column 5) is statistically significant at the 95% level. Including leadership in the index (column 6) reduces the size of the impact. However, it is still large in magnitude compared to the effects on the self-confidence of the more sociable students (Panel B of Table 8). This evidence suggests that the intervention increased less sociable students' self-confidence in their level of social skills.

I also estimate whether lower-achieving students changed their beliefs about their cognitive skills due to the higher-achieving peers treatment. The estimates of equation 8 for outcomes of self-confidence in cognitive skills are presented in Table 9. Panel A shows the results for the lower-achieving students, and Panel B for the higher-achieving students. The set of outcomes in the table are:

[18]Only a random sample of the students answered the two questions of empathy in the social network survey in August 2017.

- Belonging to the school (column 1): in the endline survey students indicated the degree to which they felt they belonged to the school (0 to 100).
- Academic ranking: students estimated their academic ranking in the school (column 2) and relative to their friends (column 3).
- Goals relative to peers: students answered the *Achievement Goal Questionnaire*.[19] Two of the factors of the test involve the student comparing herself to her peers: whether the student wants to perform well compared to other students (column 4), and whether the student avoids doing worse than other students (column 5).
- Academically skilled: whether the student considers herself one of the top five most academically skilled in the school, weighted by the ranking (column 6).
- Volunteering to receive information about scholarship and other academic opportunities, and the ranking the student gave herself for receiving this information (column 7).
- A self-confidence index that aggregates all the variables in columns 1 to 7.

My results show that lower-achieving students had lower levels of self-confidence in their cognitive skills when assigned to higher-achieving peers. The treatment effects of higher-achieving peers on the variables of self-confidence for lower-achieving students are presented in Panel A of Table 9. In column 1, I find that lower-achieving students decrease their belief that they belong to the school by 0.06 standard deviations (p-value 0.086) when assigned to higher-achieving peers. Since the COAR Network is designed for high-achieving students, the changes in beliefs are consistent with students feeling that they underachieve and are not part of the academic community because of it. Columns 2 and 3 show that these students also report having a lower academic ranking in the school and relative to their friends, although only the latter is statistically different from zero. Likewise, they are less likely to report that they want to perform better (column 4) or avoid doing worse than their peers (column 5). They are also less likely to name themselves among the top five most-skilled students (column 6), and to volunteer to receive information about scholarships and other academic opportunities (column 7). These results are all summarized by the effect on the self-confidence index in column 8; the higher-achieving peers treatment has a negative effect on self-confidence for lower-achieving roommates of 0.301 standard deviations (p-value 0.007). In comparison to the results in Panel B, the estimates for the higher-achieving peers in Panel C do not suggest that peer characteristics changed their beliefs about their skills.

Thus, my empirical evidence suggests that by interacting with more sociable peers, less sociable students gained self-confidence in their social skills. Similarly, by interacting with

[19]The details of the test are discussed in Appendix C.

higher-achieving, lower-achieving students lost self-confidence in their cognitive skills, which can explain why they have a lower academic performance.

In summary, this section provided a simple theoretical framework and empirical evidence that suggests that the type of peer had an impact on the students' self-confidence. Hence, I cannot rule out self-confidence as a mechanism driving my results on social and cognitive skills.

8 Conclusion

While there is a positive return to social skills in the labor market, there is still little evidence on how social skills form. This paper presents the results of a field experiment designed to estimate causal cognitive and social peer effects. The study was conducted in 23 out of 25 exam schools in Peru, covering a sample of approximately 6,000 students. Students were classified by baseline sociability and academic achievement using centrality measures of social networks and test scores. Unlike previous experimental designs, I have shown that it is possible to guarantee strong variation in peer characteristics by randomizing the type of peer and matching students to peers of their treatment groups.

I found that more sociable peers have a positive impact on the development of social skills. These effects are mainly driven by the impact on those who were assessed as less sociable at baseline. My results show that they became more extroverted and showed a higher level of agreeableness, the two components of the Big Five personality traits related to social skills. These effects are robust to other psychological tests that measure social skills and to measures that account for self-reported biases. Less sociable students with more sociable peers were perceived to be more friendly and popular in their school.

By contrast, I do reject positive cognitive peer effects on the formation of cognitive skills. For students who were lower achieving at baseline, the evidence suggests that the higher-achieving peers treatment has a negative effect. My rejection of positive cognitive peer effects is consistent with the evidence from quasi-experimental studies (Angrist and Lang, 2004; Abdulkadiroğlu et al., 2014; Rao, 2013). The findings contradict the positive peer effects of studies that use random allocation to groups (Sacerdote, 2011; Epple and Romano, 2011), highlighting the methodological concerns associated with this type of design, as pointed out by Angrist (2014).

I have shown that my results are consistent with the idea that students changed their self-confidence by interacting with their assigned peers. I presented a simple theoretical framework through which I can reproduce the empirical primary findings under reasonable assumptions for the model. The main intuition behind the model is that students do not know their level of skills but update their beliefs after interacting with their peers. Hence, students changed their investment decisions due to the updated beliefs about their success in social and cognitive activities. Furthermore, there is empirical evidence suggesting that self-confidence explains my results on the formation of social skills and academic achievement.

I also study the role of social interactions as a mechanism of how peer characteristics affect the formation of skills. I can rule out the level of interactions as a mechanism. While the intervention had a substantial impact on social network formation in the schools, I do not find that less sociable students interact more with more sociable peers.

Similarly, although lower-achieving students are befriending and studying with higher-achieving peers, they experience a decline in academic achievement. This result contradicts previous evidence in the literature which suggests students only benefit from higher-achieving peers when they are interacting with them (Carrell et al., 2013).

In conclusion, while the existing literature and education policies have focused on peers' academic achievement, this paper shows that in the setting of selective boarding schools in Peru, the role of sociable peers is more important. Further studies are needed to assess whether these results are valid in other contexts and the impact of these effects on labor market outcomes.

References

Abdulkadiroğlu, A., Angrist, J., and Pathak, P. (2014). The elite illusion: Achievement effects at boston and new york exam schools. *Econometrica*, 82(1):137–196.

Akee, R., Copeland, W., Costello, E. J., and Simeonova, E. (2018). How does household income affect child personality traits and behaviors? *American Economic Review*, 108(3):775–827.

Almlund, M., Duckworth, A. L., Heckman, J., and Kautz, T. (2011). Chapter 1 - personality psychology and economics1. In Hanushek, E. A., Machin, S., and Woessmann, L., editors, *Handbook of The Economics of Education*, volume 4 of *Handbook of the Economics of Education*, pages 1 – 181. Elsevier.

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics*, 30:98 – 108. Special Section articles on.

Angrist, J. D. and Lang, K. (2004). Does school integration generate peer effects? evidence from boston's metco program. *American Economic Review*, 94(5):1613–1634.

Arli, S. K. and Bakan, A. B. (2018). An investigation of the relationship between intercultural sensitivity and compassion in nurses. *International Journal of Intercultural Relations*, 63:38 – 42.

Athey, S. and Imbens, G. (2017). Chapter 3 - the econometrics of randomized experimentsa.

In Banerjee, A. V. and Duflo, E., editors, *Handbook of Field Experiments*, volume 1 of *Handbook of Economic Field Experiments*, pages 73 – 140. North-Holland.

Autor, D. (2014). Polanyi's paradox and the shape of employment growth. Working Paper 20485, National Bureau of Economic Research.

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

Banerjee, A., Chandrasekhar, A. G., Duflo, E., and Jackson, M. O. (2014). Gossip: Identifying central individuals in a social network. Working Paper 20422, National Bureau of Economic Research.

Beaman, L. and Dillon, A. (2018). Diffusion of agricultural information within social networks: Evidence on gender inequalities from mali. *Journal of Development Economics*, 133:147 – 161.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014a). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, 28(2):29–50.

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

Boisjoly, J., Duncan, G. J., Kremer, M., Levy, D. M., and Eccles, J. (2006). Empathy or antipathy? the impact of diversity. *American Economic Review*, 96(5):1890–1905.

Breza, E. and Chandrasekhar, A. (2018). Social networks, reputation, and commitment: Evidence from a savings monitors field experiment. Technical report.

Bursztyn, L., Egorov, G., and Jensen, R. (2018). Cool to be smart or smart to be cool? understanding peer pressure in education. *The Review of Economic Studies*, page rdy026.

Caeyers, B. and Fafchamps, M. (2016). Exclusion bias in the estimation of peer effects. Working Paper 22565, National Bureau of Economic Research.

Carre, A., Stefaniak, N., D'Ambrosio, F., Bensalah, L., and Besche, C. (2013). The basic empathy scale in adults (bes-a): Factor structure of a revised form. 25:679–691.

Carrell, S., Fullerton, R. L., and West, J. (2009). Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics*, 27(3):439–464.

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.

Castro-Solano, A. (2007). *Teoría y Evaluación dell Liderazgo*. Paidos.

Chandrasekhar, A. (2016). *The Oxford Handbook of the Economics of Networks*, chapter Econometrics of Network Formation. Oxford Handbook Series.

Chandrasekhar, A. and Jackson, M. O. (2016). A Network Formation Model Based on Subgraphs. *ArXiv e-prints*.

Chen, G.-M. and Starosta, W. J. (2000). The development and validation of the intercultural sensitivity scale. *Human Communication*.

Cobb-Clark, D. A. and Tan, M. (2011). Noncognitive skills, occupational attainment, and relative wages. *Labour Economics*, 18(1):1 – 13.

Colman, A. M. (2009). *A Dictionary of Psychology*.

Compte, O. and Postlewaite, A. (2004). Confidence-enhanced performance. *American Economic Review*, 94(5):1536–1557.

Cupani, M. (2009). El cuestionario de personalidad ipip-ffm: Resultados preliminares de una adaptación en una muestra de preadolescentes argentinos. 6:51–58.

de Paula Aureo, Seth, R.-S., and Elie, T. (2018). Identifying preferences in networks with bounded degree. *Econometrica*, 86(1):263–288.

Declerck, C. H. and Bogaert, S. (2008). Social value orientation: related to empathy and the ability to read the mind in the eyes. *The Journal of social psychology*, 148.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *Quarterly Journal of Economics*, 132(4):1593–1640.

Donato, K., Miller, G., Mohanan, M., Truskinovsky, Y., and Vera-Hernández, M. (2017). Personality traits and performance contracts: Evidence from a field experiment among maternity care providers in india. *American Economic Review*, 107(5):506–10.

Duflo, E., Dupas, P., and Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya. *American Economic Review*, 101(5):1739–74.

Epple, D. and Romano, R. E. (2011). Chapter 20 - peer effects in education: A survey of the theory and evidence. volume 1 of *Handbook of Social Economics*, pages 1053 – 1163. North-Holland.

Fabregas, R. (2017). A better school but a worse position? the effects of marginal school admissions in mexico city. Technical report, Harvard University. Job Market Paper.

Falk, A., Kosse, F., Deckers, T., Pinger, P., and Schildberg-Hörisch, H. (2018). The formation of prosociality: Causal evidence on the role of social environment. Technical report, University of Bonn.

Fletcher, J. M. (2013). The effects of personality traits on adult labor market outcomes: Evidence from siblings. *Journal of Economic Behavior & Organization*, 89:122 – 135.

Fryer, R. and Austen-Smith, D. (2005). An economic analysis of 'acting white'. *Quarterly Journal of Economics*.

Gambin, M. and Sharp, C. (2018). The relations between empathy, guilt, shame and depression in inpatient adolescents. *Journal of Affective Disorders*, 241:381 – 387.

Garlick, R. (2018). Academic peer effects with different group assignment policies: Residential tracking versus random assignment. *American Economic Journal: Applied Economics*, 10(3):345–69.

Goldberg, L. R. (1999). *Personality Psychology in Europe*, volume 7, chapter A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. Tilburg University Press.

Gordon, S. R. and Mwavita, M. (2018). Evaluating the international dimension in an undergraduate curriculum by assessing students' intercultural sensitivity. *Studies in Educational Evaluation*, 59:76 – 83.

Graham, B. S. (2017). An econometric model of network formation with degree heterogeneity. *Econometrica*, 85(4):1033–1063.

Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., Krivitsky, P. N., and Morris, M. (2018). *ERGM: Fit, Simulate and Diagnose Exponential-Family Models for Networks*.

Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., Krivitsky, P. N., Morris, M., and de Moll, S. B. (2016). Exponential random graph models (ergms) using statnet.

Hussam, R., Rigol, N., and Roth, B. (2017). Targeting high ability entrepreneurs using community information. Technical report, MIT M.

J. Elliot, A. and Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. 100.

John, O. P. and Srivastava, S. (1999). *Handbook of Personality: Theory and Research*, chapter The Big Five trait taxonomy: History, measurement, and theoretical perspectives, pages 102–138. Guilford Press, New York, second edition.

Jolliffe, D. and Farrington, D. P. (2006). Development and validation of the basic empathy scale. *Journal of Adolescence*, 29(4):589 – 611.

Kranton, R. E. and Sanders, S. G. (2017). Groupy versus non-groupy social preferences: Personality, region, and political party. *American Economic Review*, 107(5):65–69.

Lavy, V. and Sand, E. (2012). The friends factor: How students' social networks affect their academic achievement and well-being? Working Paper 18430, National Bureau of Economic Research.

Law, K., Wong, C.-S., and J Song, L. (2004). The construct and criterion related validity of emotional intelligence and its potential utility for management studies. 89:483–96.

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

Marsh, H. W. and Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology*, 47(1):213–231.

McCrae, R. R. and John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60(2):175–215.

Mehu, M., Grammer, K., and Dunbar, R. I. (2007). Smiles when sharing. *Evolution and Human Behavior*, 28(6):415 – 422.

Mele, A. (2017a). A structural model of dense network formation. *Econometrica*, 85(3):825–850.

Mele, A. (2017b). A structural model of homophily and clustering in social networks. Technical report, Johns Hopkins University.

Otto, P. E. and Bolle, F. (2011). Multiple facets of altruism and their influence on blood donation. *The Journal of Socio-Economics*, 40(5):558 – 563.

Rao, G. (2013). Familiarity does not breed contempt: Diversity, discrimination and generosity in delhi schools. Technical report, Department of Economics, Berkeley University.

Ribas, R., Sampaio, B., and Trevisan, G. (2018). Can better peers signal less success? the disruptive effect of perceived rank on career investment. Technical report.

Rushton, J. P., Chrisjohn, R. D., and Fekken, G. C. (1981). The altruistic personality and the self-report altruism scale. *Personality and Individual Differences*, 2(4):293 – 302.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly Journal of Economics*, 116(2):681–704.

Sacerdote, B. (2011). Chapter 4 - peer effects in education: How might they work, how big are they and how much do we know thus far? volume 3 of *Handbook of the Economics of Education*, pages 249 – 277. Elsevier.

Villadangos, M., Errasti, J., Amigo, I., Jollife, D., and García-Cueto, E. (2016). Characteristics of empathy in young people measured by the spanish validation of the basic empathy scale. *Psicothema*.

Wai, M. and Tiliopoulos, N. (2012). The affective and cognitive empathic nature of the dark triad of personality. *Personality and Individual Differences*, 52(7):794 – 799.

Weinberger, C. J. (2014). The increasing complementarity between cognitive and social skills. *The Review of Economics and Statistics*, 96(5):849–861.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., and Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, 330(6004):686–688.

Young, A. (2017). Channelling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. Technical report, London School of Economics.

Yukl, G. (2013). *Leadership in Organizations*. Pearson, eighth edition edition.

TABLE 1: Summary Statistics at Baseline

	All Students	By Sociability		By Academic Achievement	
		Less Sociable	More Sociable	Lower Achieving	Higher Achieving
Variable	(1)	(2)	(3)	(4)	(5)
Sociability (centrality measure)	-0.00	-0.64	0.64	-0.04	0.04
	(1.00)	(0.51)	(0.96)	(0.99)	(1.01)
Academic achievement (score in the admission test)	0.00	-0.08	0.08	-0.79	0.78
	(0.99)	(0.96)	(1.02)	(0.47)	(0.71)
Poor (%)	0.36	0.39	0.35	0.39	0.33
Extremely poor (%)	0.18	0.22	0.18	0.21	0.16
Rural (%)	0.26	0.31	0.23	0.30	0.22
N	6,147	1,832	1,822	3,069	3,078

Notes: This table reports summary statistics by type of student. Column 1 shows statistics for all students, and columns 2 to 5 according to the classification of students by social and cognitive skills. To classify students as more vs. less sociable, I use sociability at baseline as measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality measures a student's influence student in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Eigenvector centrality is not measured for the 2017 cohort. To classify students as higher vs. lower achieving, I use the score in the admissions test to the COAR Network, which evaluates the applicants in math and reading comprehension. For sociability and admissions test scores at baseline, standard deviations are reported in parentheses. The table includes a set of three demographic characteristics of students, including poverty, extreme poverty, and whether the student comes from a rural household. These demographic variables come from government administrative data.

TABLE 2: Balance on Cognitive and Social Skills at Baseline

	Social Skills Index			Math Score			Reading Score		
	All students	Sociability		All students	Academic Achievement		All students	Academic Achievement	
Variable		Less Sociable	More Sociable		Lower Achieving	Higher Achieving		Lower Achieving	Higher Achieving
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
More sociable peers	-0.041	-0.033	-0.050	0.019	0.019	0.019	0.010	0.022	-0.002
	(0.082)	(0.117)	(0.111)	(0.027)	(0.034)	(0.042)	(0.028)	(0.040)	(0.040)
Higher-achieving peers	0.048	0.130	0.003	0.016	-0.023	0.053	-0.000	-0.012	0.011
	(0.059)	(0.116)	(0.110)	(0.023)	(0.030)	(0.034)	(0.023)	(0.032)	(0.033)
Control mean	-0.05	-0.39	0.26	-0.11	-0.39	0.31	-0.08	-0.31	0.25
More sociable F p-value	0.89	0.57	0.91	0.89	0.82	0.92	0.89	0.82	0.92
Higher-achieving F p-value	0.73	0.77	0.49	0.73	0.80	0.40	0.73	0.80	0.40
N	6,103	1,832	1,822	4,553	2,274	2,279	4,553	2,274	2,279

Notes: This table reports balance checks of being assigned to more sociable and higher-achieving peers on social and cognitive skills outcomes for all students and subgroups by sociability and academic achievement at baseline. All regressions include strata fixed effects and control for the baseline value of the dependent variable. For the 2017 cohort all regressions include strata-by-classroom fixed effects. To classify students as more vs. less sociable, I use sociability at baseline as measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality measures a student's influence in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Eigenvector centrality is not measured for the 2017 cohort. To classify students as higher vs. lower achieving, I use the score in the admissions test to the COAR Network, which evaluates the applicants in math and reading comprehension. The control group is defined as being assigned to less sociable and lower-achieving peers. The "More sociable F p-value" and the "Higher-achieving F p-value" correspond to the F-statistic of the respective treatment of multivariate regressions that include all variables at baseline presented in Appendix Tables A.3 and A.4. Standard errors are clustered at the group of peers-by-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 3: Treatment Effects on Assigned Peers, Roommates, and Friends

	Assigned Peers				Roommates				Friends			
	Number		Baseline Characteristics		Number		Baseline Characteristics		Number		Baseline Characteristics	
	More sociable	Higher achieving	Sociability	Academic Achievement	More sociable	Higher achieving	Sociability	Academic Achievement	More sociable	Higher achieving	Sociability	Academic Achievement
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
More sociable peers	3.179***	-0.013	0.883***	0.088***	1.643***	0.020	0.557***	0.092***	0.483***	-0.004	0.063***	-0.010
	(0.107)	(0.106)	(0.016)	(0.017)	(0.051)	(0.049)	(0.019)	(0.020)	(0.108)	(0.096)	(0.011)	(0.010)
Higher-achieving peers	0.015	2.955***	0.039***	0.941***	-0.048	1.620***	0.020	0.590***	-0.109	0.414***	-0.004	0.058***
	(0.073)	(0.083)	(0.011)	(0.014)	(0.036)	(0.039)	(0.014)	(0.016)	(0.080)	(0.083)	(0.008)	(0.008)
Control mean	0.38	0.98	-0.22	-0.53	0.60	1.39	-0.14	-0.36	2.83	6.34	0.03	-0.06
N	6,103	6,103	6,103	6,103	6,103	6,103	6,103	6,103	6,103	6,103	6,103	6,103

Notes: This table reports the effect of being assigned to more sociable and higher-achieving peers identified at baseline on the number of more sociable and higher-achieving assigned peers, roommates, and friends, and on the average sociability and academic achievement for each of these groups. Assigned peers are students in the *groups of peers* to which the student was assigned, roommates are students in the same dormitory for small room and neighbors—students in the same or adjacent bunk bed—for large rooms. All regressions include strata fixed effects and control for the baseline value of the dependent variable. For the 2017 cohort all regressions include strata-by-classroom fixed effects. To classify students as more vs. less sociable, I use sociability at baseline as measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality measures a student's influence in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Eigenvector centrality is not measured for the 2017 cohort. To classify students as higher vs. lower achieving, I use the score in the admissions test to the COAR Network, which evaluates the applicants in math and reading comprehension. The control group is defined as being assigned to less sociable and lower-achieving peers. Standard errors are clustered at the group of peers-by-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 4: Reduced-Form Estimates on Social Skills

	Big Five Personality Traits					Peers' Perceptions				Other Measures of social skills	**Social Skills Index**
	Openness	Conscientiousness	Emotional Stability	Extraversion	Agreeableness	Leadership	Friendliness	Popularity	Shyness		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: All Students											
More sociable peers	0.066**	0.026	0.035	0.067**	0.066**	0.002	0.037	0.026	0.009	0.048*	0.067**
	(0.032)	(0.032)	(0.031)	(0.031)	(0.031)	(0.019)	(0.024)	(0.019)	(0.017)	(0.028)	(0.027)
Higher-achieving peers	0.024	0.031	0.025	-0.006	-0.007	0.050***	-0.023	0.008	-0.016	-0.008	-0.006
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.019)	(0.024)	(0.020)	(0.017)	(0.023)	(0.022)
RI p-value more sociable peers	0.036	0.464	0.298	0.048	0.066	0.944	0.148	0.268	0.674	0.156	0.030
RI p-value higher-achieving peers	0.340	0.252	0.342	0.856	0.854	0.028	0.326	0.820	0.444	0.856	0.878
Control mean	-0.05	0.01	0.01	-0.03	-0.02	-0.05	-0.02	-0.04	0.02	-0.01	-0.01
N	4,942	4,942	4,942	4,942	4,942	3,544	3,544	3,544	3,544	6,103	6,103
Panel B: Less Sociable Students at Baseline											
More sociable peers	0.080*	0.023	0.060	0.084*	0.123***	0.004	0.061**	0.049**	0.013	0.085**	0.114***
	(0.045)	(0.046)	(0.042)	(0.045)	(0.046)	(0.021)	(0.027)	(0.022)	(0.029)	(0.039)	(0.039)
Higher-achieving peers	-0.039	0.047	0.051	-0.056	0.024	-0.012	-0.094***	-0.060***	-0.022	0.016	0.008
	(0.043)	(0.044)	(0.041)	(0.044)	(0.045)	(0.022)	(0.028)	(0.023)	(0.030)	(0.038)	(0.038)
RI p-value more sociable peers	0.092	0.642	0.204	0.078	0.018	0.910	0.046	0.050	0.712	0.074	0.004
RI p-value higher-achieving peers	0.386	0.342	0.324	0.264	0.654	0.600	0.002	0.010	0.476	0.724	0.856
Control mean	-0.15	-0.17	-0.16	-0.16	-0.24	-0.24	-0.31	-0.19	0.17	-0.21	-0.26
N	1,536	1,536	1,536	1,536	1,536	1,763	1,763	1,763	1,763	1,832	1,832
Panel C: More Sociable Students at Baseline											
More sociable peers	0.053	0.028	0.010	0.051	0.012	0.003	0.023	0.004	0.005	0.010	0.019
	(0.042)	(0.042)	(0.045)	(0.039)	(0.041)	(0.032)	(0.039)	(0.031)	(0.019)	(0.039)	(0.037)
Higher-achieving peers	0.083**	0.021	0.020	0.006	-0.033	0.108***	0.053	0.075**	-0.009	-0.012	-0.011
	(0.041)	(0.042)	(0.045)	(0.039)	(0.041)	(0.030)	(0.037)	(0.031)	(0.019)	(0.039)	(0.037)
RI p-value more sociable peers	0.240	0.560	0.848	0.280	0.858	0.954	0.588	0.996	0.788	0.802	0.630
RI p-value higher-achieving peers	0.070	0.688	0.678	0.946	0.460	0.002	0.244	0.046	0.718	0.718	0.754
Control mean	-0.03	-0.09	-0.04	0.04	-0.02	0.10	0.24	0.06	-0.18	-0.01	0.01
N	1,570	1,570	1,570	1,570	1,570	1,781	1,781	1,781	1,781	1,822	1,822
p-value of difference	0.655	0.941	0.416	0.594	0.072	0.903	0.498	0.263	0.816	0.175	0.079

Notes: This table reports the effect of being assigned to more sociable and higher-achieving peers identified at baseline based on personality traits, peers' perceptions, and social skills outcomes. All regressions include strata fixed effects and control for the baseline value of the dependent variable. For the 2017 cohort all regressions include strata-by-classroom fixed effects. To classify students as more vs. less sociable, I use sociability at baseline as measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality measures a student's influence in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Eigenvector centrality is not measured for the 2017 cohort. To classify students as higher vs. lower achieving, I use the score in the admissions test to the COAR Network, which evaluates the applicants in math and reading comprehension. The control group is defined as being assigned to less sociable and lower-achieving peers. Personality traits correspond to the Big Five. Measures of peers' perceptions correspond to the number of peers who think the student is in the top five of leadership, friendliness, popularity, and shyness (school-by-grade). The table presents two social skills indexes for robustness. The first is constructed using PCA on all the variables that measure social skills (see Appendix C for details), excluding personality traits and measures of peers' perceptions. The second index is constructed using PCA on all the variables that measure social skills (see Appendix C for details). The table also reports the RI p-values of each treatment. The treatments were randomly reassigned on 1,000 permutations of students. Each outcome variable was then regressed on the reassigned treatment indicators. The true coefficients were then compared to the distribution of coefficients induced by reassignment in order to generate the RI p-values. The last row reports the p-value of equal treatment effects of more sociable peers for less sociable and more sociable students at baseline. Standard errors are clustered at the group of peers-by-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 5: Reduced-Form Estimates on Cognitive Skills

	Grades		Test Scores	
	Math	Reading	Math	Reading
	(1)	(2)	(3)	(4)
	Panel A: All Students			
More sociable peers	0.004	0.038	-0.035	0.021
	(0.035)	(0.038)	(0.024)	(0.032)
Higher-achieving peers	-0.022	-0.003	-0.028	-0.023
	(0.025)	(0.027)	(0.019)	(0.024)
RI p-value more sociable peers	0.924	0.338	0.208	0.548
RI p-value higher-achieving peers	0.360	0.830	0.154	0.284
Control mean	-0.05	-0.05	-0.04	-0.03
N	4,407	4,410	4,414	4,436
	Panel B: Lower-Achieving Students at Baseline			
More sociable peers	-0.004	0.086	-0.046	0.024
	(0.049)	(0.052)	(0.032)	(0.044)
Higher-achieving peers	-0.081**	-0.049	-0.049*	-0.068**
	(0.035)	(0.039)	(0.026)	(0.033)
RI p-value more sociable peers	0.924	0.118	0.154	0.616
RI p-value higher-achieving peers	0.034	0.170	0.056	0.032
Control mean	-0.27	-0.21	-0.29	-0.09
N	2,195	2,196	2,196	2,209
	Panel C: Higher-Achieving Students at Baseline			
More sociable peers	0.013	-0.009	-0.024	0.018
	(0.050)	(0.054)	(0.036)	(0.048)
Higher-achieving peers	0.031	0.041	-0.010	0.019
	(0.036)	(0.037)	(0.028)	(0.035)
RI p-value more sociable peers	0.828	0.920	0.592	0.680
RI p-value higher-achieving peers	0.482	0.362	0.750	0.640
Control mean	0.26	0.18	0.32	0.05
N	2,212	2,214	2,218	2,227
p-value of difference	0.037	0.097	0.293	0.065

Notes: This table reports the effect of being assigned to more sociable and higher-achieving peers identified at baseline on grades and test scores. All regressions include strata fixed effects and control for the baseline value of the dependent variable. For the 2017 cohort all regressions include strata-by-classroom fixed effects. To classify students as more vs. less sociable, I use sociability at baseline as measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality measures a student's influence in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Eigenvector centrality is not measured for the 2017 cohort. To classify students as higher vs. lower achieving, I use the score in the admissions test to the COAR Network, which evaluates the applicants in math and reading comprehension. The control group is defined as being assigned to less sociable and lower-achieving peers. Grades are standardized at the school-by-grade level and test scores at the grade level. The table also reports the RI p-values of each treatment. The treatments were randomly reassigned on 1,000 permutations of students. Each outcome variable was then regressed on the reassigned treatments indicators. The true coefficients were then compared to the distribution of coefficients induced by reassignment in order to generate the RI p-values. The last row reports the p-value of equal treatment effects of higher-achieving peers for lower- and higher-achieving students at baseline. Standard errors are clustered at the group of peers-by-type level; *** p-value$<$0.01, ** p-value$<$0.05, * p-value$<$0.1.

Table 6: 2SLS Estimates of Social and Cognitive Peer Effects on Students' Outcomes

	Social Skills Index	Grades		Test Scores	
		Math	Reading	Math	Reading
	(1)	(2)	(3)	(4)	(5)
	Panel A: All Students				
Roommates' sociability	0.132***	0.010	0.068	-0.052	0.061
	(0.050)	(0.064)	(0.070)	(0.044)	(0.057)
Roommates' academic achievement	-0.021	-0.059	-0.015	-0.045	-0.047
	(0.038)	(0.044)	(0.047)	(0.033)	(0.041)
N	5,988	4,343	4,346	4,350	4,372
	Panel B: Less Sociable Students at Baseline				
Roommates' sociability	0.237***	-0.001	-0.037	-0.129**	0.057
	(0.075)	(0.097)	(0.108)	(0.064)	(0.090)
Roommates' academic achievement	-0.017	-0.135*	0.049	-0.057	-0.104
	(0.063)	(0.077)	(0.087)	(0.053)	(0.074)
N	1,793	993	994	995	1,008
	Panel C: More Sociable Students at Baseline				
Roommates' sociability	0.031	0.016	0.152*	-0.010	0.067
	(0.064)	(0.084)	(0.090)	(0.059)	(0.070)
Roommates' academic achievement	-0.016	-0.042	-0.048	0.030	-0.036
	(0.061)	(0.084)	(0.088)	(0.062)	(0.067)
N	1,782	1,015	1,014	1,014	1,021
	Panel D: Lower-Achieving Students at Baseline				
Roommates' sociability	0.254***	0.015	0.153*	-0.065	0.062
	(0.068)	(0.087)	(0.093)	(0.056)	(0.074)
Roommates' academic achievement	-0.051	-0.150**	-0.091	-0.082*	-0.122**
	(0.058)	(0.064)	(0.071)	(0.047)	(0.058)
N	2,999	2,169	2,170	2,171	2,184
	Panel E: Higher-Achieving Students at Baseline				
Roommates' sociability	0.011	0.013	-0.022	-0.031	0.065
	(0.075)	(0.096)	(0.104)	(0.069)	(0.089)
Roommates' academic achievement	0.004	0.015	0.053	-0.020	0.017
	(0.052)	(0.060)	(0.061)	(0.047)	(0.056)
N	2,989	2,174	2,176	2,179	2,188

Notes: This table reports 2SLS estimates of roommates' average sociability and academic achievement at baseline on students' outcomes. There are two endogenous variables: roommates' sociability and roommates academic achievement at baseline. I instrument for these using indicators for whether the student was assigned to the more sociable or higher-achieving peers treatment. Roommates are students in the same dormitory for small dormitories and neighbors—students in the same or adjacent bunk bed—for large dormitories. All regressions include strata fixed effects and control for the baseline value of the dependent variable. For the 2017 cohort all regressions include strata-by-classroom fixed effects. To classify students as more vs. less sociable, I use sociability at baseline as measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality measures a student's influence in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Eigenvector centrality is not measured for the 2017 cohort. To classify students as higher vs. lower achieving, I use the score in the admissions test to the COAR Network, which evaluates the applicants in math and reading comprehension. The control group is defined as being assigned to less sociable and lower-achieving peers. The social skills index is constructed using PCA on all the variables that measure social skills (see Appendix C for details). Grades are standardized at the school-by-grade level and test scores at the grade level. Standard errors are clustered at the group of peers-by-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 7: Reduced Form Estimates on Social Support

Dependent variable:	I feel supported by my friends	I am satisfied with my dormitory	I am friends with my roommates	I am empathetic with my roommates	My roommates are empathetic with me	My roommates are worried about me	**Quality Index**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Panel A: Less Sociable Students at Baseline			
More sociable peers	0.089**	0.138***	0.063	0.096	0.217***	0.020	0.358**
	(0.041)	(0.050)	(0.051)	(0.080)	(0.077)	(0.045)	(0.139)
Higher-achieving peers	0.022	-0.019	0.017	0.118	0.081	0.067	0.126
	(0.042)	(0.050)	(0.050)	(0.081)	(0.082)	(0.043)	(0.135)
Control mean	-0.10	-0.14	-0.11	-0.13	-0.20	-0.12	-0.55
N	1,646	1,615	1,614	648	647	1,670	1,832
				Panel B: More Sociable Students at Baseline			
More sociable peers	0.027	0.068	0.110**	0.070	0.032	0.032	0.217
	(0.044)	(0.047)	(0.048)	(0.086)	(0.082)	(0.041)	(0.137)
Higher-achieving peers	-0.037	-0.051	-0.011	0.188**	0.109	-0.010	0.031
	(0.043)	(0.046)	(0.046)	(0.088)	(0.081)	(0.040)	(0.130)
Control mean	0.02	0.05	-0.00	-0.12	-0.12	0.04	-0.01
N	1,660	1,646	1,645	696	696	1,706	1,822
p-value of difference	0.314	0.308	0.502	0.880	0.114	0.845	0.463

Notes: This table reports the effect of being assigned to more sociable and higher-achieving peers on the perceived quality of interactions with friends and roommates. All regressions include strata fixed effects and control for the baseline value of the dependent variable. For the 2017 cohort all regressions include strata-by-classroom fixed effects. To classify students as more vs. less sociable, I use sociability at baseline as measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality measures a student's influence in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Eigenvector centrality is not measured for the 2017 cohort. To classify students as higher vs. lower achieving, I use the score in the admissions test to the COAR Network, which evaluates the applicants in math and reading comprehension. The control group is defined as being assigned to less sociable and lower-achieving peers. All the dependent variables are standardized at the school-by-grade level. The quality index is the sum of all the other dependent variables. The last row reports the p-value of equal treatment effects of more sociable peers for less and more sociable students at baseline. Standard errors are clustered at the group of peers-by-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 8: Reduced-Form Estimates on Self-Confidence in Social Skills

Dependent variable:	Self-perception				Self-confidence Index	
	I am in the top five of:				Excluding leadership	All variables
	Leadership	Friendliness	Popularity	Shyness		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Less Sociable Students at Baseline					
More sociable peers	-0.044	0.037	0.061	0.072*	0.169**	0.129
	(0.042)	(0.045)	(0.045)	(0.043)	(0.078)	(0.097)
Higher-achieving peers	-0.003	-0.043	0.003	-0.020	-0.061	-0.066
	(0.043)	(0.045)	(0.044)	(0.043)	(0.076)	(0.094)
Control mean	-0.05	-0.01	-0.01	-0.06	-0.07	-0.12
N	1,681	1,681	1,681	1,681	1,832	1,832
	Panel B: More Sociable Students at Baseline					
More sociable peers	-0.087*	-0.025	0.001	-0.003	-0.040	-0.128
	(0.045)	(0.047)	(0.045)	(0.040)	(0.080)	(0.101)
Higher-achieving peers	0.053	-0.016	-0.071	0.040	-0.048	-0.001
	(0.046)	(0.048)	(0.045)	(0.039)	(0.077)	(0.101)
Control mean	0.02	-0.00	0.00	-0.02	-0.02	-0.00
N	1,710	1,710	1,710	1,710	1,822	1,822
p-value of difference	0.474	0.339	0.337	0.203	0.062	0.069

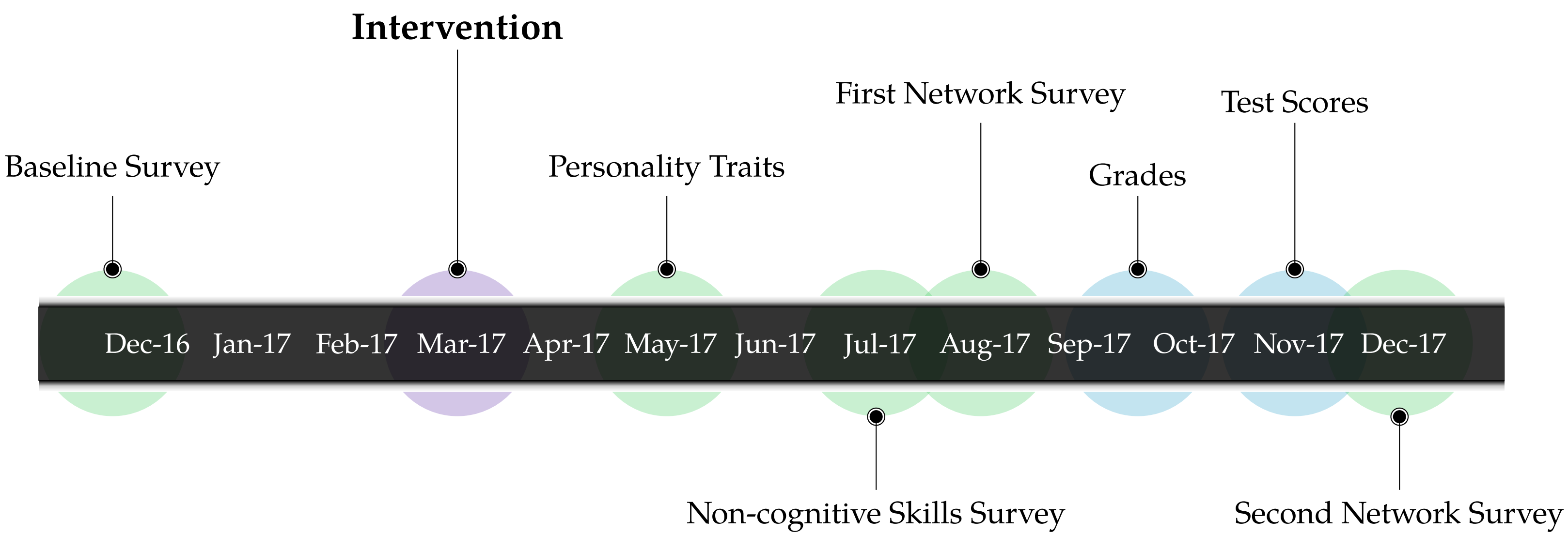
Notes: This table reports the effect of being assigned to more sociable and higher-achieving peers identified on self-confidence in social skills. All regressions include strata fixed effects and control for the baseline value of the dependent variable. For the 2017 cohort all regressions include strata-by-classroom fixed effects. To classify students as more vs. less sociable, I use sociability at baseline as measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality measures a student's influence in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Eigenvector centrality is not measured for the 2017 cohort. To classify students as higher vs. lower achieving, I use the score in the admissions test to the COAR Network, which evaluates the applicants in math and reading comprehension. The control group is defined as being assigned to less sociable and lower-achieving peers. All the dependent variables are standardized at the school-by-grade level, and weighted by the ranking that the student assigned to herself in the survey. In column 5, the self-confidence index is the sum of all the other dependent variables excluding leadership. In column 6, the self-confidence index is the sum of all the other dependent variables. The last row reports the p-value of equal treatment effects of more sociable peers for less and more sociable students at baseline. Standard errors are clustered at the group of peers-by-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table 9: Reduced-Form Estimates of Self-Confidence in Cognitive Skills

Dependent variable:	What is my level of belonging to the school	**What is my academic ranking:** in the school?	among friends?	**Goals with respect to peers:** I try to perform better	I avoid doing worse	**Self-perception:** I am in the top five of academically skilled	I volunteer to receive academic information	**Self-confidence Index**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Lower-Achieving Students at Baseline				
More sociable peers	-0.003	-0.037	-0.037	-0.015	0.008	-0.028	0.012	-0.118
	(0.041)	(0.042)	(0.044)	(0.048)	(0.048)	(0.037)	(0.042)	(0.123)
Higher-achieving peers	-0.066*	-0.035	-0.089**	-0.068*	-0.032	-0.017	-0.072**	-0.318***
	(0.035)	(0.035)	(0.036)	(0.041)	(0.041)	(0.033)	(0.035)	(0.106)
Control mean	0.02	-0.08	-0.05	0.03	-0.04	-0.07	0.02	0.35
N	2,801	2,813	2,805	2,680	2,680	2,829	2,829	3,050
				Panel B: Higher-Achieving Students at Baseline				
More sociable peers	0.009	-0.019	0.027	-0.079	-0.050	-0.085*	0.007	-0.107
	(0.043)	(0.042)	(0.043)	(0.048)	(0.044)	(0.050)	(0.045)	(0.133)
Higher-achieving peers	0.044	0.004	-0.023	-0.006	0.002	0.065	0.033	0.109
	(0.035)	(0.033)	(0.035)	(0.038)	(0.036)	(0.040)	(0.037)	(0.108)
Control mean	-0.02	0.11	0.09	0.00	0.06	0.08	-0.01	0.72
N	2,856	2,855	2,854	2,763	2,763	2,867	2,867	3,053
p-value of difference	0.027	0.429	0.191	0.270	0.537	0.118	0.041	0.005

Notes: This table reports the effect of being assigned to more sociable and higher-achieving peers on self-confidence in cognitive skills. All regressions include strata fixed effects and control for the baseline value of the dependent variable. For the 2017 cohort all regressions include strata-by-classroom fixed effects. To classify students as more vs. less sociable, I use sociability at baseline as measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality measures a student's influence in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Eigenvector centrality is not measured for the 2017 cohort. To classify students as higher vs. lower achieving, I use the score in the admissions test to the COAR Network, which evaluates the applicants in math and reading comprehension. The control group is defined as being assigned to less sociable and lower-achieving peers. All the dependent variables are standardized at the school-by-grade level. The self-perception measures are weighted by the ranking that the student assigned to herself in the survey. The self-confidence index is the sum of all the other dependent variables. The last row reports the p-value of equal treatment effects of higher-achieving peers for lower and higher-achieving students at baseline. Standard errors are clustered at the group of peers-by-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Figure 1: Timeline of the Project



Notes: This figure presents the timeline of the project. The green circles represent instances of data collection with surveys, the blue circles the collection of administrative data, and the purple circle the intervention.

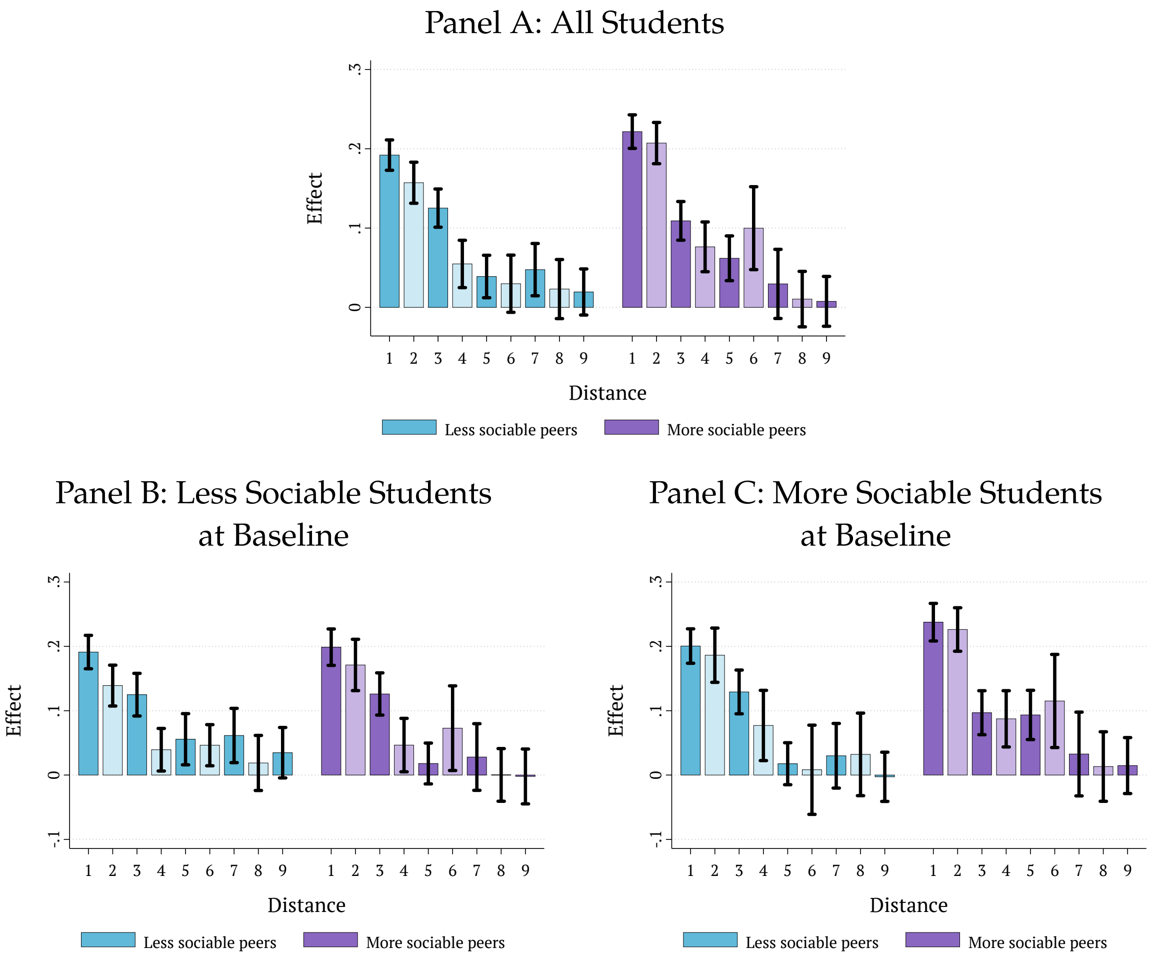
Figure 2: Groups of Peers in the Experimental Design

Type of Peers

Student Type	more sociable higher-achieving	more sociable lower-achieving	less sociable higher-achieving	less sociable lower-achieving
more sociable higher-achieving	Group 1	Group 2	Group 3	Group 4
more sociable lower-achieving	Group 2	Group 5	Group 6	Group 7
less sociable higher-achieving	Group 3	Group 6	Group 8	Group 9
less sociable lower-achieving	Group 4	Group 7	Group 9	Group 10

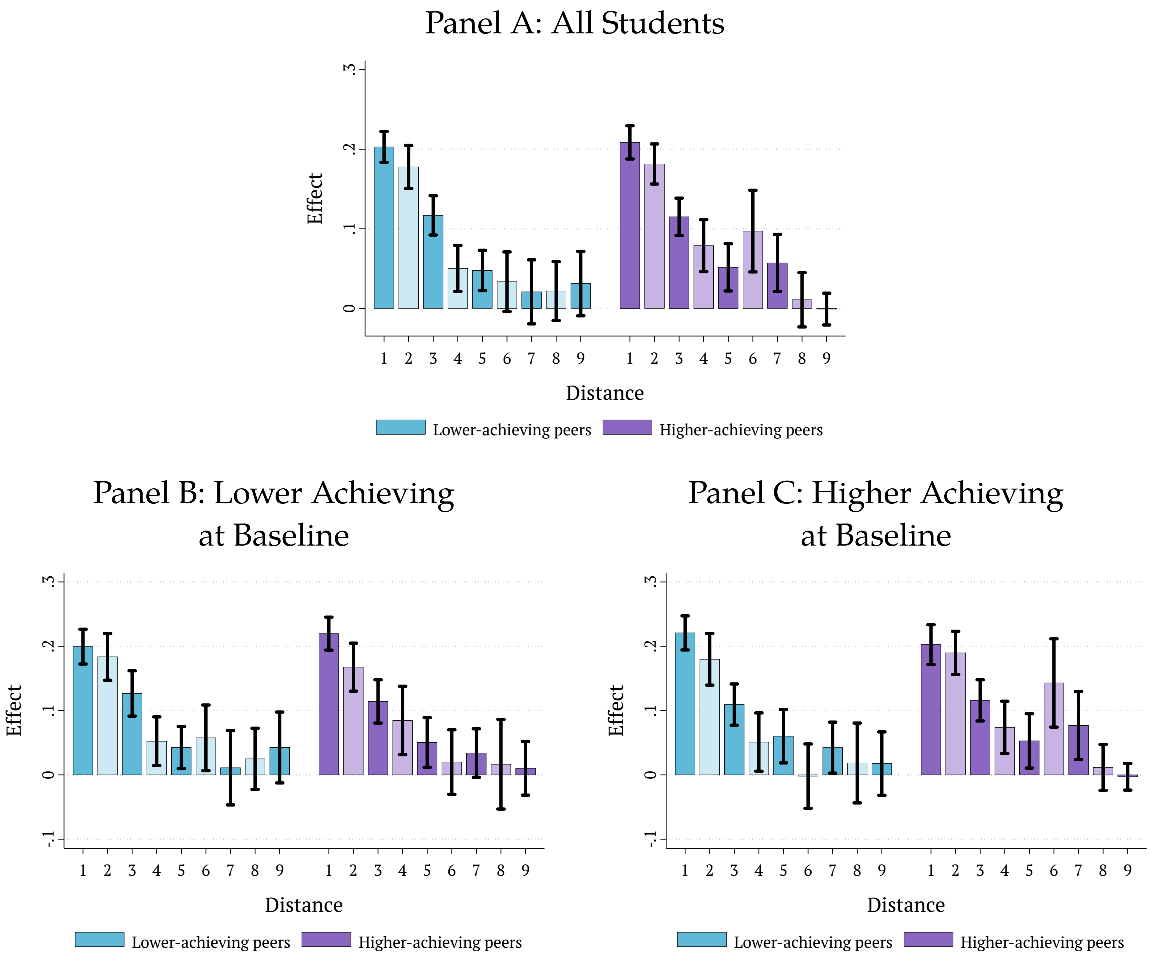
Notes: This figure shows the ten *groups of peers* in my experimental design. It represents all possible combinations between student type and type of peers. Rows are described by student types, and columns illustrate the types of peers to which they were randomly assigned. The diagonal of the matrix is composed by groups of a single type. The matrix is symmetric by virtue of the fact that students are matched with peers of the assigned type.

Figure 3: Effects of Proximity on Link Formation by Level of Sociability



Notes: This figure shows the impact of distance between a pair of students on the likelihood that they become friends, study together, or engage in social activities such as playing or dancing. Distance is captured by nine distance dummies, and 95% confidence intervals are displayed for all proximity effects. The figure plots the impact by treatment (more sociable peers) vs. control (less sociable peers) status in sociability. By construction in the experimental design, students are at an odd distance from their treatment peers, and at an even distance from the peers of their type. Panel A presents the results for all students, Panel B for the less sociable students at baseline, and Panel C for the more sociable students at baseline. To classify students as more vs. less sociable, I use sociability at baseline measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. All estimations control for strata fixed effects. Standard errors are double clustered at the student level and at the group-by-type level.

FIGURE 4: Effects of Proximity on Link Formation by Academic Achievement



Notes: This figure shows the impact of distance between a pair of students on the likelihood that they become friends, study together, or engage in social activities such as playing or dancing. Distance is captured by nine distance dummies, and 95% confidence intervals are displayed for all proximity effects. The figure plots the impact by treatment (higher-achieving peers) vs. control (lower-achieving peers) status in academic achievement. By construction in the experimental design, students are at an odd distance from their treatment peers, and at an even distance from the peers of their type. Panel A presents the results for all students, Panel B for lower-achieving students, and Panel C for higher-achieving students. To classify students as higher vs. lower achieving, I use the score in the admissions test to the COAR Network, which assess the competencies of applicants in math and reading comprehension. All estimations control for strata fixed effects. Standard errors are double-clustered at the student level and at the group-by-type level.

A Supplementary Material

Table A.1: Summary Statistics for the Networks Survey at Baseline

	All Students	By Sociability		By Academic Achievement	
		Less Sociable	More Sociable	Lower Achieving	Higher Achieving
Variable	(1)	(2)	(3)	(4)	(5)
Dorm preferences degree	6.49	5.16	7.83	6.27	6.71
	(3.47)	(2.67)	(3.66)	(3.52)	(3.41)
Dorm preferences mutual degree	1.69	1.35	2.04	1.57	1.82
	(1.54)	(1.27)	(1.69)	(1.49)	(1.57)
Dorm preferences centrality	-0.00	-0.23	0.23	-0.03	0.03
	(1.00)	(0.75)	(1.14)	(0.97)	(1.02)
Friendships degree	7.87	5.67	10.10	7.62	8.12
	(5.11)	(2.75)	(5.91)	(5.11)	(5.10)
Friendships mutual degree	2.10	1.65	2.55	1.98	2.23
	(1.68)	(1.37)	(1.83)	(1.70)	(1.65)
Friendships centrality	-0.00	-0.51	0.51	-0.06	0.06
	(1.00)	(0.57)	(1.07)	(0.98)	(1.00)
Study partnerships degree	4.76	3.71	5.82	4.54	4.98
	(2.66)	(1.94)	(2.86)	(2.52)	(2.77)
Study partnerships mutual degree	1.11	0.92	1.29	1.04	1.18
	(1.08)	(0.95)	(1.18)	(1.05)	(1.12)
Study partnerships centrality	0.00	-0.37	0.37	-0.07	0.07
	(1.00)	(0.66)	(1.13)	(0.92)	(1.06)
Social partnerships degree	5.67	4.44	6.90	5.61	5.73
	(3.13)	(2.29)	(3.36)	(3.20)	(3.05)
Social partnerships mutual degree	1.23	1.00	1.46	1.21	1.25
	(1.23)	(1.05)	(1.36)	(1.24)	(1.23)
Social partnerships centrality	0.00	-0.37	0.37	-0.01	0.01
	(1.00)	(0.69)	(1.11)	(1.01)	(0.98)
Any partnership degree	11.07	8.23	13.92	10.78	11.35
	(5.66)	(3.19)	(6.15)	(5.61)	(5.70)
Any partnership mutual degree	3.34	2.60	4.07	3.20	3.48
	(2.18)	(1.66)	(2.38)	(2.19)	(2.16)
Any partnership centrality	0.00	-0.71	0.71	-0.05	0.05
	(1.00)	(0.44)	(0.89)	(0.97)	(1.01)
Peers who named the student as a leader	2.63	1.54	3.73	1.91	3.35
	(5.16)	(3.31)	(6.32)	(4.21)	(5.86)
Peers who named the student as friendly	2.71	1.91	3.50	2.61	2.80
	(2.81)	(1.91)	(3.31)	(2.76)	(2.86)
Peers who named the student as popular	2.43	1.51	3.35	1.99	2.86
	(5.32)	(3.44)	(6.58)	(4.48)	(6.02)
Peers who named the student as skilled	2.62	1.80	3.45	1.20	4.04
	(6.45)	(5.97)	(6.80)	(3.14)	(8.32)
Peers who named the student as shy	2.00	2.52	1.49	2.18	1.83
	(4.67)	(5.16)	(4.04)	(4.62)	(4.71)
N	3,654	1,832	1,822	1,822	1,832

Notes: This table reports summary statistics for the social networks survey at baseline. Information is presented by categories of student student. Standard errors are reported in parentheses. Column 1 shows statistics for all the students and columns 2 to 5 according to the classification of students by social and cognitive skills. To classify students as more vs. less sociable, I use sociability at baseline measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality measures a student's influence in the network. A student with a high eigenvector score means that she is connected to other students who themselves have a high influence as well. To classify students as higher vs. lower achieving, I use the score in the admissions test to the COAR Network, which evaluates the applicants in math and reading comprehension. The dorm preferences network is based on the question: "who would you like to have as roommates?". The friendship network is based on the question: "who are your friends?". The study partnerships network is based on the question: "with whom have you studied?". The social partnerships network is based on the question: "with whom have you engage in social activities with such as playing or dancing?". The any partnership network aggregates the four questions.

Table A.2: Correlation of Sociability and Social Skills Outcomes

	Big Five Personality Traits					Peers' Perception				Other measures of social skills	Social Skills Index Before the Intervention	Social Skills Index After the Interventipn
	Openness	Conscientiousness	Emotional Stability	Extraversion	Agreeableness	Leadership	Friendliness	Popularity	Shyness			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Academic achievement at baseline	0.090***	-0.001	0.064***	-0.014	-0.013	0.217***	0.042**	0.105***	-0.066***	0.047***	0.064***	0.042**
	(0.021)	(0.022)	(0.022)	(0.019)	(0.020)	(0.020)	(0.016)	(0.022)	(0.019)	(0.016)	(0.017)	(0.016)
Sociability at baseline	0.103***	0.061**	0.032*	0.142***	0.120***	0.230***	0.359***	0.215***	-0.103***	0.099***	0.139***	0.142***
	(0.024)	(0.025)	(0.017)	(0.021)	(0.017)	(0.021)	(0.020)	(0.029)	(0.021)	(0.018)	(0.018)	(0.018)
Social-fit score	0.072***	0.022	0.001	0.063***	0.027	0.137***	0.073***	0.104***	-0.113***	0.027	0.066***	0.046**
	(0.018)	(0.020)	(0.020)	(0.019)	(0.020)	(0.015)	(0.016)	(0.017)	(0.020)	(0.018)	(0.017)	(0.019)
Interview score	0.092***	0.073***	0.071***	0.091***	0.049***	0.068***	0.057***	0.049***	-0.036**	0.074***	0.118***	0.088***
	(0.019)	(0.016)	(0.016)	(0.018)	(0.018)	(0.015)	(0.015)	(0.017)	(0.017)	(0.015)	(0.016)	(0.015)
N	3,106	3,106	3,106	3,106	3,106	3,637	3,637	3,637	3,637	3,654	3,654	3,654

Notes: This table reports standardized estimates of an OLS regression on social skills outcomes of sociability at baseline and the score in the three tests of the admission process to the COAR Network. All regressions include school-by-grade-by-gender fixed effects. Sociability at baseline is measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality is a measure of the influence of a student in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Academic achievement at baseline is the score in the admission test to the COAR Network, which evaluates the applicants in math and reading comprehension. Personality traits correspond to the Big Five. Measures of peers' perception correspond to the number of peers who think the student is in the top 5 of leadership, friendliness, popularity, and shyness (school-by-grade). The table presents two social skills indexes for robustness. The first one is constructed using Principal Component Analysis (PCA) on all the variables that measure social skills (see Appendix C for details), excluding personality traits and measures of peers' perception. The second index is contructed using Principal Component Analysis (PCA) on all the variables that measure social skills (see Appendix C for details).

Table A.3: Balance Tests for the More Sociable Peers Treatment

Variable	All Students		Less Sociable at Baseline		More Sociable at Baseline	
	Control mean	Difference	Control mean	Difference	Control mean	Difference
Admission test	-0.016	0.002	-0.064	-0.024	0.058	0.028
		(0.019)		(0.027)		(0.026)
Interview score	14.065	0.009	14.053	0.019	14.082	-0.000
		(0.042)		(0.059)		(0.061)
Social-fit score	20.001	0.095	19.977	0.071	20.035	0.119
		(0.064)		(0.088)		(0.092)
Female	0.589	0.003	0.594	0.002	0.582	0.004
		(0.003)		(0.004)		(0.003)
Not poor	0.426	-0.014	0.393	-0.008	0.475	-0.020
		(0.014)		(0.021)		(0.020)
Poor	0.364	0.026*	0.378	0.034	0.344	0.019
		(0.015)		(0.022)		(0.020)
Extremely poor	0.210	-0.012	0.229	-0.026	0.182	0.001
		(0.012)		(0.018)		(0.017)
Rural	0.279	-0.014	0.315	-0.028	0.224	0.000
		(0.012)		(0.018)		(0.017)
Subsidized health insurance	0.508	0.008	0.552	-0.011	0.443	0.027
		(0.016)		(0.022)		(0.022)
Average math at baseline	-0.049	0.019	-0.157	0.012	0.115	0.027
		(0.027)		(0.037)		(0.038)
Average reading at baseline	-0.049	0.010	-0.169	0.041	0.134	-0.021
		(0.028)		(0.041)		(0.038)
Sociability index baseline	-0.042	-0.041	-0.293	-0.033	0.333	-0.050
		(0.080)		(0.117)		(0.111)
Indegree baseline network	-0.077	-0.021	-0.434	-0.003	0.460	-0.040
		(0.027)		(0.033)		(0.044)
Outdegree baseline network	-0.094	0.023	-0.419	0.016	0.395	0.029
		(0.026)		(0.027)		(0.046)
Centrality baseline network	0.303	0.003	0.203	-0.002	0.453	0.008
		(0.003)		(0.003)		(0.006)
Peers' perception leader	2.502	-0.184	1.587	-0.113	3.879	-0.257
		(0.151)		(0.132)		(0.272)
Peers' perception friendly	2.553	-0.009	1.837	0.200**	3.631	-0.221
		(0.080)		(0.083)		(0.136)
Peers' perception popular	2.203	0.110	1.466	0.126	3.313	0.093
		(0.159)		(0.142)		(0.285)
Peers' perception shy	2.083	0.029	2.560	-0.131	1.366	0.192
		(0.144)		(0.227)		(0.178)
Baseline grit	43.707	-0.248	43.340	-0.171	44.251	-0.325
		(0.198)		(0.285)		(0.274)
Baseline Rosenberg scale	32.991	0.101	32.777	0.119	33.306	0.081
		(0.154)		(0.224)		(0.212)
Baseline Read the Mind	20.521	-0.060	20.224	0.184	20.960	-0.304*
		(0.130)		(0.186)		(0.180)
Multivariate F p-value		0.785		0.571		0.726

Notes: This table reports balance checks of being assigned to more sociable peers on baseline characteristics. All regressions include strata fixed effects, control for the baseline value of the dependent variable, and include the higher-achieving peers treatment. For the 2017 cohort all regressions include strata-by-classroom fixed effects. To classify students as more vs. less sociable, I use sociability at baseline as measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality is a measure of the influence of a student in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Eigenvector centrality is not measured for the 2017 cohort. To classify students as higher vs. lower achieving, I use the score in the admission test to the COAR Network, which evaluates the applicants in math and reading comprehension. The control group is defined as being assigned to less sociable and lower-achieving peers. The "F p-value" correspond to the F-statistic of the more sociable peers treatment of multivariate regressions that include all the variables at baseline. Standard errors are clustered at the group of peers-by-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table A.4: Balance Tests for the Higher-Achieving Peers Treatment

Variable	All Students		Lower Achieving at Baseline		Higher Achieving at Baseline	
	Control mean	Difference	Control mean	Difference	Control mean	Difference
Admission test	-0.163	0.013	-0.787	-0.006	0.764	0.031
		(0.016)		(0.018)		(0.026)
Interview score	8.400	-0.013	8.589	-0.011	8.117	-0.015
		(0.033)		(0.042)		(0.052)
Social-fit score	11.973	-0.061	12.161	0.012	11.693	-0.131*
		(0.049)		(0.061)		(0.077)
Female	0.575	-0.000	0.575	0.004	0.575	-0.004
		(0.002)		(0.004)		(0.003)
Not poor	0.443	0.010	0.404	0.001	0.502	0.019
		(0.012)		(0.017)		(0.018)
Poor	0.369	-0.017	0.387	-0.005	0.342	-0.030*
		(0.013)		(0.018)		(0.017)
Extremely poor	0.188	0.007	0.209	0.003	0.156	0.011
		(0.011)		(0.016)		(0.014)
Rural	0.276	-0.003	0.319	-0.021	0.212	0.014
		(0.013)		(0.019)		(0.016)
Subsidized health insurance	0.504	0.013	0.517	0.045**	0.485	-0.019
		(0.015)		(0.021)		(0.022)
Average math at baseline	-0.093	0.016	-0.373	-0.023	0.319	0.053
		(0.023)		(0.030)		(0.034)
Average reading at baseline	-0.064	-0.000	-0.281	-0.012	0.256	0.011
		(0.023)		(0.032)		(0.033)
Sociability index baseline	-0.023	0.048	-0.065	0.060	0.040	0.038
		(0.059)		(0.083)		(0.083)
Indegree baseline network	0.008	-0.050*	-0.075	-0.032	0.131	-0.068*
		(0.027)		(0.038)		(0.039)
Outdegree baseline network	-0.001	-0.013	-0.016	-0.059	0.021	0.033
		(0.028)		(0.040)		(0.038)
Centrality baseline network	0.330	-0.003	0.324	-0.007	0.338	0.000
		(0.003)		(0.004)		(0.004)
Peers' perception leader	2.493	0.006	1.930	-0.036	3.331	0.050
		(0.148)		(0.178)		(0.236)
Peers' perception friendly	2.691	0.010	2.618	-0.001	2.801	0.021
		(0.078)		(0.112)		(0.109)
Peers' perception popular	2.362	-0.013	2.007	-0.024	2.890	-0.006
		(0.158)		(0.186)		(0.255)
Peers' perception shy	2.020	0.036	2.110	0.160	1.886	-0.089
		(0.146)		(0.211)		(0.203)
Baseline grit	43.568	0.158	43.330	0.431	43.921	-0.108
		(0.196)		(0.271)		(0.282)
Baseline Rosenberg scale	33.013	0.139	32.854	0.127	33.249	0.153
		(0.154)		(0.221)		(0.215)
Baseline Read the Mind	20.602	-0.265**	20.248	-0.231	21.127	-0.295
		(0.128)		(0.180)		(0.183)
Multivariate F p-value		0.810		0.857		0.625

Notes: This table reports balance checks of being assigned to higher-achieving peers on baseline characteristics. All regressions include strata fixed effects, control for the baseline value of the dependent variable, and include the more sociable peers treatment. For the 2017 cohort all regressions include strata-by-classroom fixed effects. To classify students as more vs. less sociable, I use sociability at baseline as measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality is a measure of the influence of a student in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Eigenvector centrality is not measured for the 2017 cohort. To classify students as higher vs. lower achieving, I use the score in the admission test to the COAR Network, which evaluates the applicants in math and reading comprehension. The control group is defined as being assigned to less sociable and lower-achieving peers. The "F p-value" correspond to the F-statistic of the higher-achieving peers treatment of multivariate regressions that include all the variables at baseline. Standard errors are clustered at the group of peers-by-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table A.5: Effects of Proximity on Social Interactions

	All Students	By Sociability		By Academic Achievement	
		Less Sociable	More Sociable	Lower Achieving	Higher Achieving
Variable	(1)	(2)	(3)	(4)	(5)
Neighbor	0.161***	0.163***	0.141***	0.170***	0.151***
	(0.012)	(0.017)	(0.017)	(0.016)	(0.017)
More sociable peers	-0.000	0.002	-0.002	0.001	-0.002
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)
Higher-achieving peers	-0.000	-0.003*	0.002	-0.000	0.000
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)
Neighbor $\times$ more sociable peers	0.037**	0.019	0.048**	0.003	0.070***
	(0.015)	(0.020)	(0.021)	(0.020)	(0.022)
Neighbor $\times$ higher-achieving peers	0.001	-0.022	0.035	0.011	-0.007
	(0.013)	(0.020)	(0.022)	(0.018)	(0.019)
Control mean	0.11	0.11	0.11	0.11	0.11
N students	6,134	1,832	1,822	3,064	3,071
N pairs of students	751,758	223,719	223,147	375,896	375,862

Notes: This table reports the effect of being neighbors on the list on the likelihood that students will become friends, study together or engage in social activities. Effects are reported at the pair-of-students level, and include the interaction between this effect and the effect of having more sociable peers, as well as the effect of having higher-achieving peers. All regressions include strata fixed effects and control for whether the pair of students had any of the mentioned social interactions at baseline. To classify students as more vs. less sociable, I use sociability at baseline as measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality measures a student's influence in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Eigenvector centrality is not measured for the 2017 cohort. To classify students as higher vs. lower achieving, I use the score in the admissions test to the COAR Network, which evaluates the applicants in math and reading comprehension. Standard errors are two-way clustered at the student and at the group of peers-by-type level; *** $p - value < 0.01$, ** $p - value < 0.05$, * $p - value < 0.1$.

Table A.6: Attrition Differentials

Variable	Big 5 (1)	Non-cognitive skills survey (2)	1st network survey (3)	2nd network survey (4)	Grades (5)	Test Scores (6)
			Panel A: All Students			
More sociable peers	-0.021*	0.003	0.005	0.005	0.001	0.001
	(0.011)	(0.009)	(0.006)	(0.006)	(0.006)	(0.008)
Higher-achieving peers	-0.009	0.003	-0.009**	-0.003	0.001	-0.001
	(0.009)	(0.007)	(0.004)	(0.005)	(0.004)	(0.005)
Control mean	0.79	0.87	0.90	0.90	0.96	0.96
N	6,103	6,103	6,103	6,103	6,009	4,602
			Panel B: Less Sociable Students at Baseline			
More sociable peers	-0.055***	0.007	-0.003	-0.003	-0.001	-0.006
	(0.016)	(0.013)	(0.009)	(0.010)	(0.008)	(0.012)
Higher-achieving peers	-0.026*	-0.017	-0.019**	-0.022**	-0.005	-0.020
	(0.016)	(0.013)	(0.009)	(0.010)	(0.008)	(0.013)
Control mean	0.87	0.91	0.91	0.90	0.96	0.95
N	1,832	1,832	1,832	1,832	1,773	1,080
			Panel C: More Sociable Students at Baseline			
More sociable peers	0.013	-0.001	0.012*	0.013*	0.004	0.008
	(0.015)	(0.013)	(0.006)	(0.007)	(0.007)	(0.011)
Higher-achieving peers	0.008	0.007	-0.006	0.006	0.002	0.005
	(0.014)	(0.012)	(0.006)	(0.007)	(0.007)	(0.011)
Control mean	0.84	0.91	0.92	0.90	0.97	0.95
N	1,822	1,822	1,822	1,822	1,787	1,073
			Panel D: Lower-Achieving Students at Baseline			
More sociable peers	-0.006	-0.014	0.007	0.010	0.004	0.004
	(0.017)	(0.014)	(0.009)	(0.010)	(0.008)	(0.012)
Higher-achieving peers	-0.019	-0.009	-0.010	-0.010	0.002	-0.007
	(0.014)	(0.011)	(0.007)	(0.008)	(0.006)	(0.008)
Control mean	0.78	0.85	0.89	0.90	0.95	0.95
N	3,049	3,049	3,049	3,049	2,986	2,301
			Panel E: Higher-Achieving students at Baseline			
More sociable peers	-0.037**	0.019	0.002	-0.000	-0.001	-0.002
	(0.015)	(0.012)	(0.007)	(0.008)	(0.008)	(0.011)
Higher-achieving peers	0.002	0.014	-0.009	0.003	0.001	0.004
	(0.012)	(0.009)	(0.005)	(0.006)	(0.006)	(0.007)
Control mean	0.82	0.89	0.90	0.91	0.97	0.97
N	3,054	3,054	3,054	3,054	3,023	2,301

Notes: This table reports the effect of being assigned to more sociable and higher-achieving peers on data availability in the Big Five, the non-cognitive skills survey, the two Networks surveys, grades and test scores after the intervention. All regressions include strata fixed effects. For the 2017 cohort all regressions include strata-by-classroom fixed effects. To classify students as more vs. less sociable, I use sociability at baseline as measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality is a measure of the influence of a student in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Eigenvector centrality is not measured for the 2017 cohort. To classify students as higher vs. lower achieving, I use the score in the admission test to the COAR Network, which evaluates the applicants in math and reading comprehension. The control group is defined as being assigned to less sociable and lower-achieving peers. Standard errors are clustered at the group of peers-by-type level; *** p-value<0.01, ** p-value<0.05, * p-value<0.1.

Table A.7: Reduced Form Estimates on Personality Traits (Accounting for Attrition)

Dependent variable:	Attrition Big Five	Big Five Personality Traits				
		Openness	Conscientiousness	Emotional Stability	Extraversion	Agreeableness
	(1)	(2)	(3)	(4)	(5)	(6)
More sociable peers	-0.007	0.077*	0.033	0.051	0.084*	0.108**
	(0.016)	(0.046)	(0.048)	(0.043)	(0.046)	(0.046)
Higher-achieving peers	-0.025	-0.041	0.070	0.036	-0.061	0.017
	(0.016)	(0.045)	(0.046)	(0.043)	(0.046)	(0.046)
Control mean	0.87	-0.04	-0.04	-0.08	-0.01	-0.05
N	1,622	1,370	1,370	1,370	1,370	1,370

Notes: This table reports the effect of being assigned to more sociable and higher achieving peers identified at baseline on personality traits. Effects are presented for the less sociable students at baseline only, and account for attrition in the Big Five. All regressions include strata fixed effects and control for the baseline value of the dependent variable. For the 2017 cohort all regressions include strata-by-classroom fixed effects. To classify students as more vs. less sociable, I use sociability at baseline as measured by the eigenvector centrality of an aggregate social network of dorm preferences, friendships, study, and social partnerships. Eigenvector centrality is a measure of the influence of a student in the network. A student with a high eigenvector score is connected to other students who themselves have a high influence as well. Eigenvector centrality is not measured for the 2017 cohort. To classify students as higher vs. lower achieving, I use the score in the admission test to the COAR Network, which evaluates the applicants in math and reading comprehension. The control group is defined as being assigned to less sociable and lower-achieving peers. Personality traits correspond to the Big Five. Standard errors are clustered at the group of peers-by-type level; *** p-value$<$0.01, ** p-value$<$0.05, * p-value$<$0.1.

Table A.8: Treatments and Outcomes Available by Cohort

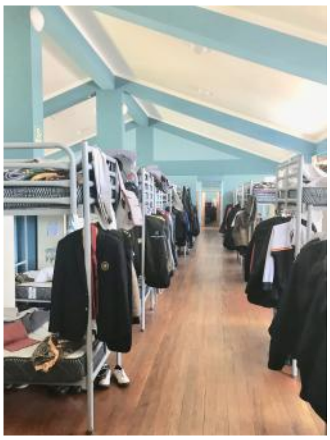
Treatments	**Outcomes**	
	Social Skills	Cognitive Skills
More sociable peers	2015, 2016	2016
Higher-achieving peers	2015, 2016, 2017	2016, 2017
Complementarity	2015, 2016	2016

Figure A.1: Figure Appendix: Dorm Structure

School in Lima

School in Piura

School in Cusco



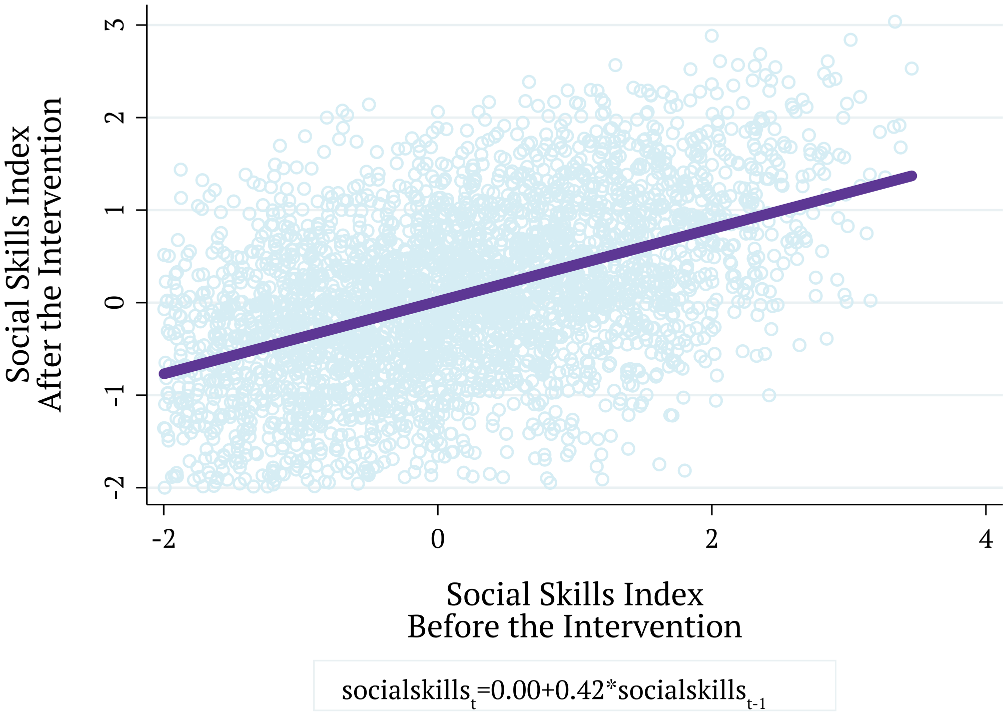
Notes: This figure displays pictures of the dorms for the schools in Lima, Piura, and Cusco.

FIGURE A.2: Correlation of Cognitive and Social Skills

achievement=0.00+0.11*socialskills

Notes: This figure shows a scatter plot of academic achievement and sociability at baseline for the 2015-16 cohorts by student type. A one standard deviation of the social skills index predicts an increase in 0.11 standard deviations of academic achievement at baseline.

Figure A.3: Correlation of the Sociability Index Before and After the Intervention



Notes: This figure shows a scatter plot and the linear prediction of the sociability index before and after the intervention. A one standard deviation of the social skills index before the intervention predicts an increase of 0.42 in the social skills index after the intervention.

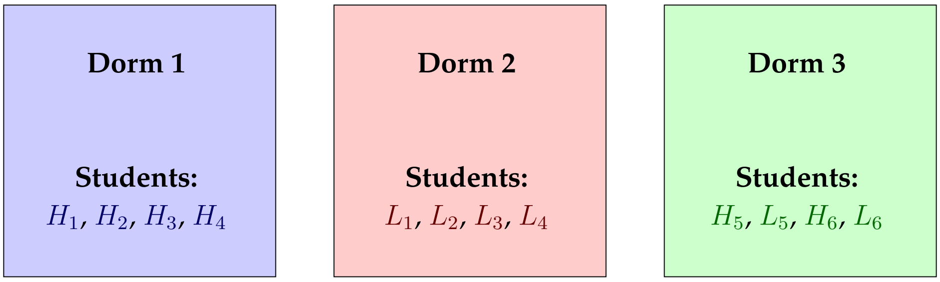
B Use of the Lists to Allocate Students to Dorms and Classrooms

This section explains in detail how the order of students on the lists determines their allocation to dormitories. To start, recall the simple two-type example in which students were either H and L. As described in section 3.3.2, this case allows for three groups of combined types: Group 1 (only Hs), Group 2 (a mixed group of Hs and Ls), and Group 3 (only Ls). Let us assume that the random ordering of the groups on the list is: Group 1-Group 3-Group 2. For the purpose of illustration, I will assume that there are 12 students, four in each group. After numbering the students by type, the order on the list would be the following: $H_1 - H_2 - H_3 - H_4 - L_1 - L_2 - L_3 - L_4 - H_5 - L_5 - H_6 - L_6$. Notice that on the list, the order of the students in Group 2 alternates the type of student in the form H-L.

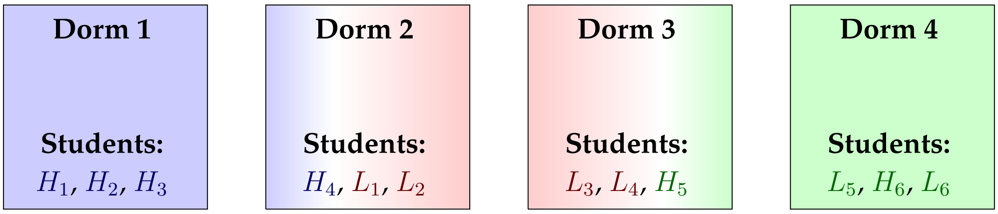
Figure B.1 shows an illustrative example of how the lists were used to allocate students to dorms and classrooms. Panel A describes the allocation when each room holds four students, Panel B when dorm rooms hold three students and Panel C for a big dorm room of 12 students. The order on the list is used to determine the allocation. When dorm rooms hold four students, dorms and *groups of roommates* have the same size so there is perfect compliance with the initial allocation. Yet, when dorm rooms hold three students, the first three students of type H –who were assigned to Group 1 (only Hs)—are allocated to room 1. The fourth student assigned to Group 1 is assigned to a room with two students who are type L. Notice that in this case there is no perfect compliance. In the big dorm, all students $H_1 - H_4$ are in beds close to each other. However, students H_3 and H_4 are also close to students of type L, while students H_1 and H_2 are surrounded only by peers who are type H.

FIGURE B.1: Illustrative Examples of the Allocation to Dorms

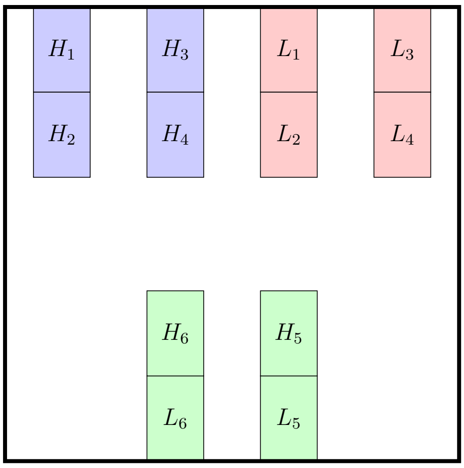
Panel A: Dorms of 4 Students



Panel B: Dorms of 3 Students



Panel C: Big Dorm of 6 Bunk Beds



Student:	H_1	H_2	H_3	H_4	L_1	L_2	L_3	L_4	H_5	L_5	H_6	L_6
3 dorms of 4 students:	D1	D1	D1	D1	D2	D2	D2	D2	D3	D3	D3	D3
4 dorms of 3 students:	D1	D1	D1	D2	D2	D2	D3	D3	D3	D4	D4	D4

Notes: This figure displays three examples of how the randomization to groups was used to allocate students to dorm rooms and classrooms.

C Psychological Tests

This section describes in detail the psychological tests that were used to construct the sociability index.

In addition to the Big Five personality traits and the peers' perceptions measures described in section 5.1, the tests used to construct the sociability index are:

- **Altruism:** The altrusim self-reported scale was developed by Rushton et al. (1981). The test used in the COAR network is composed of 17 items. The score on the test is found to predict criteria such as peer ratings of altruism, completing an organ donor card, and paper-and-pencil measures of prosocial orientation (Rushton et al., 1981). More recent evidence shows that the score on the test is related to spontaneous smiles—which is an important signal in the formation and maintenance of cooperative relationships (Mehu et al., 2007). Likewise, there is evidence that the score in the test is related to charity giving but not to blood donor donation behavior (Otto and Bolle, 2011).

- **Leadership:** The leadership scale corresponds to the leader behavior questionnaire developed in Spanish by Castro-Solano (2007). It is based on the theory of Yukl (2013). The scale measures three components of leadership: (1) behaviors guided towards tasks, (2) behaviors guided towards others, and (3) behaviors guided towards changes. In my data, there is a positive correlation between the score on the scale and the number of peers who perceived the subject as a leader.

- **Empathy:** The empathy scale corresponds to the Basic Empathy Scale developed by Jolliffe and Farrington (2006). The scale is composed of two factors: cognitive and emotional empathy. The scale has been validated in other contexts: when applied to adults (Carre et al., 2013) and the Spanish version of it (Villadangos et al., 2016). It has also been affirmed that students who report higher scores in socially aversive personalities (psychopathy, narcissism, and Machiavellianism) have a low score on the scale (Wai and Tiliopoulos, 2012). Likewise, Gambin and Sharp (2018) show that a low score on the test is associated with guilt and depressive symptoms.

- **Intercultural Sensitivity:** This 24-item scale of intercultural sensitivity was developed by Chen and Starosta (2000). The authors define intercultural sensitivity as: "*a person's ability to develop a positive emotion towards understanding and appreciating cultural differences that promotes appropriate and effective behavior in intercultural communication*." The scale is composed of two factors: positive and negative reactions to intercultural interactions. Evidence shows that there is a positive correlation between intercultural sensitivity and compassion in nurses (Arli and Bakan, 2018), that American student scores depend on religious affiliation and the number of times they have traveled outside the US (Gordon and Mwavita, 2018), and that Iranian university students have demonstrated a strong relationship between intercultural sensitivity and ethnic background.

- **Emotional Intelligence:** Emotional intelligence is defined as individuals' ability to recognize their own emotions and those of others, discern between different feelings and label them appropriately, use emotional information to guide thinking and behavior, and manage and/or adjust emotions to adapt to environments or achieve one's goal(s) (Colman, 2009). The emotional intelligence test corresponds to the scale developed by Law et al. (2004). The test is composed of 16 items and has four factors: self-emotional appraisal, uses of emotion, regulation of emotion, and others' emotional appraisal.

- **The Read the Mind in the Eyes Test:** The objective of this test is to assess how well people can read others' emotions just by looking at pictures of their eyes. It is a multiple choice test with 36 items. For each item, the respondent has to identify the corresponding emotion expressed in a pair of eyes; four choices are given for each question. According to Deming (2017), this test is a reliable measure of social skills since it is positively correlated with performance in groups (Declerck and Bogaert, 2008). However, this measure could potentially have problems due to the cultural differences between the context where the test was developed and the context of my study. In particular, according to the website Lab in the Wild: `socialintelligence.labinthewild.org`, the test was developed in Great Britain and the images were taken from British magazines in the 1990s. Therefore, the test may not produce accurate results when administered to people who are not native speakers of English or those who come from cultures that are very different from Britain's.

While not part of the construction of the sociability index, students completed other assessments:

- **Achievement Goals:** Achievement goals are conceptualized as cognitive–dynamic aims that focus on competence. Students answered the *The Achievement Goal Questionnaire* (J. Elliot and Murayama, 2008). The test is composed of 12 items and has four factors: mastery approach goal items, mastery avoidance goal items, performance approach goal items, and performance avoidance goal items. The last two items are related to goals in comparison with peers.

D Proofs

Proof of Proposition 1: In equilibrium we have that:

$$\phi_r^* = \kappa(\phi_r^*)\mu_r, \tag{1a}$$
$$\phi_o^* = \kappa(\phi_r^*)\mu_o. \tag{1b}$$

To show (1) and (2) in the proposition, we apply the implicit function theorem to 1a and 1b:

$$\frac{d\phi_r^*}{d\mu_r} = \kappa'(\cdot)\mu_r\frac{d\phi^*}{d\mu_r} + \kappa(\cdot), \tag{2a}$$
$$\frac{d\phi_o^*}{d\mu_r} = \kappa'(\cdot)\mu_o\frac{d\phi^*}{d\mu_r}, \tag{2b}$$

where $\frac{d\phi^*}{d\mu_r} = \gamma\frac{d\phi_r^*}{d\mu_r} + (1-\gamma)\frac{d\phi_o^*}{d\mu_r}$.

Solving this system of equations yields:

$$\frac{d\phi_r^*}{d\mu_r} = \frac{\kappa(\cdot) + \kappa'(\cdot)\mu_r(1-\gamma)\frac{\gamma\kappa(\cdot)\kappa'(\cdot)\mu_o}{1-(\gamma\kappa'(\cdot)\mu_r+(1-\gamma)\kappa'(\cdot)\mu_o)}}{1-\kappa'(\cdot)\mu_r}, \tag{3a}$$
$$\frac{d\phi_o^*}{d\mu_r} = \frac{\gamma\kappa(\cdot)\kappa'(\cdot)\mu_o}{1-(\gamma\kappa'(\cdot)\mu_r+(1-\gamma)\kappa'(\cdot)\mu_o)}. \tag{3b}$$

Since we assumed that $\rho_a'(\cdot) < 1$ for $a = \{r, o\}$. Then, we know that $\rho_r'(\cdot) = \kappa'(\cdot)\mu_r < 1$. Likewise, $\rho_o'(\cdot) = \kappa'(\cdot)\mu_o < 1$. This implies that $\gamma\kappa'(\cdot)\mu_r + (1-\gamma)\kappa'(\cdot)\mu_o < 1$. This assures that the denominators in both 3a and 3b are greater than zero. Notice that if $\rho_r'(\cdot) > 1$, then the right hand side of 2a is larger than the left hand side, and we are in a corner solution where $\phi_r^* = 1$. Similarly, in case $\gamma\rho_r'(\cdot) + (1-\gamma)\rho_o'(\cdot) > 1$, then the right hand side of 2b is greater than the left hand side and $\phi_o^* = 1$.

Since $\kappa(\cdot) \in (0,1)$, and it is an increasing function in ϕ, then $\kappa(\cdot) > 0$, $\kappa'(\cdot) > 0$. Similarly, μ_r and μ_o are probabilities: $\mu_r \in (0,1]$, and $\mu_o \in (0,1]$. Hence, the numerators in both expressions 3a and 3b are greater than zero. This guarantees that:

$$\frac{d\phi_r^*}{d\mu_r} > 0, \tag{4a}$$
$$\frac{d\phi_o^*}{d\mu_r} > 0. \tag{4b}$$

Hence, (1) and (2) of proposition 1 hold.

For (3), we have that $\kappa(\phi)$ is an increasing function in ϕ. We have that:

$$\phi^* = \gamma\phi_r^* + (1-\gamma)\phi_o^*$$

$$\frac{d\phi^*}{d\mu_r} = \gamma\frac{d\phi_r^*}{d\mu_r} + (1-\gamma)\frac{d\phi_o^*}{d\mu_r}.$$

Since $\gamma \in (0, 1)$, and $\frac{d\phi_r^*}{d\mu_r} > 0$, $\frac{d\phi_o^*}{d\mu_r} > 0$, then $\frac{d\phi^*}{d\mu_r} > 0$. Since $\kappa(\phi)$ is an increasing function in ϕ, then $\frac{d\kappa(\phi^*)}{d\mu_r} = \kappa'(\cdot)\frac{d\phi^*}{d\mu_r} > 0$.

Proof of corollary 1: The likelihood that the agent invests in equilibrium is given by $G(\phi_a^*)$. Since $G(\cdot)$ is a c.d.f. with no flat regions then, $G(\cdot)$ is a strictly increasing function and $\phi_a^*(\cdot)$ increases with μ_r for $a = \{r, o\}$. Hence, the agent invests more in both types of activities.

Proof of Proposition 2: Under the assumption of social success, for less sociable students $\mu_r(l, l) < \mu_r(l, h)$. By applying proposition 1 and corollary 1 we obtain the result. For more sociable students $\mu_r(h, l) = \mu_r(h, h)$, and therefore they do not experience changes in self-confidence or investment decisions.

Proof of Proposition 3: Under the assumption of cognitive success, for lower-achieving students $\mu_r(l, h) < \mu_r(l, l)$. By applying proposition 1 and corollary 1 we obtain the result. For higher-achieving students $\mu_r(h, l) = \mu_r(h, h)$, and therefore they do not experience changes in self-confidence or investment decisions.

E Network Formation Model

This section explores the impact of the allocation of students on network formation, including dependencies in the formation of links.

E.1 Exponential Random Graph Models (ERGM)

For the descritpion of this section, I follow Handcock et al. (2018) and Handcock et al. (2016)[20]. Exponential random graph models (ERGMs) are a family of statistical models, which are suitable for the analysis of data on social networks. The assumption behind these models is that the structure of a network graph Y can be explained by any statistic $g(Y)$, such as the number of ties, centrality measures, and so on. The ERGM assigns a probability to graphs on the basis of these statistics. The general form for an ERGM can be written as:

$$P(Y=y)=\frac{exp(\theta' g(y))}{k(\theta)} \tag{5}$$

where Y is the random variable for the state of the network —with realization y—, and $g(y)$ is a vector of statistics for the network; θ is a vector of associated coefficients to those statistics, and $k(\theta)$ is a normalizing constant that represents the quantity in the numerator over all possible realizations of network Y.

Equation 5 can be re-expressed in terms of the conditional log-odds of a single tie between two students as:

$$logit(Y_{ij}=1|y^c_{ij})=\theta'\delta(y_{ij})),$$

where Y_{ij} is the random variable for the state of the dyad between students i and j; y^c_{ij} represent all the other dyads in the network other than y_{ij}. The vector $\delta(y_{ij}))$ contains the "change statistic" for each model term $g(y)$.

$$\delta(y_{ij})=g(y^+_{ij})-g(y^-_{ij})$$

where y^+_{ij} gets determined by y^c_{ij} and setting y_{ij} to 1, and y^-_{ij} is defined by y^c_{ij} and settinh y_{ij} set to 0. That is, $\delta(y_{ij}))$ equals the value of $g(y)$ when $y_{ij}=1$ minus the value of $g(y)$ when $y_{ij}=0$, but all other dyads are as in $g(y)$. Therefore the coefficient θ can be interpreted as the log-odds of an individual tie conditional on all others.

I perform the estimation using the package statnet in R[21]. In my analysis, I consider the following set of statistics as part of the set $g(Y)$: the number of edges in the network, the number of neighbors, and the interaction between the number of neighbors with treatments and baseline characteristics. I estimate a separate model by the baseline classifica-

[20] This tutorial can be accessed: https://statnet.org/trac/raw-attachment/wiki/Sunbelt2016/ergm_tutorial.html.

[21] http://statnet.org/

tion of: (i) less and more sociable students, and (ii) lower- and higher- achieving students.

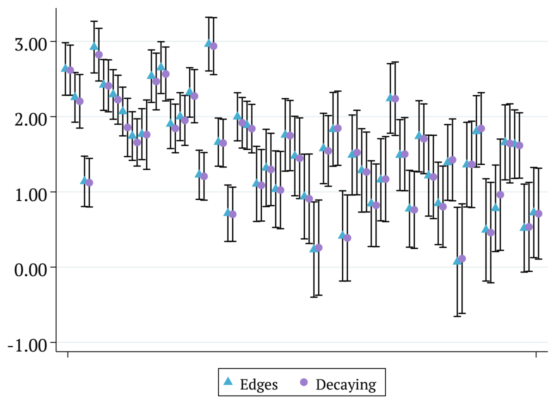
$$g_1(y) = \{edges, neighbors \times treatments \times baseline\}$$

In addition to these terms, I estimate a second model that includes a geometrically-weighed edgewise shared partner term that accounts for link dependencies.

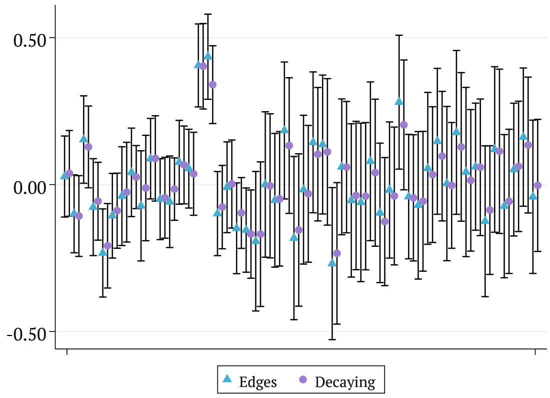
Figures E.1 and E.2 show the estimates for the less sociable students, and lower-achieving students, respectively. The figure presents three panels, one for each coefficient. Panel A shows the results for the neighbor, and panels B and C for heteroeneous effects of neighbors by each treatment status. Overall, I find very similar estimates regardless of including or not the link-dependencies term.. Therefore, the conclusions of social interactions are robust to link dependencies. There is a substantial impact of being neighbors on the likelihood of forming social interactions and there are no heterogeneous impacts by treatment neither for the less sociable nor the lower-achieving students.

Figure E.1: Estimates of ERGM by Baseline Sociability

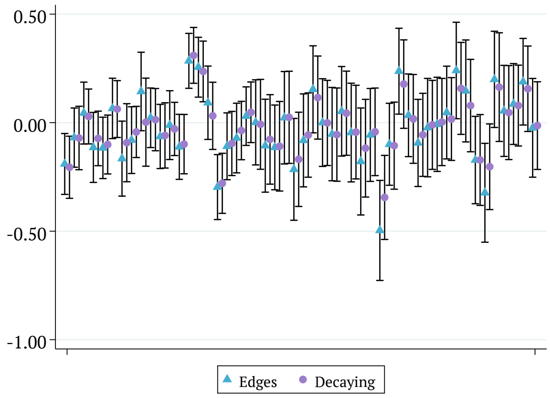
Panel A: The coefficient of neighbors
for less sociable students



Panel B: The coefficient of neighbors
× more sociable peers
for less sociable students

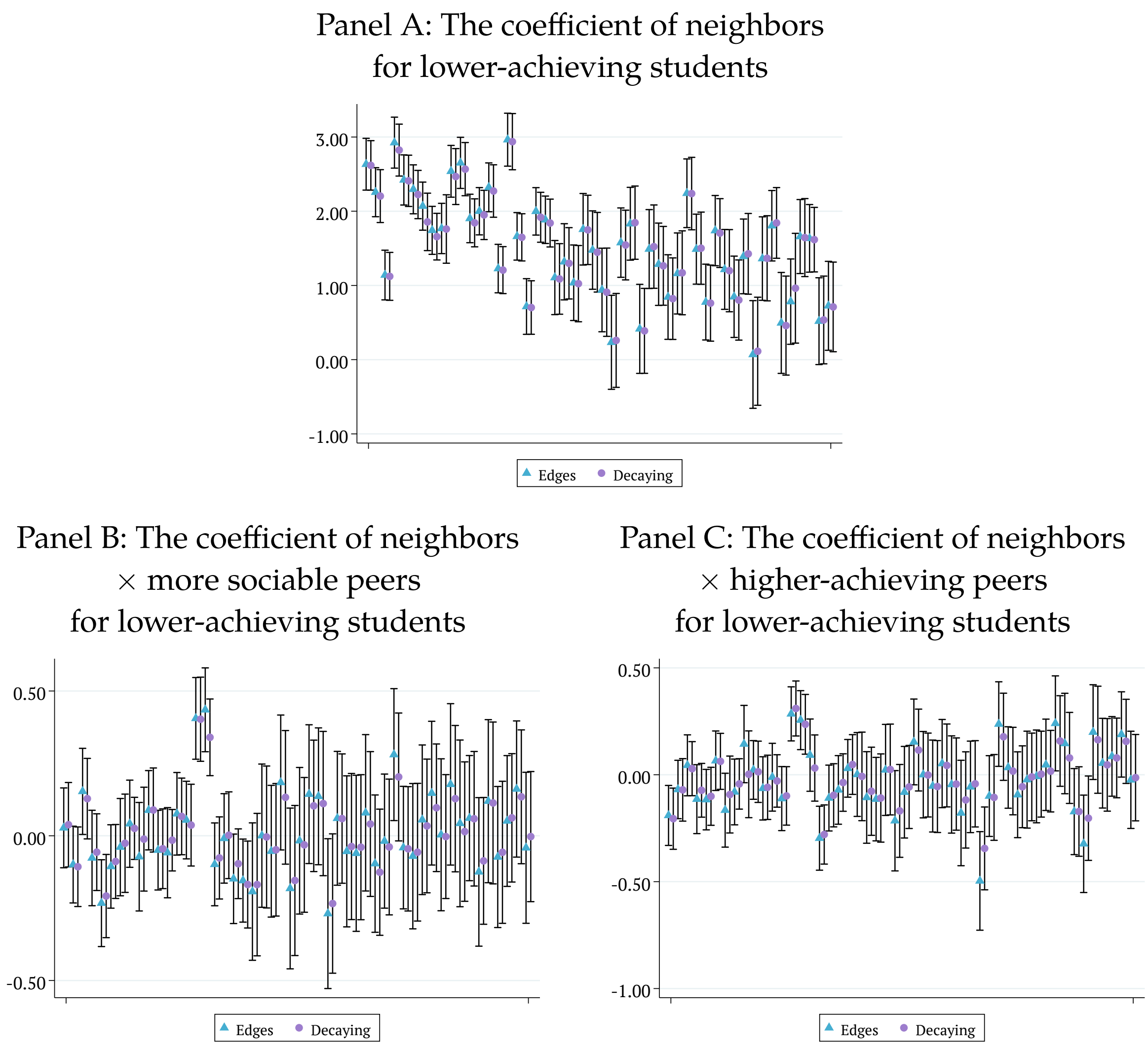


Panel C: The coefficient of neighbors
× higher-achieving peers
for less sociable students



Notes: This figure shows the estimates of an ERGM for each network. 95% confidence intervals are displayed for all estimates. The model includes the number of edges in the graph, and the neighbor coefficient interacted with treatments status and baseline characteristics of sociability. The plot presents estimates for two models: one with edges (blue circles) without link dependencies, and a second model including link dependencies with a decaying geometrically-weighed edgewise shared partner term of 0.1 (purple circles). The network aggregates whether students are friends, study together, or engage in social activities such as playing or dancing.

Figure E.2: Estimates of ERGM by Baseline Academic Achievement



Notes: This figure shows the estimates of an ERGM for each network. 95% confidence intervals are displayed for all estimates. The model includes the number of edges in the graph, and the neighbor coefficient interacted with treatments status and baseline characteristics of sociability. The plot presents estimates for two models: one with edges (blue circles) without link dependencies, and a second model including link dependencies with a decaying geometrically-weighed edgewise shared partner term of 0.1 (purple circles). The network aggregates whether students are friends, study together, or engage in social activities such as playing or dancing.