Trust and the Dynamics of Network Formation

Juan Camilo Cárdenas, Danisz Okulicz,
Davide Pietrobon, Tomás Rodríguez, Tatiana Velasco*

Abstract

We evaluate the effect of reciprocal trust within pairs of individuals—gauged by total potential earnings in a trust experiment—on the probability of relationship formation, in comparison with well-known determinants of social ties, such as time of exposure and homophily along demographic traits. We measured trust and trustworthiness for every individual in an incoming cohort of undergraduate students before they began interacting. Using relationship data sourced from surveys and campus entry/exit times between one month and two years after the trust experiment, we find that reciprocal trust is neither a statistically nor an economically significant factor in determining the students' social networks. Instead, time of exposure, prior acquaintance, and other demographic characteristics play important and persistent roles in relationship formation.

Keywords: trust, prosocial beliefs, network formation, social capital.

JEL Classification: C83 C91 D85 Z13

*Cárdenas: Universidad de los Andes and University of Massachusetts Amherst, juancamilo@umass.edu. Okulicz: Karlsruhe Institute of Technology, d.okulicz@gmail.com. Pietrobon: Lund University, d.pietrobon91@gmail.com. Rodríguez: Universidad de los Andes, t.rodriguezb@uniandes.edu.co. Velasco: Teachers College - Columbia University, tv2225@tc.columbia.edu. We thank Francesco Bogliacino, Gary Charness, Julien Daubanes, Giacomo De Giorgi, José Alberto Guerra, Dunia López-Pintado, Matthew Jackson, Eric Quintane, Alessandro Tarozzi, Michela Tincani, and Leeat Yariv for their comments. We thank Anastasia Agafnikova, Juan Camilo Cháves and Lucas Gómez for excellent research assistance. We thank seminar participants at Universitat Autónoma de Barcelona, Universidad del Rosario, Universidad Nacional de Colombia, LANET Conference 2018, and University of Geneva. Okulicz acknowledges the financial support of the fellowship BES-2016-077805 and the research project from the Ministerio de Economía, Industria y Competitividad–Feder (ECO2015-63679-P). Pietrobon acknowledges the financial support of Fundació Markets, Organizations and Votes in Economics (BES-2015-075756) and the SNF (100018-182243). Rodríguez acknowledges the financial support of Universidad de los Andes through FAPA. Velasco acknowledges partial financial support from the National Academy of Education / Spencer Foundation Dissertation Fellowship. Declarations of interest: none.

Social networks are a fundamental aspect of human life and influence many economic situations, such as peer effects, information transmission, and job search (Jackson, Rogers and Zenou, 2017; Bailey et al., 2018). Several studies in the network formation literature show that we are more likely to befriend individuals who are similar to us in various characteristics (homophily) and those we are exposed to by chance (McPherson, Smith-Lovin and Cook, 2001; Marmaros and Sacerdote, 2006; Currarini, Jackson and Pin, 2009). Despite significant advances in understanding how relationships form (see Jackson et al. (2023)), much remains to be established regarding how individual and pairwise characteristics interact in the creation of social connections. In this paper, we examine the role of reciprocal trust—the degree to which trust from each individual in a pair is reciprocated by the other—in shaping network formation.

The willingness of individuals to trust others is frequently argued to play a pivotal role in facilitating cooperation and fostering the formation of social capital (Putnam (1995)). The idea is that when this trust is met with trustworthiness (i.e., it is reciprocated) it allows a pair to generate surplus that would otherwise remain unrealized. High-trust individuals hold more positive beliefs about interactions with strangers and are more willing to invest in potential reciprocation. High *reciprocal* trust between a pair of individuals occurs when both members exhibit high levels of trust and this trust is validated by each other's trustworthiness. According to the trust and social capital narrative, such dyads should be uniquely positioned to achieve greater benefits from their interactions, even amidst uncertainties about the returns to cooperation and in the absence of external enforcement mechanisms. Due to their enhanced capacity to capitalize on strategic interactions, pairs with higher reciprocal trust should be more likely to establish relationships, all else being equal. In this paper, we provide a rigorous empirical test of this hypothesis.

We investigate whether pairs' reciprocal trust—the extent to which the trust of each individual in the pair is reciprocated by the other—is a determinant of social networks formation among an incoming cohort of first-year undergraduate students at a university in Bogotá, Colombia. Our findings suggest that pairs' reciprocal trust—as measured by the the sum of the amounts that the individuals in each pair would have received in the role of sender when interacting with each other

in a trust experiment—plays a negligible role in the formation of social ties among our subjects. We fail to reject the null hypothesis that reciprocal trust does not affect link formation probabilities, and we retain sizeable power when doing so (conditional on the hypothesis that the true effect of reciprocal trust is comparable in size to other significant determinants of social network formation).

On the other hand, our results demonstrate that time of exposure (measured by the number of course credits shared between students), previous acquaintance, and several demographic characteristics significantly influence network formation. We find that a one standard deviation increase in the shared number of course credits due to class assignment[1] is associated with a 7 to 10 percentage point increase in the likelihood of friendship formation. Prior acquaintanceship also increases the probability of forming new relationships. Finally, our study uncovers a distinct pattern of homophily based on socioeconomic status and hometown within the students' networks, highlighting the significant role demographic traits play in shaping network structures. Overall, our results point to a picture where relationships are more the outcome of chance and demographics than the result of pairs' reciprocal trust. These results provide insights on improving integration between people starting higher education in the presence of segregation and socioeconomic differences.

This paper uses data from an entire incoming cohort of first-year economics undergraduate students at a university in Bogotá. The data was collected in two stages. In the first stage, we asked each of the students comprising the entire cohort to participate in activities to measure their trust and trustworthiness towards strangers before they had significant chances to get to know each other and socialize. This feature of the data collection strategy allows us to avoid the possibility of reverse causality from relationships to pairs' reciprocal trust.[2] Specifically, we conducted our measurements of trust and trustworthiness on the university welcome day, which is the first day in which students formally attend the university campus.[3] These activities comprised (1) a trust

[1] All of our subjects take the same six courses in the first semester but are allocated to different class sections for each of them, corresponding to different instructors, locations or time slots. As a result, any two students have a number of credits that are "shared" because they are assigned to the same sections in some of their courses.

[2] For example, friendships with prosocial people may be associated with an increase in one's prosocial behavior. Thus, if we measured the students' trust and trustworthiness after relationships between them were established, we could not have ruled out the possibility that those particular relationships helped shaping their trust and trustworthiness towards strangers.

[3] In the second stage of the data collection, we asked each subject to name each of the other participants that he or

experiment, taken from Berg, Dickhaut and McCabe (1995), and (2) two survey questions adapted from Glaeser et al. (2000).[4] We focus on trust and trustworthiness because they allow us to construct a pairwise-specific measure of "reciprocal trust" which measures the extent to which the trust of each individual in a pair is reciprocated by the other individual (see Subsection 2.1).

In the second stage of the data collection process—conducted at the end of the first academic semester (i.e., four months after the measurement of trust and trustworthiness)—we administered a survey to elicit five types of social networks representing different relationships (greeting, having lunch together, studying together, confiding in, and friendship). This sample of students provides a dataset encompassing 1,485 potential undirected connections, a measure of reciprocal trust for each pair, and various individual and dyadic factors that are likely to play a critical role in network formation.

We also utilize administrative data from turnstiles at all entry and exit points on the university campus, which record students' entry and exit times. This data allows us to construct measures of the students' social networks at various points in time. Following the methodology outlined by Velasco (2023), we classify a pair of students as linked if they swipe their university IDs at the same campus entrance, in the same direction (either entering or exiting), within a time window of three seconds or less, and if this pair of IDs is observed entering or exiting the campus together at least twice within an academic semester. These methodology allows us to track the development of students' networks in the periods following our survey on network elicitation and to explore the degree to which reciprocal trust influences the subsequent characteristics of these networks. We additionally use this data to investigate whether reciprocal trust might explain the students' social networks prior to our network elicitation survey. To construct the short-term networks, we adopt a more flexible definition of turnstile-based interactions than Velasco (2023)'s, considering pairs of students who are recorded moving together within a three-second window at least once a month.

she knew from before starting university. Hence, we can control for whether subjects knew each other from before the university welcome day.

[4]Albeit our survey questions have been experimentally validated (Glaeser et al., 2000), we measured trust and trustworthiness using both a lab experiment and survey questions hoping that the use of different data collection methods would reduce measurement error.

This approach allows us to track and analyze monthly interactions from the date of the experiment until December 2017.

We estimate linear probability models (LPMs) to identify how pairs' reciprocal trust, demographic characteristics, and exogenous variation in time of exposure predict link formation probability in the networks elicited.[5] The estimates of reciprocal trust on link formation probabilities are negative and statistically insignificant. Moreover, we can safely assert that reciprocal trust does not have an impact on link formation probability as quantitatively meaningful as other characteristics, such as time of exposure, knowing each other from before, hometown, and differences in socioeconomic status.[6] Using turnstile data to analyze the impact of reciprocal trust on relationship formation over time, our findings confirm that reciprocal trust is neither a statistically nor an economically significant factor in determining students' social networks, both in the short term and the long term.

Related literature

This paper contributes to the expanding body of research on the empirical determinants and dynamics of network formation (Jackson et al. (2023)). A common theme within this field is homophily along demographic lines, whereby individuals tend to form connections with others who share similar characteristics (see Jackson (2010), Bramoullé, Galeotti and Rogers (2016), and Jackson, Rogers and Zenou (2017)). We contribute to this endeavor by studying the link between pairs' reciprocal trust and relationship formation. We focus on trust and trustworthiness due to their perceived importance in enhancing pairs' abilities to cooperate in social dilemmas (Putnam (1995)). Broadly, our analysis confirms the importance of homophily in shaping networks, highlighting segregation based on socioeconomic status and hometown across various student social

[5]We use dyadic-robust standard errors to account for possible correlation between unobservables affecting link formation. See Giné et al. (2010), Santos and Barrett (2011), and Giné and Mansuri (2018) for applications of QAP in economics.

[6]For example, we retain more than 99% power when failing to reject the null that reciprocal trust does not affect relationship formation probability, conditional on the hypothesis that the true effect of the reciprocal trust has the same size as the effect of the time of exposure.

networks. Conversely, we find that reciprocal trust between pairs—measured by total potential earnings in a trust experiment—plays a negligible role in the formation of relationships among students.

We also speak to the social capital literature, which frequently relates trust and social networks. Trust has often been bundled into the very definition of social capital.[7] Other times, measures of trust have been used as proxies for social capital. While trust and networks might both play a role in determining social capital, we contribute to this literature by shedding light on the interconnection between pairs' reciprocal trust and network formation. Having more relationships, or being embedded in social networks with certain structural properties, may encourage people to trust more (Buskens (1998) and Jackson, Rodriguez-Barraquer and Tan (2012)). Kosse et al. (2020) provides causal evidence on the positive effect of enriching a person's social environment on his or her trust. We contribute to this literature by examining the link between pairs' reciprocal trust and relationship formation in real-life networks among students over an extended timeframe, including months and years after the trust game has been played.[8]

Lastly, our paper contributes to a growing body of literature examining the causes of segregation in education, including that occurring within institutions. First, our finding that pre-existing networks strongly predict social interactions complements prior research that finds high school networks largely explain students' participation in social spaces like college clubs (Michelman, Price and Zimmerman, 2022). Second, we contribute to the evidence on how exposure to peers shapes social interactions within college. Our finding that exposure to peers significantly increases the chances of social interactions in the short- and long-terms in college, even after controlling for other student and dyadic characteristics, complements those from Marmaros and Sacerdote (2006); Baker, Mayer and Puller (2011); Mayer and Puller (2008), and represents new evidence on

[7]For example, Inglehart (1997) defines social capital as "a culture of trust and tolerance, in which extensive networks of voluntary associations emerge." Putnam (1995) defines it as "features of social organization such as networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit." Woolcock (1998) defines it as "the information, trust, and norms of reciprocity inhering in one's social networks."

[8]Di Cagno and Sciubba (2010) analyzes the relationship between trust and network formation through a trust game and a network-connection game, finding that measured trust levels are higher when network formation follows the trust game than when the trust game follows network formation. They attribute the difference to the idea that continuation play induces players to trust each other.

the persistence of relationships formed early in college.[9]

1 Design and protocols

We collected data from incoming first-year undergraduate students choosing economics as their major at a university in Bogotá. Our design consisted of two stages. We conducted the first stage on August 4, 2017, and its main goal was to measure the students' trust and trustworthiness. Crucially, we carried out this stage on the university welcome day, which is the very first day in which incoming students formally attend the university campus. The rationale behind this choice was to measure the students' trust and trustworthiness before they had significant opportunities to socialize, to avoid the possibility of reverse causality from relationships to trust and trustworthiness. We conducted the second stage online between December 7, 2017, and January 5, 2018, at the end of the first academic semester, and its main aim was to elicit social networks among the students. In what follows, we describe the design of the two stages in detail.

Our sample choice aimed at three goals. First, we chose a group of people for whom we could accurately measure trust and trustworthiness before they had significant opportunities to socialize. Second, we wanted our subjects to have many chances to get to know each other over an extended period of time after the measurement of trust and trustworthiness. Finally, we selected people for whom we could collect detailed information on many characteristics, at both the individual and the relationship level. Our strategy allows us to obtain measures of the subjects' trust and trustworthiness in a controlled setting and gather precise information on numerous other variables of interest.

First stage. We directed the first stage to the 81 students comprising the entire incoming undergraduate economics cohort of the first semester of 2017, and its main goal was to measure their

[9] Mayer and Puller (2008); Baker, Mayer and Puller (2011) use Facebook data and to document the factors related to network formation, finding race is a strong predictor of social interactions. Marmaros and Sacerdote (2006) uses random dorm assignment and detailed data on students' communications and finds exposure to peers in dorms is a strong predictor of interactions, particularly along racial lines.

trust and trustworthiness. We conducted this stage in a single session on the university welcome day. The session lasted 90 minutes. Out of the 81 intended subjects, 72 were present on the welcome day. All of these 72 students agreed to participate in the experiment. We gave each student four paper handouts labeled A, B, C, and D. The Online Appendix contains an English translation of the handouts. Handout A is a general description of the activity and an informed consent form that we required the students to complete for participating in the session. Handout B is a detailed description of the trust experiment. Handout C is a form for recording the students' strategies in the experiment. Finally, handout D is a questionnaire with eight questions on generalized trust, particularized trust towards friends and neighbors, and particularized trustworthiness towards friends and neighbors,[10] and six (1-3 and 6-8) questions on individual characteristics.[11]

In the trust experiment, we endowed every participant with $e = 20,000$ Colombian Pesos ($\$COP$) (about $USD\$7$). In every anonymously created sender-receiver pair, each sender had to decide how much money s to transfer to the receiver in a range from 0 to e in $\Delta := COP\$2,000$ increments. For each possible s chosen by the sender, the receiver would receive $3s := r(s)$; i.e., three times the money sent to him or her by the sender. The receiver had to decide how much money to send back to the receiver, $f(r(s))$, for each possible s he or she could have received, following the convention of the strategy method in the trust game. For each s, the sender could send back any amount in a range from 0 to $r(s)$ in Δ increments. The monetary payoffs at the end of the game for a sender-receiver pair in which the sender uses strategy s and the receiver uses strategy $f(r(s))$ are $e - s + f(3s)$ to the sender and $e + 3s - f(3s)$ to the receiver.

We described the two roles in the trust experiment to all participants. We informed them that each had to report how they would behave both as a sender and as a receiver, as we would then assign these roles randomly,[12] and randomly match senders and receivers to implement their re-

[10]The psychology literature refers to trust towards strangers as generalized trust and distinguishes it from particularized trust, which is trust directed towards particular groups of people. Unless otherwise noted, we refer to trust towards strangers simply as trust, as it is customary in the economics literature.

[11]These were sex, age, number of siblings, number of friends outside the university, number of people in the incoming cohort of first-year economics undergraduate students knew from before the first day of university, and self-assessed happiness on a 0–3 scale.

[12]After we collected the subjects' choices, we randomly assigned half of the subjects to the role of sender and the

ported strategies and realize monetary payoffs.[13] Handout B included instructions for the strategies available to the sender and the receiver, the functions used to calculate the monetary payoffs, and a detailed example. We read out loud the instructions and the example and conducted a question-and-answer session right afterward. We then instructed the students to fill out handout C, which contained the strategy sets for the sender and the receiver. Specifying the strategy for the role of sender entailed stating one among 11 (0-10) transfer options in Δ units. Specifying the strategy for the role of receiver entailed stating 11 *contingent* transfers, one for each of the 11 possible amounts received from the sender. For each possible amount that he or she might receive, the receiver could choose to send back to the sender an amount ranging from 0 to the entire amount in Δ units.

After the experiment, the students filled out a survey contained in handout D. First, the survey contained eight questions aimed to measure generalized trust, particularized trust towards friends and neighbors, and particularized trustworthiness towards friends and neighbors. We report the questions below.

4. To what extent do you agree with the following statements (on a 1–5 scale, where 1 denotes total disagreement and 5 total agreement):

a. One cannot trust strangers.

b. When dealing with strangers it is important to be careful and not to readily trust them.

5.a. How many among your 10 closest friends have you lent money to?

5.b. How many among your 10 closest friends have lent money to you?

5.c. To how many among your 10 closest friends have you lent your belongings (e.g., books, CDs, clothing, bicycle)?

other half to the role of receiver. In case of an odd number of participants (this was not the case) we planned to include an extra artificial player relying on a pair of strategies (one for each role) randomly chosen from the set of strategy pairs submitted by the subjects.

[13] We chose to have every participant assume both roles to record the behavior of as many senders as possible. No student knew during or was revealed after the experiment to whom he or she was paired with. To avoid the possibility that the monetary payoffs realized in the experiment could contaminate the formation of relationships between the subjects, we only informed them about the payoffs realized and made the payments four months after the experiment.

5.d. How many among your 10 closest friends have lent their belongings (e.g., books, CDs, clothing, bicycle) to you?

5.e. How many among your 10 closest neighbors would you trust with your house keys?

5.f. How many among your 10 closest neighbors would trust you with their house keys?

Questions 4.a and 4.b measure generalized trust, questions 5.a and 5.c measure particularized trust towards friends, question 5.e measures particularized trust towards neighbors, questions 5.b and 5.d measure particularized trustworthiness towards friends, and question 5.f measures particularized trustworthiness towards neighbors. We adapted all the questions aimed to measure generalized and particularized trust from Glaeser et al. (2000). However, note that what we refer to as questions measuring particularized trust (questions 5.a, 5.c, and 5.e), Glaeser et al. (2000) identifies as questions measuring past trusting behavior. We think of these questions as aimed to measure particularized trust because they explicitly refer to particular groups of people (i.e., friends and neighbors) to which trust is directed, instead of unknown individuals (i.e., strangers). Our aim in collecting this information was to have additional (non-lab) measures of trust and trustworthiness that we could use in combination with our main (lab) measures of trust and trustworthiness to reduce possible measurement error concerns. The survey also included five questions on demographic characteristics (sex, age, number of siblings, number of friends outside the university, number of people in the cohort that the person knew from before starting university) and one question on self-assessed happiness.

Second stage. We conducted the second stage of the data collection process four months after the first stage (i.e., at the end of the first academic semester), and its goal was to elicit some of the networks of relationships among the students comprising the entire incoming cohort of 2017. Additionally, we asked the participants questions on individual characteristics. We sent emails to the students asking them to complete an incentivized survey.[14] We elicited social networks as

[14]We offered each respondent a $COP\$20,000$ (about $USD\$7$) voucher for a fast food restaurant on the university campus. The Online Appendix includes an English translation of the survey, which was conducted using Qualtrics.

follows. First, we presented each student with the list of names of the other students invited to complete the survey (in random order), and we asked him or her to indicate the students who he or she greeted (henceforth, hello partners). Specifically, for each student on the list, we asked him or her to tick a box if they would say hi to that student upon encountering him or her. Secondly, we presented each student with his or her list of hello partners and, for each of them, we asked the student to check one or more of six boxes acknowledging the following relationships: (1) "I met this person before starting university," (2) "I frequently have lunch with this person," (3) "I frequently study or work together with this person," (4) "I share my personal feelings with this person," (5) "I believe this person is a friend of mine," and (6) "None of the previous options apply to my relationship with this person."[15] Thanks to box (1) we can control for whether relationships formed before our intended socialization period (the first academic semester), and so we end up with five possible relationships (greeting, having lunch together, studying together, confiding in, and friendship).

Besides questions to elicit networks, the survey included questions on many individual characteristics that we suspect to play a role in relationship formation. The rationale behind this design is that isolating the impact of reciprocal trust on network formation requires controlling for variables that might affect the creation of social links and correlate with trust and trustworthiness. In particular, we collected information on the number of siblings, the number of friends enrolled in the same university met before starting university, the number of friends enrolled in the same university met after starting university, the number of friends not enrolled in the same university, weekly hours spent socializing with friends enrolled in the same university, weekly hours spent socializing with friends not enrolled in the same university, weekly hours spent doing physical activities, hobbies, age, eye color, hair color, height, weight, whether wearing glasses, whether wearing tattoos, whether wearing piercings, whether smoking, whether attending parties, whether their hometown is Bogotá, and four personality questions. In the latter questions, we asked the students to rate on a scale from 1 to 5 how much they perceived themselves as realistic, introverted, inhibited, and

[15] We decided to ask the students to check these boxes only for their hello partners to avoid overwhelming them with questions, which could have increased the likelihood of them opting not to complete the survey.

shy. Finally, we asked the students to rate on a scale from 1 to 5 how much they agreed with the following statements: "I am very sociable," "I am satisfied with my social life," "making friends at university is easier than I thought," and "I am satisfied with the number of friends I have." In addition to the data collected with our survey questions, our empirical analysis uses administrative data from the university on several student characteristics, such as the scores obtained at the high school exit examination, their GPAs at the end of the first academic semester, and their socioeconomic status. Moreover, we use the administrative data to obtain information on the time that each pair of students are exposed to each other because assigned to the same classrooms during the first semester.

Out of the 81 students comprising the entire cohort, 72 participated in the activities to measure their trust and 70 out of the 72 provided complete answers to the trust questionnaires.[16] Out of these 72 students, 58 participated in the activities to measure their networks. We could obtain complete administrative information for 55 students out of the latter 58. The administrative data contains student characteristics at the moment of college entry such as age, gender, test scores from the high school exit exam students take prior to college enrollment, the household stratum that proxies the student socioeconomic status as well as the student's class schedule which we use to construct measures of exposure to other students. This student sample results in a final dataset comprising a potential 1,485 undirected relationships among the students.

Tracking social networks over time: Interactions elicited through the turnstile data. Beyond survey-derived networks, we also harness university administrative records, specifically leveraging data from student ID swipes at campus turnstiles. Adopting the methodology of Velasco (2023), we match anonymized student IDs from our sample with turnstile data to identify pairwise interactions through synchronized campus movements. Specifically, we classify a pair of students as linked if their IDs are swiped within a three-second interval at the same turnstile and in the same direction

[16]Two students did not provide complete answers to the the question on how much money they would return in the role of the receiver. As explained below, our measures of dyadic reciprocal trust can only be computed with complete answers.

(entering or exiting), provided this pattern occurs at least twice during the academic semester. This approach minimizes measurement error and closely approximates the interactions typically captured through surveys (see details of the validation process using our survey data in Appendix A). We use this approach to capture long-term networks, that is, students' interactions between 2017-2 and 2019-2. To capture interactions occurring between August and November of 2017 (short-term networks), we relax this definition and classify a pair of students as linked if their IDs are swiped within a three-second interval, at the same turnstile and going in the same direction in the given month.

We have turnstile-based interaction data for all 70 students who provided complete responses in the trust experiment. For 64 of these students, we also have information on some of their "basic controls" (socioeconomic status, hometown, and high school exit exams). We discuss the results of our baseline specification with the extended sample of 70 students (2,415 dyads) and 64 students (2,016 dyads) in Appendix C.

The interactions elicited through turnstile data serve two key purposes in augmenting our survey-based network findings. Firstly, they act as a robustness check, validating the patterns observed in the survey-elicited interactions. Secondly, they provide insights into the dynamic nature of social interactions and the evolving role of reciprocal trust in shaping social networks over time. In particular, we can use the turnstile-based networks to keep track the real-time evolution of social relationships between students, starting from one-month post-admission up to five semesters thereafter.

2 Empirical analysis

In this section, we present our empirical analysis of the link between reciprocal trust and the probability of relationship formation. We begin by stating a precise definition of reciprocal trust (Subsection 2.1). In Subsection 2.2, we describe the subjects' characteristics, behavior in the trust experiment, and networks. In Subsection 2.3, we present our baseline specification. In Subsection

2.4, we examine the effects of reciprocal trust on relationship formation over time. Overall, our results suggest that reciprocal trust does not play a relevant role in relationship formation among our subjects, while time of exposure, prior acquaintance, and homophily along some demographic traits, (such as socioeconomic status and hometown) are important determinants of social ties. In Appendix C we verify that our results are robust to changes in the way we measure reciprocal trust, to the social relationships we analyze, and to the sample of dyads that we consider.

2.1 Reciprocal trust

Throughout the analysis, we study the relation between the likelihood that a link between a pair of agents i and j forms and $RecipTrust_{ij}$—a measure of reciprocal trust between i and j. We define $RecipTrust_{ij}$ as the sum of the amounts that i and j would have received in the role of the sender when interacting with each other in the trust experiment. To be precise, suppose that i and j interact in the trust experiment with i as sender and j as receiver, and let $EfTrust_{ij}$ be the total amount that i would obtain in the experiment (i.e., the amount that he[17] would receive back from j, computed using i's sender strategy and j's receiver strategy). $EfTrust_{ij}$ is large to the extent that i sends a large amount to j and j, in response, returns a large amount to i. This is the case because the amount that j can send back to i is limited by the amount that he receives from i in the first place. $EfTrust_{ij}$ is thus a measure of i's trust in an anonymous partner that he would see effectively reciprocated if that partner happened to be j. $EfTrust_{ji}$ is defined analogously. Finally, we let,

$$RecipTrust_{ij} = EfTrust_{ij} + EfTrust_{ji}.$$

We are able to compute $RecipTrust_{ij}$ for every pair of agents because we implemented the strategic version of the trust experiment: we asked each subject to specify the amount they would send as a sender and the amount they would return as a receiver in response to each possible received amount. $RecipTrust_{ij}$ captures how "productive" the partnership between i and j in the

[17] We use the pronoun "he" only for economy of language.

trust experiment would have been for the senders, assuming both individuals played the role of sender with the other as the receiver.[18]

2.2 Students' characteristics, behavior, and networks

Students' characteristics. Table 1 presents summary statistics for the individual variables collected during both the first and second stages of the data collection process. The statistics reported refer to the sample of 55 students (1) who participated in the first stage, (2) who filled out the online survey we administered in the second stage, and (3) for whom we could obtain administrative data. Table 2 provides summary statistics for the administrative variables for the same sample of students. We proxy socioeconomic status with an administrative classification referred to as "estratificación socioeconómica" (socioeconomic stratification), which classifies residential real estates into six categories, ranging from 1 (corresponding to the poorest socioeconomic status) to 6 (corresponding to the richest one). The high school exit examination, officially referred to as the SABER 11 examination, is a standardized test administered to every graduating high school cohort in Colombia. This examination is similar to the SAT and ACT examinations in the United States, and its score ranges from 0 to 500.

In our sample, most of the students come from wealthy families in Bogotá. Only for 34.5% of the students' socioeconomic status is 4 or less, and the average socioeconomic status is 5 out of 6. The average score obtained at the high school exit examination is about 400, which usually falls in the top percentiles of the country-level score distribution.

Table 3 presents a balance test indicating that the observable characteristics of the 55 students who participated in the lab experiment, completed the online survey administered in the second stage, and for whom we obtained administrative data are, on average, similar to those of the students for whom we have administrative data but did not complete the second stage.

[18] This is formally equivalent to attaching equal probabilities to either of them being the sender.

Table 1: Summary statistics for the individual characteristics

First stage

Variable	Avg.	S.d.
Sex (1 = male)	0.7	0.46
Age	17.8	0.57
Number of siblings	1.19	0.77
Number of friends outside university	27.06	35.24
Number of people in the cohort known from before	2.38	3.31
Happiness (0-3)	2.36	0.57

Second stage

Variable	Avg.	S.d.
Number of friends at this university met before starting university	11.2	8
Number of friends at this university met after starting university	10.8	9.14
Number of friends not at this university	18.42	19.27
Weekly hours spent socializing with friends at this university	12.5	11.02
Weekly hours spent socializing with friends not at this university	9.02	14.59
Weekly hours spent doing physical activity	4.4	4.63
Optimistic/Realistic (1—5)	3.2	1.24
Extroverted/Introverted (1—5)	3	1.02
Assertive/Inhibited (1—5)	2.45	1
Outgoing/Shy (1—5)	2.35	1.09
Height (cm)	173.56	8.88
Weight (kg)	65.55	10.31
Wearing glasses (1 = yes)	0.33	0.47
Wearing tattoos (1 = yes)	0.02	0.13
Wearing piercings (1 = yes)	0.2	0.40
Smoker (1 = yes)	0.29	0.46
Attending parties (0 = never, 1 = once in a while, 2 = frequently)	1.42	0.53
Sociable (1—5)	3.65	0.84
Satisfied with own social life (1—5)	4.02	1.05
Difficulty making friends in university (1—5)	3.53	1.21
Satisfied with own number of friends	3.27	1.19

Summary statistics under the “first stage" header refer to student characteristics obtained from survey questions asked on the “university welcome day,” immediately after the trust experiment took place. Summary statistics under the “second stage” header refer to characteristics obtained through survey questions administered at the end of the academic semester, immediately after eliciting the subjects’ social networks. The first-stage summary statistics are based on a sample of 64 students who participated in the lab experiment during the first stage and for whom we have complete administrative data. The second-stage summary statistics refer to the sample of 55 students who participated in the first and second stages, and for whom we have administrative data.

Table 2: Summary statistics for the individual characteristics

Variable	Avg.	S.d.
Socioeconomic status (1—6)[a]	5	0.96
Sex (1 = male)	0.69	0.47
Exit exams	398.36	23.55
Hometown (1 = Bogota)	0.76	0.43

This table reports summary statistics for the sample of 55 students who participated in the lab experiment, who filled out the online survey we administered in the second stage, and for whom we could obtain administrative data.

[a] We proxy socioeconomic status with an administrative classification referred to as "estratificación socioeconómica" (socioeconomic stratification), which classifies residential real estates into six categories, ranging from 1 (corresponding to the poorest socioeconomic status) to 6 (corresponding to the richest one).

Table 3: Balance tests: Analysis sample vs. sample of students who did not complete the second stage

	Analysis sample			Excluded participants			
	n	Avg.	S.d.	*n*	Avg.	S.d.	Diff.
Money sent[a]	55	4.96	2.62	9	4.66	2.6	0.3
Socioeconomic status (1—6)	55	5	0.96	9	4.75	0.89	0.25
Gender	55	0.69	0.47	9	0.78	0.44	-0.09
GPA	55	4.08	0.35	9	3.72	0.89	0.36
High-school exit exam	55	398.4	23.55	9	390	22.15	8.36

This table presents balance tests comparing the average observable characteristics between two groups of students: (1) the 55 students who participated in the lab experiment, completed the online survey administered in the second stage, and for whom we obtained administrative data, and (2) the students who participated in the first stage and for whom we have administrative data but did not complete the second stage of the data collection process.

[a] Each unit represents two thousand pesos.

Students' behavior in the trust experiment. Figure 1 shows several summary statistics for the students' behavior in the trust experiment. On average, the senders sent about half of his or her endowment of $COP\$20,000$ (Std. Dev. is $COP\$5,251.58$). Overall, our subjects' behavior in the laboratory squares well with the literature.

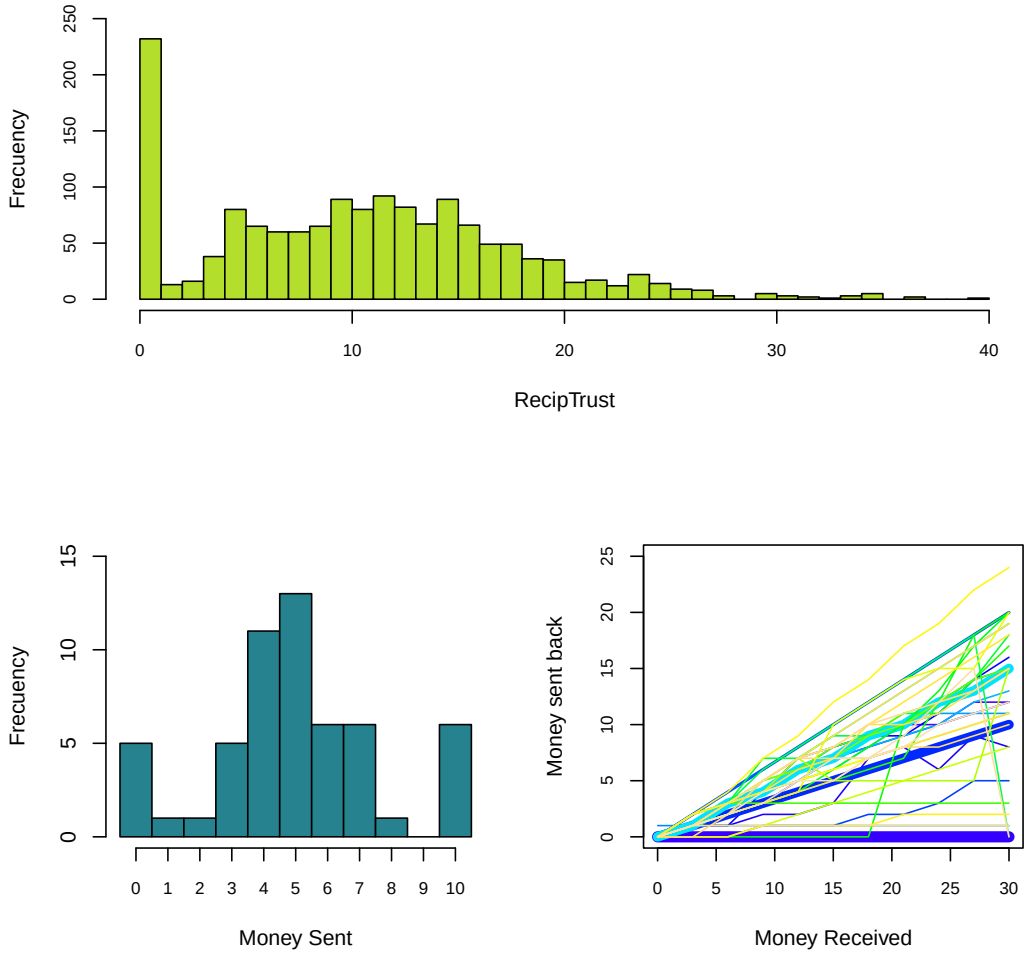


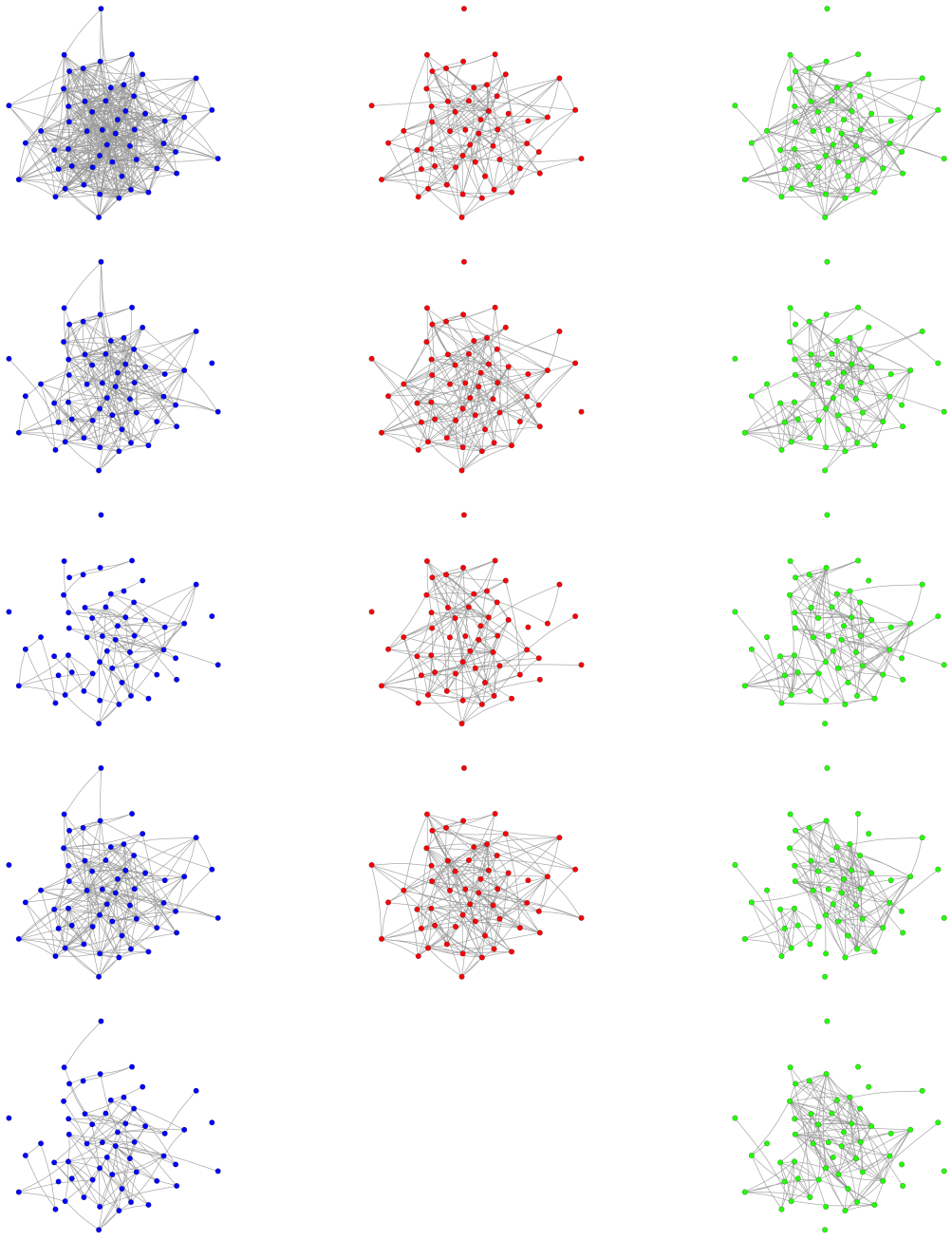
Figure 1: Top: Frequency of $RecipTrust_{ij}$ among the 1485 dyads in our sample. Bottom left: Frequencies of *money sent* (as senders) by the students in our sample. Bottom right: Profiles of *money sent back* (as receivers) as a function of *money received*. The width of the line is proportional to the number of students that responded with that strategy profile. In all graphs, money is measured in units of two thousand pesos.

During the first stage of the data collection process (right after the trust experiment took place),

we also asked survey questions aimed at measuring generalized (4.a and 4.b), and particularized trust towards friends (5.a and 5.c) and neighbors (5.e).[19] In Appendix B, we compare our lab-based trust measure with alternative measures obtained from the survey answers.

[19]We distribute the surveys only after the experiment was finished to avoid biasing subjects.

Figure 2: Survey- and turnstile-based networks.



This figure shows the networks involving the 1485 dyads used in the baseline specification (Section 2.3). The left column shows survey-elicited networks (greeting, having lunch together, studying together, confiding in, and friendship). The center column displays short-term turnstile-based networks (August, September, October, and November 2017). The right column features long-term turnstile-based networks (2017-2, 2018-1, 2018-2, 2019-1, and 2019-2, where -1 denotes the first semester and -2 denotes the second semester).

Students' networks. Figure 2 displays the undirected survey-elicited networks, the short-term turnstile-based networks, and the long-term turnstile-based networks for the sample of 55 students who participated in both stages of the data collection process and for whom we could obtain administrative data.[20] Tables 4, 5 and 6 display some summary statistics for the social networks. The average degree in the (survey-based) friendship network is 9.53, the average local clustering is 0.45, the global clustering is 0.39, and the average path length is 2.12.[21] The characteristics of the networks we retrieve square well with the literature (Jackson, 2010). In particular, they all exhibit high degrees of clustering and low average path lengths. The greeting network is denser than the other networks, the having-lunch-together and confiding-in networks are sparser, and the studying together and friendship networks sit in between the two extremes. In all networks, there is one giant component, and the greeting network is connected. The average path lengths are similar across survey- and turnstile-based networks. The mean degrees and the clustering coefficients of the long-term survey-elicited networks are all within the range of the clustering coefficients observed in the survey networks. The short-term turnstile-based networks tend to have larger mean degrees and smaller clustering coefficients than the long-term turnstile-based networks. This pattern is consistent with the idea that short-term networks tend to reflect relationships in formation, some of which do not last, and which may not yet be embedded in the kind of cohesive sets of social ties which display high clustering.

2.3 Baseline specification

We test whether pairs of individuals exhibiting higher reciprocal trust are more likely to form new connections. We use linear probability models (LPMs) to estimate the effect of $RecipTrust_{ij}$

[20] For comparability, the nodes in all networks are displayed using a Fruchterman-Reingold layout of the greeting network.

[21] The degree of an individual i is number of links of i . A person's neighbors are all the individuals linked to i. To calculate an individual's local clustering, we divide the number of edges between his or her neighbors by the number of links that could exist between them. Let a triplet be three individuals that are connected by either two or three links. Say that a triplet is closed if three edges are connecting these individuals. The global clustering of a network is the share of closed triplets among all triplets. A path from i to j is the smallest number of links that need to be crossed to go from i to j.

Table 4: Summary statistics for the survey-based networks

Network	Variable	Avg.	S.d.
Friendship	Degree	9.53	4.96
	Clustering (local)	0.45	0.19
	Clustering (global)	0.39	-
	Path length	2.12	0.73
Having lunch together	Degree	3.71	2.73
	Clustering (local)	0.42	0.30
	Clustering (global)	0.40	-
	Path length	3.99	2.01
Studying together	Degree	8.69	4.99
	Clustering (local)	0.44	0.21
	Clustering (global)	0.34	-
	Path length	2.11	0.68
Confiding in	Degree	4.25	2.74
	Clustering (local)	0.34	0.27
	Clustering (global)	0.30	-
	Path length	2.93	1.11
Greeting	Degree	18.47	7.99
	Clustering (local)	0.54	0.10
	Clustering (global)	0.50	-
	Path length	1.67	0.50

This table reports summary statistics for the survey-based networks. The number of observations is 1485 dyads for each network.

on the presence or absence of relationships in our networks. In the following, we use capital letters for random variables, small letters for possible realizations, and bold letters for vectors. We assume that for, each unordered pair of subjects $\{i, j\}$, the probability that an undirected link between i and j forms is

$$Y_{ij} = \beta^0 + \boldsymbol{\beta}^1(\boldsymbol{X}_i + \boldsymbol{X}_j) + \boldsymbol{\beta}^3 \boldsymbol{Z}_{ij} + \varepsilon_{ij}, \tag{1}$$

where $Y_{ij} = 1$ indicates that i and j have a relationship in the network in question, $\boldsymbol{X}_i$ and $\boldsymbol{X}_j$ are vectors of individual-level characteristics, and $\boldsymbol{Z}_{ij}$ is a vector of pairwise-level characteristics.[22]

[22]Note that since we deal with undirected networks, neither endpoint of any given dyad is special, and it follows that it would be meaningless to allow for a different relationship between $\Pr(Y_{ij})$, and $\boldsymbol{X}_i$ and $\boldsymbol{X}_j$.

Table 5: Summary statistics for the short-term turnstile-based networks

Network	Variable	Avg.	S.d.
Aug. 2017	Degree	6.15	3.15
	Clustering (local)	0.33	0.22
	Clustering (global)	0.27	-
	Path length	2.41	0.82
Sept. 2017	Degree	6.80	3.37
	Clustering (local)	0.28	0.18
	Clustering (global)	0.24	-
	Path length	2.22	0.72
Oct. 2017	Degree	5.27	2.99
	Clustering (local)	0.22	0.14
	Clustering (global)	0.22	-
	Path length	2.51	0.86
Nov. 2017	Degree	7.42	3.26
	Clustering (local)	0.30	0.13
	Clustering (global)	0.26	-
	Path length	2.19	0.70

This table reports summary statistics for each of the short-term turnstile-based networks (August, September, October, and November 2017). The number of observations is 1485 dyads for each network.

If we were to assume that ε_{ij} is independent of $\varepsilon_{k\ell}$, for each $ij \neq k\ell$, then we could estimate Equation (1) with a standard OLS regression. However, when observations $\left(Y_{ij}\right)_{i,j=1,\ldots,N,i\neq j}$ correspond to the presence of links between N individuals, it is generally unsafe to assume that unobservables are independent across pairs of individuals. Specifically, the unobservables of pairs that share a common individual are likely to be correlated. As a result, standard OLS regressions produce consistent point estimates but underestimate p-values.[23] Acknowledging the possibility of autocorrelation in the networks' adjacency matrices, even after adjusting for observed traits, we lean toward a conservative approach that uses the dyadic-robust variance estimator (as detailed in

[23]Hence, p-values would too easily lead to rejecting the null hypothesis that an explanatory variable does not predict the probability that a link forms.

Table 6: Summary statistics for the long-term turnstile-based networks

Network	Variable	Avg.	S.d.
2017-2	Degree	7.05	3.49
	Clustering (local)	0.36	0.19
	Clustering (global)	0.31	-
	Path length	2.28	0.77
2018-1	Degree	5.60	3.27
	Clustering (local)	0.45	0.30
	Clustering (global)	0.44	-
	Path length	2.75	1.07
2018-2	Degree	5.13	3.10
	Clustering (local)	0.54	0.28
	Clustering (global)	0.50	-
	Path length	3.37	1.56
2019-1	Degree	5.45	3.87
	Clustering (local)	0.56	0.23
	Clustering (global)	0.51	-
	Path length	3.24	1.50
2019-2	Degree	5.49	4.17
	Clustering (local)	0.53	0.23
	Clustering (global)	0.54	-
	Path length	2.90	1.33

This table reports summary statistics for the long-term turnstile-based networks (2017-2, 2018-1, 2018-2, 2019-1, and 2019-2, where suffix -1 denotes the first semester and -2 denotes the second semester).

Fafchamps and Gubert (2007) and Tabord-Meehan (2019)) to refine standard errors.

We also add a battery of controls for several individual and pairwise characteristics. As for individual characteristics, we use information on sex, hometown, age, eye color, hair color, height, weight, whether wearing glasses, whether wearing tattoos, whether wearing piercings, number of siblings, score obtained at the high school exit examination.[24] As for pairwise characteristics, we

[24]While we collect information about few other characteristics, e.g., students' habits, like attending parties. However, we do not include them in our main specification as they may be "bad controls." For example, a more trusting student may be more willing to attend parties, which in turn affects the value of of $RecipTrust_{ij}$ with any other student j, and his or her chances of forming relationships. We run the analysis disregarding the "bad controls" problem and

have information on whether the students reported knowing each other from before our intended socialization period, and the amount of time they spent together in the same classrooms during the first semester, as measured by the number of university credits that the students share.[25] Moreover, for each individual characteristic X and unordered pair of individuals $\{i, j\}$, we can also control for the presence of homophily in that characteristic, as defined by $\Delta X_{ij} := |X_i - X_j|$.

To ease the comparison of the effect of different covariates, we standardize each non-binary variable by subtracting its average from the variable and dividing the result by the standard deviation of the variable.[26] Thus, we can interpret the marginal effects in the regressions below as resulting from one standard deviation increases in the original variables.

2.3.1 Results

We estimate several linear probability models using both the (survey-based) friendship network and the first-semester turnstile-based network. Specifically, for each of these two networks, we run three different models. First, we use only $\mathit{RecipTrust}_{ij}$ as an explanatory variable. Next, we introduce pairwise-level characteristics and homophily along individual characteristics as controls. Finally, we include both pairwise and individual-level characteristics as controls. We restrict our sample to the 1485 dyads involving the subjects for whom we have complete first stage, second stage, and administrative data. We report the results in Table 7.

The first two columns of Table 7 present the results of two linear probability models that regress the (survey-based) friendship network and the first-semester turnstile-based network on the reciprocal trust between individuals i and j. In the second two columns, we introduce pairwise-level controls in the regression. First, we include a dummy variable indicating whether the students

the results remain largely unaffected.

[25] All of our subjects take the same six courses in the first semester but are allocated to different classrooms, corresponding to different time schedules. Selection is likely not an issue in this context. While students can express preferences for different classrooms, they can only do so before the first semester starts, and these preferences are not necessarily reflected in the final allocation of students to classrooms. The majority of students do not know each other from before; hence, it is unlikely that the variation in time of exposure arises from students purposefully choosing to attend the same classrooms. Even if that were the case, in our regressions we control for whether students knew each other from before whenever we introduce time of exposure as an explanatory variable.

[26] Notice that when $X\Delta$ is a non-binary variable, we standardize it after having computed the difference.

knew each other beforehand and their time of exposure in class. To account for the possibility that reciprocal trust may affect only pairs with significant exposure to each other, we include the interaction between individuals' time of exposure and reciprocal trust. Additionally, we incorporate multiple variables reflecting differences in individual characteristics, which may be important due to homophilic motives. For readability, we omit the coefficients associated with some controls.[27] Finally, the last two columns of the table show the results of the specification that includes all (pairwise- and individual-level) controls, including the individual controls X_i and X_j that we used to construct the dyadic difference ΔX_{ij}, which we included in the two middle columns to control for homophily in variable X (for each X). Following the typical approach used in network regressions with undirected connections, we represent each individual's level control as the sum of the values of the variable for both endpoints within the dyad under examination.

[27] Section 1 of the Online Appendix reports the table including the coefficients associated with all the controls in the regression.

Table 7: Baseline regressions: Survey-elicited friendship network and turnstile-inferred first semester network on reciprocal trust and various controls

	(i)		(ii)		(iii)	
	Friends	Turnst.	Friends	Turnst.	Friends	Turnst.
Constant	0.176^{***}	0.131^{***}	0.155^{***}	0.112^{***}	0.433^{***}	0.368^{**}
	(0.022)	(0.015)	(0.029)	(0.033)	(0.124)	(0.146)
Reciprocal trust	0.008	0.006	-0.006	0.002	-0.002	0.004
	(0.013)	(0.011)	(0.013)	(0.012)	(0.015)	(0.013)
Knew e.o. from before			0.516^{***}	0.334^{***}	0.539^{***}	0.361^{***}
			(0.069)	(0.066)	(0.077)	(0.066)
Exposure			0.072^{***}	0.094^{***}	0.067^{***}	0.096^{***}
			(0.016)	(0.017)	(0.016)	(0.018)
Reciprocal trust $\times$			-0.002	0.001	-0.001	-0.003
Exposure			(0.013)	(0.015)	(0.015)	(0.013)
Socioeconomic					0.009	0.013
status					(0.022)	(0.019)
Socioeconomic			-0.029^{**}	-0.022	-0.035^{**}	-0.021
status Δ			(0.014)	(0.016)	(0.016)	(0.014)
Sex					-0.049	-0.063^{**}
					(0.045)	(0.032)
Sex Δ			0.005	-0.006	-0.015	-0.024
			(0.034)	(0.02)	(0.026)	(0.024)
Exit exams					-0.027	0.002
					(0.019)	(0.018)
Exit exams Δ			0	0.02^{**}	-0.009	0.007
			(0.013)	(0.008)	(0.016)	(0.013)
Hometown					-0.078	-0.056
					(0.046)	(0.046)
Hometown Δ			-0.034	-0.012	-0.093^{**}	-0.055
			(0.028)	(0.026)	(0.045)	(0.044)

This table reports three specifications of the linear probability model for the (survey-based) friendship network and the first-semester turnstile-based network. The first two columns display the results of a linear probability model using only $RecipTrust_{ij}$ as an explanatory variable. The next two columns introduce pairwise-level controls in the regression. The last two columns show the results of the specification that includes all individual-level controls, including the individual-level controls X_i and X_j used to construct the dyadic differences ΔX_{ij} included in the two middle columns to control for homophily in variable X. For readability, we omit the coefficients associated with some controls (Section 1 of the Online Appendix reports the table including the coefficients associated with all the controls in the regression). Dyadic-robust standard errors are shown in parentheses.

The first two columns show that the estimated coefficients of $RecipTrust_{ij}$ are nearly zero and statistically insignificant. In terms of magnitudes, a one standard deviation increase in reciprocal trust between i and j is associated with a 0.008 increase in the probability of a link between them in the survey-elicited network and a 0.006 increase in the turnstile-inferred network. Reciprocal trust remains insignificant even after introducing both pairwise-level controls (in the second two columns) and combined pairwise-level and individual-level controls (in the third two columns). The change in the sign of the point estimate of $RecipTrust_{ij}$ across specifications is unsurprising, given its proximity to zero. Regarding the controls, we find that prior acquaintance is a significant and substantial predictor of relationship formation. Specifically, if either subject indicates knowing the other from before, the likelihood of a link at the end of the first academic semester increases by between 0.33 and 0.36 in the survey-elicited network and between 0.52 and 0.54 in the turnstile-based network. This result demonstrates the persistence of friendships and the ease of befriending an already acquainted person. Additionally, when individuals spend more time together due to being assigned to the same classrooms, they are significantly more likely to form a link. A one standard deviation increase in time spent together due to being in the same class-sections increases the probability of a link by 0.07 in the survey-elicited network and by 0.09 in the turnstile-inferred network. Given the average densities of these networks are 0.18 and 0.13, respectively, these effects are substantial, resulting in a 40% and 70% increase in the probability of a link. These results align well with the evidence presented in Marmaros and Sacerdote (2006), which finds that first-year students tend to interact and form long-term friendships with peers who are easily accessible. Additionally, the significant positive effect of exposure time on link formation probability supports Girard, Hett and Schunk (2015)'s finding that students in the same study groups tend to form friendships among themselves.

Homophily in socioeconomic status is also significant. In the survey-elicited network, a one standard deviation increase in the difference between i and j's socioeconomic statuses decreases the probability of a link by 0.030 to 0.035. Given the average network density, this translates to a 20% decrease in the probability of a link. Although the estimates are noisier for the turnstile-

inferred networks, resulting in only marginal significance, the magnitudes are similar (-0.022 and -0.021). Similarly, homophily in hometown is a significant determinant of relationship formation. On average, if i and j both come from Bogotá or both come from outside Bogotá, they have a 0.03 to 0.09 higher chance of forming a link in the survey-elicited network. These findings are consistent with a large body of empirical evidence (McPherson, Smith-Lovin and Cook (2001)). Finally, we find that students i and j are less likely to be linked if they come from Bogotá. This is intuitive, as students from Bogotá likely already have an established network of friends in town, reducing their need to form new friendships.

Does the insignificance of reciprocal trust arise from a weak association with link formation or from high standard errors? To address this question, we analyze the power of our t-test. Suppose the true partial correlation between i and j's reciprocal trust and the presence of a link between them, $\beta^1_{\text{Recip Trust}}$, is 0.08, a magnitude comparable to that of the correlation between time of exposure and link formation. Given our sample size and the dyadic-robust standard errors of the estimated $\beta^1_{\text{Recip Trust}}$, the probability of failing to reject the null hypothesis that $\beta^1_{\text{Trust}} = 0$ is less than 0.001. More generally, given our sample size and standard errors, the minimum detectable partial correlation of reciprocal trust with the existence of a link, with 80% power, is approximately 0.016. This is slightly less than the effect of socioeconomic dissimilarity. Therefore, we can confidently assert that if the true effect of reciprocal trust on link formation probability is positive, it is very likely to be smaller than the impact of variables such as prior acquaintance, time of exposure, socioeconomic dissimilarity, and hometown dissimilarity.

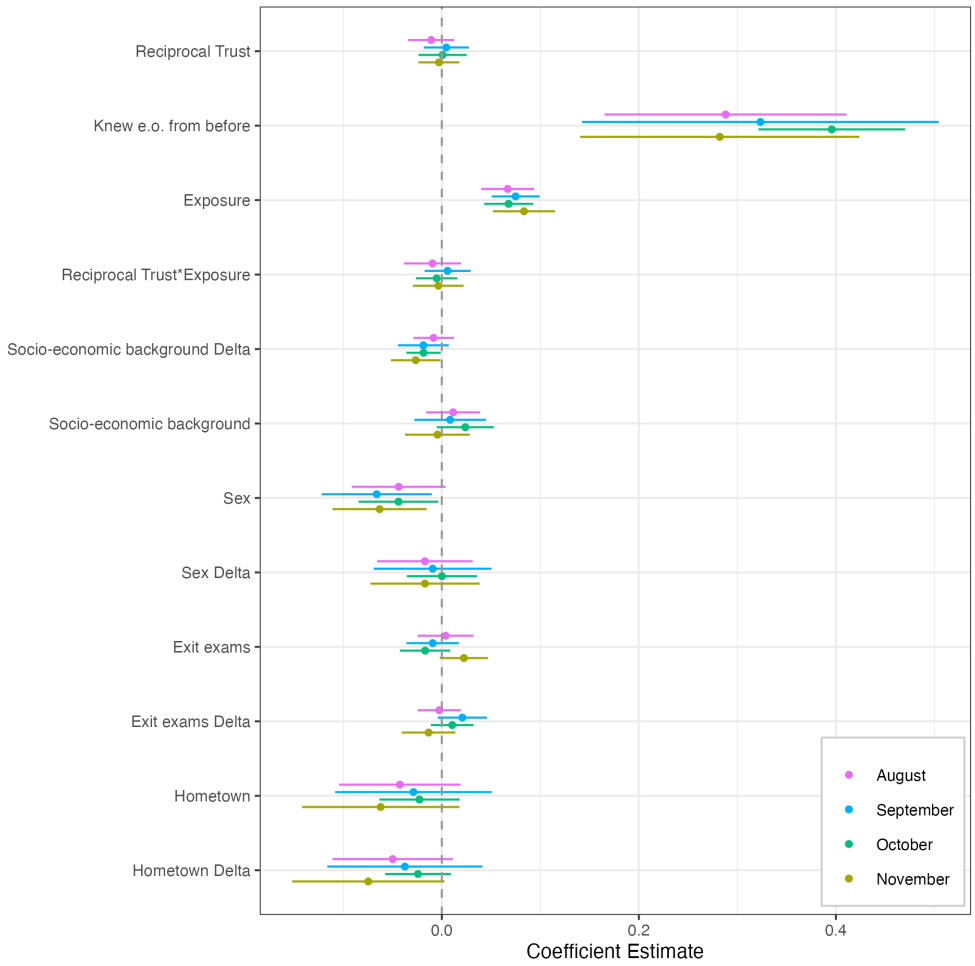
2.4 Predicting relationship formation in the short term and the long term

Our analysis uses a one-semester period to define social ties, which aligns with Christakis (2015)'s findings for U.S. colleges, which highlight a critical initial window of less than a month for forming acquaintances before relationships solidify. However, the dynamics at non-residential institutions, such as the one we study, may differ significantly.

Given the potentially less intense socialization experiences at these universities, our ties may require prolonged acquaintance periods. As a result, reciprocal trust could become a relevant predictor of relationships later in their university journey, not necessarily by the end of the first semester. Alternatively, persistent classroom interactions and shared friendships could lead to gradual camaraderie among students. In this scenario, reciprocal trust would primarily influence relationships in the very early stages, potentially even shorter than our one-semester analysis period.

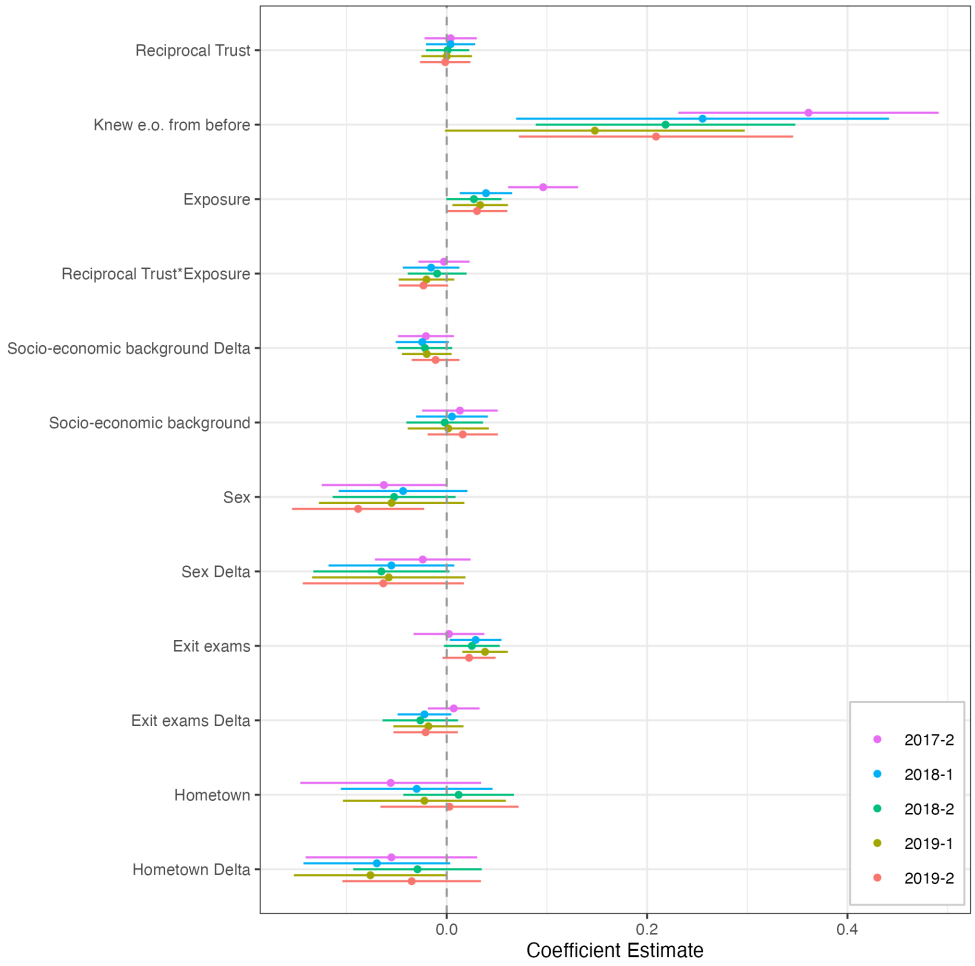
For these reasons, we leverage the turnstile data to capture both short-term and long-term relationship formation: specifically, at one, two, and three months into the first semester, and then two through five semesters post-admission. Figure 3 shows the point estimates and 80% confidence intervals from the baseline regressions of the short-term turnstile data on reciprocal trust and all the pairwise- and individual-level controls. Figure 4 shows the point estimates and 80% confidence intervals from the baseline regressions of the long-term turnstile data on reciprocal trust and all the pairwise- and individual-level controls. For readability and consistency, we report only the coefficient estimates and confidence intervals for the same controls shown in Table 7. The coefficients are very similar to those presented in the third column of Table 7. In the short run, our evidence indicates a negligible influence of reciprocal trust, with coefficients ranging narrowly between -0.003 and 0.005 during the September to November span. The August data reveals a slightly more pronounced coefficient at -0.011, yet it pales in comparison to the coefficients linked with robust predictors of link formation—factors like knowing each other from before, time of exposure duration, gender, hometown, and differences in hometown.

Figure 3: Coefficient estimates from the baseline regressions of the short-term turnstile-based networks on reciprocal trust and various controls



This figure shows the estimated coefficients and 80% confidence intervals from the baseline regressions of the short-run turnstile networks on reciprocal trust and all the pairwise-level and individual-level controls. For readability and consistency, we report only the coefficient estimates and confidence intervals for the same controls shown in Table 7 (see Section 2 for further details).

Figure 4: Coefficient estimates from the baseline regressions of the long-term turnstile-based networks on reciprocal trust and various controls



This figure shows the estimated coefficients and 80% confidence intervals from the baseline regressions of the long-run turnstile networks on reciprocal trust and various controls. For readability and consistency, we report only the coefficient estimates and confidence intervals for the same controls shown in Table 7 (see Section 2 for further details).

2.5 Robustness checks

In Appendix C, we discuss several robustness checks and demonstrate that our results remain consistent when using alternative measures of reciprocal trust, different survey-based networks as the dependent variable, and various samples of dyads based on the turnstile-based networks.

3 Conclusions

We collected data on trust and trustworthiness toward strangers from a cohort of incoming freshman economics students at a university in Bogotá, Colombia. This data was gathered on the students' first formal day on campus, before they had substantial opportunities to socialize. At the end of their first academic semester, we collected survey data on five social networks among them. We used administrative data on students' co-movements across turnstiles at campus entry and exit points to track their networks both in different points in time. For each pair of students in our sample, we computed a measure of reciprocal trust as the sum of the amounts that the individuals in the pair would have received in the role of sender when interacting with each other in a trust experiment. For each of the networks we elicited or inferred from the turnstile data, we estimated the effect of reciprocal trust on the probability of relationship formation.

We find strong evidence against the hypothesis that reciprocal trust predicts link formation. Our results suggest that if reciprocal trust is indeed a determinant of relationship formation, its influence is likely much weaker compared to several other characteristics that are well understood to play a significant role. In particular, factors such as time of exposure, prior acquaintance, and similar socioeconomic status are far more important for relationship formation than the students' ability to cooperate in social dilemmas, as measured by total potential earnings in a trust experiment. This holds true regardless of whether we measure relationships in the short term or long term, and regardless of the type of relationship we examine.

Overall, our results cautiously suggest that the emphasis on trust and trustworthiness as key facilitators of social relationships in the social capital literature may be overstated. Relationships seem to result more from people's tendency to associate with those similar to them or from random chance than from their inclination to trust and reciprocate the trust of strangers.

References

Bailey, Michael, Rachel Cao, Theresa Kuchler, Johannes Stroebel and Arlene Wong. 2018. "Social Connectedness: Measurement, Determinants, and Effects." *Journal of Economic Perspectives* 32(3):259–80.

Baker, Sara, Adalbert Mayer and Steven L Puller. 2011. "Do more diverse environments increase the diversity of subsequent interaction? Evidence from random dorm assignment." *Economics Letters* 110(2):110–112.

Berg, Joyce, John Dickhaut and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10(1):122–142.

Bramoullé, Yann, Andrea Galeotti and Brian W Rogers. 2016. *The Oxford handbook of the economics of networks*. Oxford University Press.

Buskens, Vincent. 1998. "The Social Structure of Trust." *Social Networks* 20(3):265–289.

Christakis, Nicholas A. 2015. "Making friends in new places." *New York Times* .
URL: *https://www.nytimes.com/2015/08/02/education/edlife/making-friends-in-new-places.html*

Currarini, Sergio, Matthew O. Jackson and Paolo Pin. 2009. "An Economic Model of Friendship: Homophily, Minorities, and Segregation." *Econometrica* 77(4):1003–1045.

Di Cagno, Daniela and Emanuela Sciubba. 2010. "Trust, Trustworthiness and Social Networks: Playing a Trust Game when Networks are Formed in the Lab." *Journal of Economic Behavior & Organization* 75(2):156–167.

Fafchamps, Marcel and Flore Gubert. 2007. "The Formation of Risk Sharing Networks." *Journal of Development Economics* 83(2):326–350.

Giné, Xavier and Ghazala Mansuri. 2018. "Together We Will: Experimental Evidence on Female Voting Behavior in Pakistan." *American Economic Journal: Applied Economics* 10(1):207–35.

Giné, Xavier, Pamela Jakiela, Dean Karlan and Jonathan Morduch. 2010. "Microfinance Games." *American Economic Journal: Applied Economics* 2(3):60–95.

Girard, Yann, Florian Hett and Daniel Schunk. 2015. "How Individual Characteristics Shape the Structure of Social Networks." *Journal of Economic Behavior & Organization* 115:197–216.

Glaeser, Edward L., David I. Laibson, Jose A. Scheinkman and Christine L. Soutter. 2000. "Measuring Trust." *The Quarterly Journal of Economics* 115(3):811–846.

Inglehart, Ronald. 1997. *Modernization and Post-Modernization: Cultural, Economic and Political Change in 43 Societies*. Princeton University Press.

Jackson, Matthew O. 2010. *Social and Economic Networks*. Princeton University Press.

Jackson, Matthew O., Brian W. Rogers and Yves Zenou. 2017. "The Economic Consequences of Social-Network Structure." *Journal of Economic Literature* 55(1):49–95.

Jackson, Matthew O, Stephen M Nei, Erik Snowberg and Leeat Yariv. 2023. The dynamics of networks and homophily. Technical report National Bureau of Economic Research.

Jackson, Matthew O, Tomas Rodriguez-Barraquer and Xu Tan. 2012. "Social Capital and Social Quilts: Network Patterns of Favor Exchange." *American Economic Review* 102(5):1857–97.

Kosse, Fabian, Thomas Deckers, Pia Pinger, Hannah Schildberg-Hörisch and Armin Falk. 2020. "The Formation of Prosociality: Causal Evidence on the Role of Social Environment." *Journal of Political Economy* 128(2):434–467.

Marmaros, David and Bruce Sacerdote. 2006. "How Do Friendships Form?" *The Quarterly Journal of Economics* 121(1):79–119.

Mayer, Adalbert and Steven L Puller. 2008. "The Old Boy (and Girl) Network: Social Network Formation on University Campuses." *Journal of Public Economics* 92(1-2):329–347.

McPherson, Miller, Lynn Smith-Lovin and James M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27(1):415–444.

Michelman, Valerie, Joseph Price and Seth D Zimmerman. 2022. "Old boysâĂŹ clubs and upward mobility among the educational elite." *The Quarterly Journal of Economics* 137(2):845–909.

Putnam, Robert D. 1995. "Bowling Alone: America's Declining Social Capital." *Journal of Democracy* 6(1):65–78.

Santos, Paulo and Christopher B. Barrett. 2011. "Persistent Poverty and Informal Credit." *Journal of Development Economics* 96(2):337–347.

Tabord-Meehan, Max. 2019. "Inference with Dyadic Data: Asymptotic Behavior of the Dyadic-Robust t-Statistic." *Journal of Business & Economic Statistics* 37(4):671–680.

Velasco, Tatiana. 2023. "The Effects of College Desegregation on Academic Achievement and StudentsâĂŹ Social Interactions: Evidence from Turnstile Data." *Job Market Paper* .
URL: *https://tativelasco.com/static/Velasco*$_t$*urnstiles*$_d$*iversity*2023.*pdf*

Woolcock, Michael. 1998. "Social Capital and Economic Development: Toward a Theoretical Synthesis and Policy Framework." *Theory and Society* 27(2):151–208.

A Validating social networks based on the turnstile data

Table 8 compares the survey-based and turnstile-based networks during the fall semester of 2017. It presents summary statistics for several turnstile-based interactions, defined by one to five entry or exit movements of students on or off campus within a three-second window among the students in our sample. The "twice" column refers to the definition of comovement used to construct the turnstile-based networks in Section 2.3. We compare the turnstile-based networks with several survey-based networks: greeting, studying together, having lunch together, and friendship. The table provides two critical statistics for each pair of turnstile- and survey-based networks: the proportion of survey-based links not corroborated by the turnstile-based networks (*survey pairs Linked unmatched with turnstile links*), and the proportion of survey-based pairs not linked in the survey-based network and yet identified as linked in the turnstile-based network (*survey pairs Not-linked flagged as turnstile links*).

The data presented in Table 8 reveal two important aspects of the turnstile-based links. Firstly, these links rarely misidentify pairs of students as connected when they are not considered so in the survey. Notably, in the "twice" column, for all types of survey-elicited pairs, the proportion of unlinked survey pairs flagged as linked by the turnstiles never exceeds 10 percent. However, a significant number of survey-identified links are not detected by the turnstiles. For example, while only 16 percent of the "having lunch together" connections are missed by the turnstiles, this figure rises to 45 percent for "friends" connections and 64 percent for "greeting" interactions. Secondly, this pattern suggests that turnstile-based links are more adept at capturing closer interactions, such as those who have lunch together, rather than more casual interactions like greetings or mere acquaintances. These findings imply that turnstile-based interactions provide a unique perspective on student relationships, capturing a dimension of interaction not fully represented in survey-based

measures, yet still offering significant insight into the nature of student connections.

Table 8: Comparison between survey- and turnstile-based interactions during the fall semester of 2017

Time window	Three seconds between ID swipes				
Comovement frequency	Once	Twice	Three	Four	Five
1. Turnstiles links					
No. of links	454	217	144	120	96
No. of students in the sample	55	55	55	55	55
2. Friends					
Total survey links			262		
Survey- and turnstile-based links matched	198	**144**	123	107	91
Share of survey pairs:					
Linked unmatched with turnstile links	0.24	**0.45**	0.53	0.59	0.65
Not-linked flagged as turnstile links	0.21	**0.06**	0.02	0.01	0.00
3. Greeting					
Pairs in survey			509		
Survey- and turnstile-based links matched	302	**182**	135	115	96
Share of survey pairs:					
Linked unmatched with turnstile links	0.41	**0.64**	0.73	0.77	0.81
Not-linked flagged as turnstile links	0.16	**0.04**	0.01	0.01	0.00
4. Studying together					
Number of turnstile pairs:			239		
Survey- and turnstile-based links matched	177	**134**	110	96	86
Share of survey pairs:					
Linked unmatched with turnstile links	0.26	**0.44**	0.54	0.60	0.64
Not-linked flagged as turnstile links	0.22	**0.07**	0.03	0.02	0.01
5. Having lunch together					
Number of turnstile pairs:			102		
Survey- and turnstile-based links matched	95	**86**	78	71	62
Share of survey pairs:					
Linked unmatched with turnstile links	0.07	**0.16**	0.24	0.30	0.39
Not-linked flagged as turnstile links	0.26	**0.09**	0.05	0.04	0.02

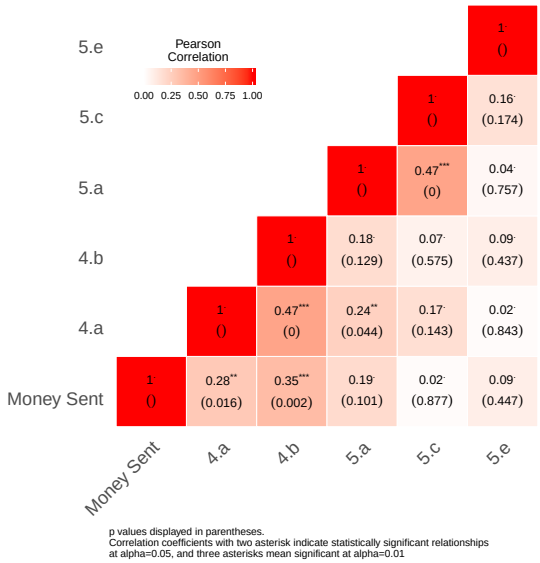
This table compares the survey-based and turnstile-based networks during the fall semester of 2017. It presents summary statistics for several turnstile-based interactions, defined by one to five entry or exit movements of students on or off campus within a three-second window among the students in our sample. The "twice" column refers to the definition of comovement used to construct the turnstile-based networks in our analysis. For each pair of turnstile- and survey-based networks, we show the proportion of survey-based links not corroborated by the turnstile-based networks (*survey pairs Linked unmatched with turnstile links*), and the proportion of survey-based pairs not linked in the survey-based network and yet identified as linked in the turnstile-based network (*survey pairs Not-linked flagged as turnstile links*).

B Trust in the lab and surveys

Figure 5 shows the correlations between the amounts of money sent by senders in the lab and the answers to survey questions 4.a, 4.b, 5.a, 5.c, and 5.e, as well as the correlations between the answers to the questions. Trust in the lab significantly correlates with generalized trust but not with particularized trust. Glaeser et al. (2000) finds that the answers to two questions on generalized trust (similar to questions 4.a and 4.b), and the answers to the questions on particularized trust (what they refer to as "past trusting behavior") are all correlated with the amount of money sent in a trust experiment. Our results confirm Glaeser et al. (2000)'s finding that generalized trust significantly predicts trusting behavior in the laboratory.[28] However, contrary to Glaeser et al. (2000), we find that particularized trust does not significantly correlate with the amount of money sent in the experiment. We believe that this discrepancy stems from a difference in Glaeser et al. (2000)'s experimental design, as their subjects knew each other's identities while playing the trust game. Moreover, in their study, individuals who arrived together at the experiment were allowed to play with each other. As a result, subjects who are friends are more likely to play together, and so particularized trust towards friends may play a crucial role in their behavior. The fact that our experiment is anonymized likely explains why past trusting behavior towards particular groups of people (i.e., friends and neighbors) does not play a significant role in predicting our subjects' behavior in the experiment.

[28]This evidence supports Glaeser et al. (2000)'s conclusion that questions on generalized trust are "more precise and meaningful than completely general, nonspecific questions regarding trust," such as the one asked in the General Social Survey.

Figure 5: Correlations between the amounts of money sent by senders in the experiment and the answers to survey questions 4.a, 4.b, 5.a, 5.c, and 5.e



This figure shows the correlations between the amounts of money sent by senders in the lab and the answers to survey questions 4.a, 4.b, 5.a, 5.c, and 5.e, as well as the correlations between the answers to the questions.

C Robustness checks

In this appendix, we discuss several robustness checks and demonstrate that our results remain consistent when using alternative measures of reciprocal trust, different survey-based networks as the dependent variable, and various samples of dyads based on the turnstile-based networks.

C.1 Other measures of reciprocal trust

The measure of reciprocal trust within dyads that we use in the baseline specification, $RecipTrust_{ij}$, may be subject to measurement error or confounded by other factors. It might also be the case that the type of trust measured by the trust experiment (how much money a subject would endow another with, in the absence of commitment or punishment technologies) is not the type of trust that matters for approaching and interacting with strangers. We address this issue by constructing alternative measures of reciprocal trust within dyads.

While trust in the trust experiment can only be measured by the amount of money sent, trustworthiness can be assessed in several ways. In Section 2, we measure the trustworthiness of an individual i as the amount that he would send back to j in response to the amount which they would receive from individual j if i was a receiver and j was a sender. Our first alternative measure of reciprical trust, which we refer to as $Exp1$, is calculated as follows:

$$Exp1 = Exp1_{ij} + Exp1_{ji},$$

where $Exp1_{ij}$ is the amount of money sent by individual i in the pair as a sender during the trust experiment plus the *average* amount of money sent back by individual j as a receiver. The second alternative measure of reciprical trust, which we refer to as $Exp2$, is calculated as follows:

$$Exp2 = Exp2_{ij} + Exp2_{ji},$$

is calculated by summing the amount of money sent by individual i in the pair as a sender during the trust experiment plus the *average* amount of money sent back by individual j as a receiver conditional on receiving at least COP $10,000. This measure considers only the average amounts sent back by each individual in the pair as a receiver in the trust experiment, but only when they received more than half of the sender's endowment. Our final measure of reciprocal trust, which we refer to as $Exp3$, is calculated as follows:

$$Exp3 = Exp3_{ij} + Exp3_{ji},$$

where $Exp3_{ij}$ represents the amount of money sent by individual i to individual j as a sender in the trust experiment. This measure disregards trustworthiness entirely and considers only the amounts sent by each individual in the pair as senders in the trust experiment.

We also construct alternative measures of reciprocal trust based on subjects' responses to survey questions about generalized trust, instead of their behavior in the trust experiment. The survey

questions on generalized trust ask, on a 1—5 scale, to what extent the subjects agree with the following statements: (4.a.) *One cannot trust strangers* and (4.b.) *When dealing with strangers, one should be careful and not readily trust them.* These measures are less context-specific than the measure we build from the trust experiment, and could thus be argued to capture a more comprehensive dimension of trust. We let

$$Surv1 = Surv1_i + Surv1_j,$$

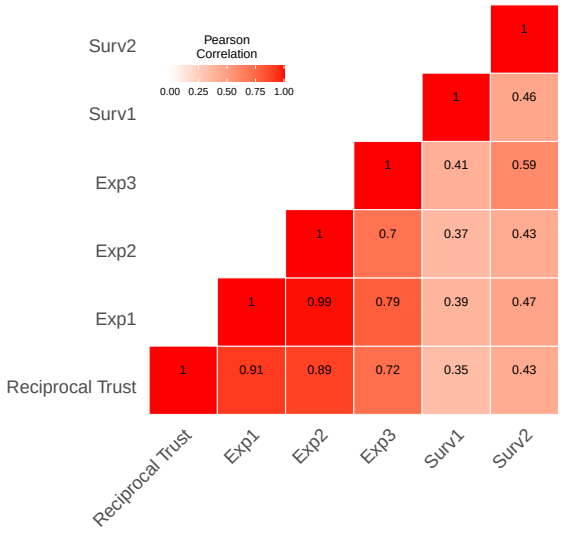
where $Surv1_i$ is the negative value of student i's answers to question 4.a. Finally, we let

$$Surv2 = Surv2_{ij} + Surv2_{ji},$$

where $Surv2_i$ is the negative value of student i's answer to question 4.b.

Figure 6 shows the correlations between our baseline measure of reciprocal trust based on the students' behavior in the trust experiment (defined in Subsection 2.1) and the five alternative measures of reciprocal trust defined above. Table 9 shows the results of running our baseline specification using the survey-based friendship network on each one of the five alternative measures of reciprocal trust, where we control for all pairwise and individual-level characteristics (analogous to Column (iii) in Table 7). Table 10 shows the results of running analogous regressions but using the first-semester turnstile-based networks.

Figure 6: Correlations among our measure of reciprocal trust, the three alternative measures based on the lab data and the two alternative measures based on survey data.



Note: This figure shows the correlations between our baseline measure of reciprocal trust (defined in Subsection 2.1) in dyads and the five alternative measures of reciprocal trust defined in Section C.1.

Table 9: Robustness regressions: Survey-elicited friendship network on alternative measures of reciprocal trust and various controls

	Recip. trust	**Exp1**	**Exp2**	**Exp3**	**Surv1**	**Surv2**
Constant	0.433^{***}	0.434^{***}	0.433^{***}	0.447^{***}	0.443^{***}	0.442^{***}
	(0.124)	(0.125)	(0.125)	(0.123)	(0.122)	(0.016)
Trust measure	-0.002	-0.003	-0.003	-0.014	-0.017	-0.037^{*}
	(0.015)	(0.021)	(0.019)	(0.022)	(0.021)	(0.017)
Knew e.o. from before	0.539^{***}	0.54^{***}	0.539^{***}	0.544^{***}	0.54^{***}	0.542^{***}
	(0.077)	(0.077)	(0.077)	(0.076)	(0.077)	(0.078)
Exposure	0.067^{***}	0.067^{***}	0.067^{***}	0.067^{***}	0.067^{***}	0.068^{***}
	(0.016)	(0.016)	(0.016)	(0.015)	(0.016)	(0.016)
Reciprocal trust $\times$	-0.001	-0.001	-0.001	-0.001	0	-0.001
Exposure	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)
Socioeconomic	0.009	0.009	0.009	0.007	0.005	0.001
status	(0.022)	(0.021)	(0.021)	(0.023)	(0.024)	(0.023)
Socioeconomic	-0.035^{**}	-0.035^{**}	-0.035^{**}	-0.035^{**}	-0.036^{**}	-0.033^{**}
status Δ	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Sex	-0.049	-0.049	-0.049	-0.048	-0.055	-0.055
	(0.045)	(0.045)	(0.044)	(0.045)	(0.048)	(0.046)
Sex Δ	-0.015	-0.015	-0.015	-0.015	-0.015	-0.012
	(0.026)	(0.026)	(0.026)	(0.025)	(0.025)	(0.025)
Exit exams	-0.027	-0.027	-0.027	-0.029	-0.029	-0.027
	(0.019)	(0.018)	(0.019)	(0.019)	(0.018)	(0.018)
Exit exams Δ	-0.009	-0.01	-0.01	-0.012	-0.009	-0.012
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Hometown	-0.078	-0.077	-0.078	-0.077^{*}	-0.077	-0.075
	(0.048)	(0.048)	(0.048)	(0.046)	(0.047)	(0.048)
Hometown Δ	-0.093^{**}	-0.093^{**}	-0.093^{**}	-0.093^{**}	-0.093^{**}	-0.093^{**}
	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)
Observations: 1485						

This table presents the results of linear probability models regressing the (survey-based) friendship network on the reciprocal trust measures constructed using three alternative measures based on the lab data and two measures based on survey data, as described in Section C. All specifications include all pairwise- and individual-level controls. For readability, the coefficients associated with some controls are omitted. Dyadic-robust standard errors are shown in parentheses.

Table 10: Robustness regressions: First-semester turnstile-based network on alternative measures of trust and various controls

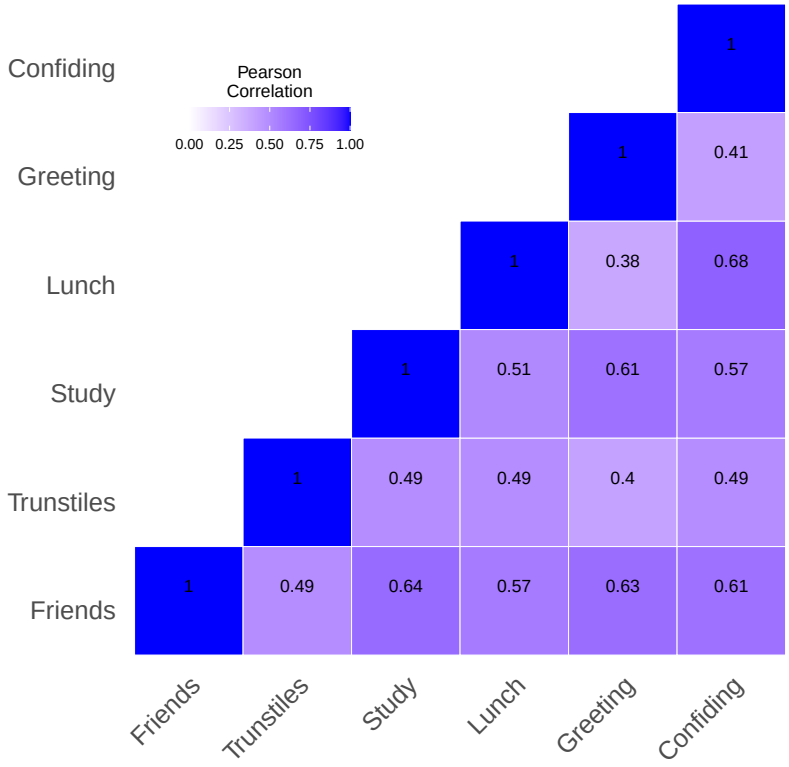
	Recip. trust	**Exp1**	**Exp2**	**Exp3**	**Surv1**	**Surv2**
Constant	0.368^{**}	0.374^{***}	0.372^{**}	0.365^{**}	0.379^{***}	0.372^{**}
	(0.146)	(0.145)	(0.145)	(0.149)	(0.144)	(0.147)
Trust measure	0.004	−0.011	−0.0014	0.004	−0.016	−0.011
	(0.013)	(0.015)	(0.015)	(0.016)	(0.015)	(0.012)
Knew e.o. from before	0.361^{***}	0.366^{***}	0.367^{***}	0.361^{***}	0.363^{***}	0.364^{***}
	(0.066)	(0.065)	(0.065)	(0.067)	(0.067)	(0.067)
Exposure	0.096^{***}	0.096^{***}	0.096^{***}	0.096^{***}	0.096^{***}	0.097^{***}
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Reciprocal trust ×	−0.003	−0.003	−0.003	−0.003	−0.002	−0.003
Exposure	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Socioeconomic	0.013	0.014	0.015	0.014	0.01	0.011
status	(0.019)	(0.019)	(0.019)	(0.019)	(0.02)	(0.019)
Socioeconomic	−0.021	−0.021	−0.021	−0.021	−0.022	−0.02
status Δ	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Sex	-0.063^{**}	-0.062^{*}	-0.062^{*}	-0.063^{**}	-0.068^{*}	-0.064^{**}
	(0.032)	(0.032)	(0.032)	(0.031)	(0.035)	(0.032)
Sex Δ	−0.024	−0.023	−0.023	−0.024	−0.023	−0.023
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Exit exams	0.002	0.002	0.002	0.003	0.001	0.002
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Exit exams Δ	0.007	0.002	0.001	0.007	0.006	0.005
	(0.013)	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)
Hometown	−0.056	−0.05	−0.049	−0.055	−0.053	−0.054
	(0.046)	(0.046)	(0.046)	(0.046)	(0.045)	(0.046)
Hometown Δ	−0.055	−0.054	−0.054	−0.055	−0.055	−0.055
	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)
Observations: 1485						

This table presents the results of linear probability models regressing the first-semester turnstile-based network on the reciprocal trust measures constructed using three alternative measures based on the lab data and two measures based on survey data, as described in Section C. All specifications include all pairwise- and individual-level controls. For readability, the coefficients associated with some controls are omitted. Dyadic-robust standard errors are shown in parentheses.

C.2 Other survey-based networks

Here, we apply our baseline strategy using different outcome variables: the "having lunch together," "studying together," and "confiding in each other" networks. Figure 7 shows the correlation of link presence among all the networks that we consider.

Figure 7: Correlation of link presence among all the networks that we consider.



Note: This figure shows the correlations in link presence between all of our survey-based networks.

Table 11: Robustness regressions: Turnstile-based networks on alternative measures of trust

	Friends	**Studying**	**Greeting**	**Lunch**	**Confiding in**
Constant	0.433^{***}	0.204	0.51^{**}	0.22^{**}	0.294^{***}
	(0.124)	(0.214)	(0.232)	(0.102)	(0.067)
Reciprocal trust	-0.002	-0.011	0.019	-0.01	-0.014
	(0.015)	(0.014)	(0.025)	(0.008)	(0.009)
Knew e.o. from before	0.539^{***}	0.384^{***}	0.689^{***}	0.198^{***}	0.305^{***}
	(0.077)	(0.098)	(0.046)	(0.062)	(0.086)
Exposure	0.067^{***}	0.081^{***}	0.118^{***}	0.041^{***}	0.046^{***}
	(0.016)	(0.015)	(0.016)	(0.011)	(0.011)
Reciprocal trust $\times$	-0.001	-0.004	0	-0.007	0.004
Exposure	(0.015)	(0.012)	(0.013)	(0.011)	(0.011)
Socioeconomic	0.009	0.019	0.042	-0.008	-0.003
status	(0.022)	(0.023)	(0.027)	(0.014)	(0.013)
Socioeconomic	-0.035^{**}	-0.05^{***}	-0.039^{**}	-0.017	-0.026^{***}
status Δ	(0.016)	(0.014)	(0.017)	(0.011)	(0.009)
Sex	-0.049	-0.064	-0.16^{**}	-0.037	-0.08^{**}
	(0.045)	(0.049)	(0.065)	(0.023)	(0.033)
Sex Δ	-0.015	-0.046^{**}	-0.029	-0.014	-0.046^{*}
	(0.026)	(0.021)	(0.037)	(0.021)	(0.025)
Exit exams	-0.027	-0.002	-0.019	-0.011	0
	(0.019)	(0.013)	(0.026)	(0.008)	(0.008)
Exit exams Δ	-0.009	-0.002	0.006	0.007	-0.006
	(0.016)	(0.013)	(0.017)	(0.01)	(0.009)
Hometown	-0.078	-0.006	-0.022	-0.04	-0.056
	(0.048)	(0.05)	(0.071)	(0.038)	(0.036)
Hometown Δ	-0.093^{**}	-0.068^{*}	-0.084^{**}	-0.066^{*}	-0.077^{**}
	(0.045)	(0.036)	(0.042)	(0.037)	(0.033)
Observations: 1485					

This table shows the results of linear probability models regressing the four networks other than friendship on reciprocal trust. All specifications include all the pairwise- and individual-level controls. The results involving the friends network are shown in the first column for comparison. For readability, we omit the coefficients associated with some controls. Dyadic-robust standard errors are shown in parentheses.

Table 11 shows the results of linear probability models regressing the four networks other than friendship on reciprocal trust. All specifications include all the pairwise- and individual-level controls. The effect of reciprocal trust on the formation of the other relationships is small and never

significant at the 10% level. The results for the other controls are consistent with those reported for the friendship network: time of exposure, homophily in socioeconomic status, and hometown are important determinants of relationship formation.

C.3 Other samples

As noted in Section 2, our baseline analysis is based on 1,485 dyads involving the 55 students for whom we have complete data from both stages of the data collection process: the experiment stage and the survey stage, as well as administrative data. The most significant attrition from the starting set of 72 students that participated in the experiment is among students who did not respond to the second-stage survey. Table 12 shows the results of the analysis using larger samples of students, relying on the turnstile-based networks and omitting controls obtained from the second-stage survey or from missing administrative data. Column (i) shows the result of our analysis using all the dyads (2415) among the 70 students who provided complete answers to the trust experiment question.[29] Columns (ii)-(iv) show the results of the analysis with 2016 dyads involving the sample of 64 students who participated in the experiment and for whom we have complete administrative data. Working with this sample forces us to omit some key controls that come from the survey. [30] The size of the coefficients associated to reciprocal trust and their significance do not change upon considering these larger samples.

[29]Two out of 72 students did not provide complete answers to the question of how much money they would return in response to the different amounts that they could have received from the sender.

[30]Specifically, lacking the survey responses for some observations, we are unable to include the pairwise- and individual-level controls related to the following variables: weight, wearing glasses, eyes, hair, height, piercings, attending parties, sibling, smoking and weekly hours of physical activity. These variables were included in the baseline regressions as shown in columns (ii) and (iii) of Figure 7.

Table 12: Robustness regressions: Turnstile-based networks on reciprocal trust and various controls using different samples

	(i)	(ii)	(iii)	(iv)
Constant	0.117^{***}	0.127^{***}	0.144^{***}	0.155^{***}
	(0.013)	(0.013)	(0.021)	(0.023)
Reciprocal trust	0.002	0.003	0	0.002
	(0.01)	(0.01)	(0.011)	(0.01)
Exposure			0.089^{***}	0.089^{***}
			(0.014)	(0.014)
Reciprocal trust $\times$			-0.003	-0.003
Exposure			(0.01)	(0.01)
Socioeconomic				0.014
status				(0.013)
Socioeconomic			-0.012	-0.011
status Δ			(0.012)	(0.011)
Sex				-0.014
				(0.013)
Sex Δ			-0.017	-0.031^{*}
			(0.017)	(0.018)
Exit exams				0.01
				(0.012)
Exit exams Δ			0.008	0.004
			(0.009)	(0.011)
Hometown				-0.006
				(0.019)
Hometown Δ			0.023	-0.038
			(0.023)	(0.027)
Observations	2415	2016	2016	2016

This table presents the results of the analysis using larger student samples, based on turnstile-based networks, and omitting controls obtained from the survey or missing administrative data. Column (i) shows the results using all 2,415 dyads among the 70 students who provided complete answers to the trust experiment question. Columns (ii)-(iv) show the results using 2,016 dyads involving the 64 students who participated in the experiment and for whom we have complete administrative data. This sample omits some key controls from the survey. Dyadic-robust standard errors are shown in parentheses.