Essays in Education, Peer Effects and Decision Making

by

Wenjing Xu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Information and Economics)
in The University of Michigan
2017

Doctoral Committee:
     Assistant Professor Erin L. Krupka, Co-Chair
     Professor Brian A. Jacob, Co-Chair
     Associate Professor Tanya Rosenblat
     Professor Charles C. Brown
     Assistant Professor Stephen G. Leider

Wenjing Xu

wjxu@umich.edu

ORCID ID: 0000-0002-4659-951X

Contents

List of Tables

List of Figures

List of Appendices

ABSTRACT

Peers play an important role in shaping behavior in many contexts. In this dissertation, I study the role of peer interactions on key educational outcomes. In Chapter One, I implement a field experiment to examine the role of social interactions in creating spillover effects. Spillover effects happen when individuals are indirectly affected by an intervention through exposure to other treated individuals. In the experiment, I randomly assign college students in an introductory statistics course to a low-cost behavioral intervention. Treated students receive advice and prompts to make exam study plans. I measure a naturally formed peer network and exploit the exogenous variation in exposure to the intervention in order to *causally* estimate spillover effects on study behaviors that are transmitted through study partners. I construct a simple social learning model. Additional behavioral evidence further supports the model and I show that the positive spillover effects on untreated students are mostly driven by treated partners who have high beliefs about the return to the applet usage. Surprisingly, I find circumstances under which social interactions reduce the treatment effect. Taken together, this paper provides causal evidence of spillover effects on behavior due to peer interactions and unpacks the complexities behind spillover effects. My results highlight that in networked environments, policy makers should take peer effects into consideration not only to correctly evaluate, but also to leverage social learning to maximize policy impacts.

In Chapter Two, I study a natural experiment that randomly assigns students into study groups and estimate the effect of studying with peers of certain characteristics. I find little evidence that peers' background academic performances have significant effects on the course final grade using the traditional linear in the mean model. I find that the group gender mix has an economically and statistically significant impact. In particular, being in groups with

more female peers leads to an increase in the course grade for both female and male students. I exploit the course website's log data, and find that one is more likely to download course materials when in more female groups. This is a plausible mechanism through which the gender mix affects the grades. I also find that studying with peers from another lecture section marginally improves one's course grade. My paper therefore provides practical suggestions for assigning students into study groups.

In Chapter Three, we use a longitudinal survey design and follow college freshman, in order to provide evidence for two separate mechanisms (homophily and influence) behind similarity in peers' behaviors. This paper demonstrates these effects for the subtle (but broadly important) underlying economic preferences, rather than the observable but potentially domain-specific behaviors previously studied. Subjects participate in three waves of an online experiment where we elicit their social network using an incentive compatible mechanism and then measure participants' levels of altruism, willingness to take risks, and willingness to delay rewards using diagnostic tasks. We find that subjects' risk and time preferences are significantly positively correlated with the preferences of their friends, consistent with peer influence on preferences. Additionally, we find that changes in subject's social networks are significantly influenced by social preferences. Subjects are more likely to add someone as a friend, and less likely to drop as a friend, the more similar their social preferences are.

Chapter 1

With a Little Help from Your Friends: A Field Experiment on Spillover Effects of Making Study Plans on Student Learning

1.1 Introduction

Social networks characterize pathways for behavior to propagate. Peers transmit information and influence every sphere of choice: academic performance, job search and career choice, productivity at work, product choice, and voting behavior (Bandiera et al., 2010; Fafchamps and Vicente, 2013; Marmaros and Sacerdote, 2002; Mobius et al., 2005; Sacerdote, 2001). In experimental settings, policy makers and economists have been interested in estimating both treatment effects as well as spillover effects. Spillovers happen when individuals are indirectly affected by an intervention through exposure to other treated individuals. While many have examined spillover effects on untreated individuals, few have studied spillovers within the treatment groups except Fafchamps and Vicente (2013) and Fafchamps et al. (2013). My experimental design allows me to ask whether social interactions amplify or abate the direct treatment effect. This can inform policy makers of better treatment assignments. In addition, unpacking the mechanisms through which peer interactions generate spillovers is also important for policy makers to better leverage social network to scale up policy impacts.

In this paper, I use a field experiment to study the spillover effects of an educational in-

tervention on student learning. I situate my study in an introductory college statistics course because it offers a large-scale social network and objective performance measures. Classrooms are a principal workplace for students, and a major domain wherein academic relationships form between students. With the experimental design, I measure two types of spillover effects: 1) how untreated students change behavior due to exposure to the intervention through peers and 2) whether social interactions reinforce or mitigate the treatment effect among the treated. The intervention features planning activities designed to help students engage in active studying for exams. Combining social network elicitation and the random assignment of the intervention, I causally identify spillover effects from existing self-selected study partners on a variety of outcomes.[1]

To define *relevant* peer groups, I adopt a network elicitation protocol that allows students to identify study partners that they interact with. I ask students to name anyone with whom they talk about the course or study for exams. I directly elicit a student's social network to define *relevant* peers while previous studies rely on geographic closeness to define peers (Bobba and Gignoux, 2016; Duflo and Saez, 2003; Miguel and Kremer, 2004). I do so because existing evidence suggests that self-selected peer groups may be critical for researchers to "look in the right place" in order to capture peer influence. Earlier work in psychology and sociology suggests that individuals with whom one associates or with whom one shares a similar identity are more influential (Cialdini and Garde, 1987; Granovetter, 1985; Tajfel and Turner, 1979).[2] Recent economics studies show that peer influence can vary with how researchers construct peer groups (Carrell et al., 2009), and that policy recommendations based on exogenous group assignment, without accounting for self-selected peer group formations, can backfire (Carrell et al., 2013). The study partnerships I elicit are pertinent for spreading course related information and study behavior.[3] These relationships, and the larger networks they create, play

---

[1] According to Harrison and List (2004), my experiment can be classified as a natural field experiment where "the environment is one where the subjects naturally undertake these tasks and where the subjects do not know that they are in an experiment."

[2] "The term identity is used to describe a person's social category - a person is a man or a woman, a black or a white, a manager or a worker. The term identity is also used to describe a person's self-image. It captures how people feel about themselves, as well as how those feelings depend upon their actions." (Akerlof and Kranton, 2005, p12)

[3] Student feedback from previous semesters shows that students who claim to have the "perfect study partner" score higher than those who claim to "know someone that they would like to study with". Students knowing no one in the class score the lowest. See Appendix Figure A.1. This evidence suggests that having a study partner,

an important role in shaping students' educational experiences and outcomes (Betts and Zau, 2004; Hanushek et al., 2003; Hoxby, 2000; Wentzel, 1998). From here on, I will refer to study partners as "partners" for short.

The intervention works in the following way. Before the exams, treated students receive both advice on the importance of planning and prompts to plan time and material use in preparation for exams. There are two types of prompts. My "time use" planning prompts ask students to write down reasonable and specific time plans about when to study for exams. My "material use" planning prompts ask students to choose what study materials to use and then to write down reasons for choosing certain materials to study for exams. These planning prompts have been shown to effectively nudge individuals towards a wide range of desirable behaviors (Rogers et al., 2015). Planning ahead reduces students' uncertainty in how they will study, by inducing them to commit to a plan. I focus on this low-cost intervention that is built upon the insights of behavioral science whereas previous studies have focused on spillovers of more substantial interventions such as providing subsidies for poor kids to go to school (Bobba and Gignoux, 2016) or offering medical treatments to eradicate life-threatening diseases (Miguel and Kremer, 2004). However, these programs are costly and not applicable to students in my experimental context (a U.S. flagship state university).

My identification relies on the random assignment of the intervention at the individual level and exploits the resulting exogenous variation in the exposure to the intervention through peers, conditional on the total number of partners a student has. My approach is different from much of the literature that causally estimates peer effects by using random peer group assignment (Carrell et al., 2009; Foster, 2006; Lu and Anderson, 2015; Sacerdote, 2001; Xu, 2016; Zimmerman, 2003).[4] Instead of manipulating the network, I manipulate an individual's exposure to the intervention in the elicited partner network.

With the intervention targeting how students prepare for exams, I choose to examine stu-

---

and knowing who that partner is, influences course achievement.

[4]This literature faces the challenge of defining relevant peers. Stinebrickner and Stinebrickner (2006) argue that randomly assigned roommates should not be viewed as relevant peers since a student "may find that there are many other students at the school who have more compatible interests and preferences than his/her [randomly assigned] roommates". This might also be one reason why Sacerdote (2001) finds that randomly assigned roommates have an impact on both college GPA and decisions to join fraternities, but Foster (2006) finds no evidence supporting influence from roommates.

dents' usage of an online learning technology (hereafter referred to as "the applet") as the primary outcome of interest. This applet has become a popular practice tool for students and is likely to be a social tool that partners use together. The communal or cooperative nature of the applet makes its usage subject to peer influence. The availability of high-frequency and objective online usage data helps unpack channels through which the partners affect a student's technology adoption. The main outcomes are the take-up of the applet and the intensity of use. I also collect other self-reported study behavior such as time spent on studying for exams.

I find that the intervention has an overall positive effect on the applet usage. Without any treated partners, a treated student is 15% more likely to use the applet compared to an untreated student. This the direct treatment effect. On the intensive margin, a treated student increases the frequency of usage by 35%. More importantly, I find a positive spillover effect on untreated students. An untreated student is about 5% more likely to use the applet when they have an additional treated partner. The spillover effect is about 30% of the direct treatment effect, both on the extensive and intensive margin. Contrary to the positive spillover effects on untreated students, I find overall insignificant spillover effects on treated students. In other words, treated students are much less likely to be affected by the exposure to treated partners than untreated students. I then exploit variations in tie strength and show that treated partners who tend to interact more intensively (e.g. mutually listed partners, exam study partners) are more influential than their counterparts (e.g. unilaterally listed partners, general talking partners).

These results motivate a model with social learning as a potential channel for spillovers, whereby students share beliefs about the usefulness of the applet.[5] I construct a simple social learning model where students choose applet usage but face uncertainty of the return to usage. I assume the intervention decreases belief uncertainties in the return because treated students have already been prompted to make a plan. With social learning, students learn from their partners about the return and update beliefs. Following the stylized results, I allow the strength of social learning to differ based on tie strength. Not only can the model explain the observed spillover effects, it also generates an additional prediction for testing: The higher a partner's

---

[5] I do this to follow previous work on tool adoption, which has shown that learning the parameters of a new technology from others is a key information transmission channel.

belief is about the return, the more the partner will encourage an individual to use the applet.

To test the model, I use the study plans that treated students write down to infer their beliefs about the return from using the applet. I classify those who plan (do not plan) to use the applet as having a high (low) belief. I find supporting evidence for the model which offers additional policy implications. First, I find that the positive spillover effects on untreated students mostly come from treated partners who also plan to use the applet. This evidence implies that exposure to treated partners is not a sufficient condition for spillovers to happen. Rather exposure to the "right" individuals – those who also have a willingness to use the applet is a key for leveraging social networks to scale up policy effects. Second, I find circumstances under which peer interactions backfire – social interactions can mitigate the treatment effect.[6] When a pair of treated partners have contrasting plans (i.e. one plans to use the applet while the other one does not), the negative effect from the one who does not plan to use the applet is large enough to offset the treatment effect. This highlights a situation where treating everyone might not be ideal and policy makers need to be more careful about treatment targeting.

Throughout, I focus the discussion around social learning as a channel that potentially transmits spillover effects within a network. There are, however, other plausible explanations. For instance, the spread of applet take-up can happen if there are economies of scale or complementarities in using the applet.[7] While the existing data has not yet allowed me to verify or falsify the alternative explanations, I note that this alternative explanation is unlikely to explain all the results, especially why social interactions can mitigate treatment effect among the treated.[8]

Regarding performance outcomes, treated students prepare better for lab discussion sections as measured by the scores from their pre-lab assignments. Having treated peers also positively affects untreated students' pre-lab assignment scores. The relative size of the spillover effect

---

[6]The literature on the interaction between spillover effects and intervention offers mixed results. Babcock and Hartman (2010) find that peer interactions reinforce the effect of financial incentives on students' propensity to exercise. Fafchamps and Vicente (2013) also find positive reinforcement of a voter education program on violence related perceptions. In contrast, Fafchamps et al. (2013) find a negative reinforcement effect of a similar voter education campaign on voter turnout.

[7]Economies of scale characterize a situation where when one partner adopts the applet, other partners' cost of using the applet decreases. Use complementarities occur if one partner's marginal utility of using the applet increases when another partner adopts it.

[8]I use Bramoullé et al. (2009) to identify the endogenous peer effect parameter but I am constrained by weak instruments.

compared to the treatment effect is about 30%, similar to that from the applet use. While the intervention improves this intermediary performance outcome, I do not find any statistically significant treatment effects on exam scores and course grades. The self-reported study habits suggest that treated students might be substituting between the applet and other materials. The lack of evidence of spillover effects on aggregate performance outcomes calls for future work on the effectiveness of the applet, and also highlight the benefit of collecting behavioral data to uncover policy spillovers.

This study contributes to a growing volume of literature on estimating peer effects in higher education, with a particular focus on self-selected peers. The ability to use objective and real time applet usage allows me to better measure behavior and provide evidence of social learning, adding to the few empirical studies that explore potential mechanisms behind peer influence. My results highlight the necessity of taking peer effects into consideration when assigning interventions in a networked environment.

1.2   Study Context and Experimental Design

1.2.1   Study Context

My study context is a large introductory statistics class at a US four-year flagship public university. This course is required for many science major students.[9] The course is worth 4 credits for three hours of lectures and an additional one-hour lab session each week. There are 6 lecture sections taught by 4 different lecturers and 68 lab sections taught by 36 different graduate teaching assistants. The syllabus provides students with detailed and important course information such as course learning objectives, schedule, grading methods, three exam dates, and other important details. Hence there is little uncertainty about the course contents. The course has an online website called *StatsOnline*, which has been used previously for multiple semesters. Both the instructors and the teaching assistants encourage students to use this website for course related activities such as looking up exam scores, browsing study tips, and doing online practice exercises. In short, this website is an important part of the course and students use it frequently.

---

[9]These majors include Computer Science, Ecology and Evolutionary Biology, Information Science, Bio-Psychology, Neuroscience, and Environment majors.
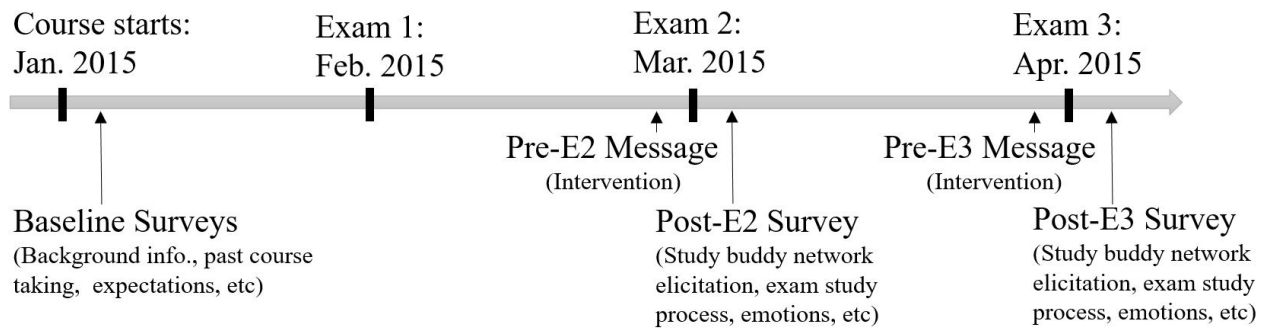
Figure 1.1: Timeline

The course is not curved.

1.2.2 Experimental Design

My experimental design features three waves of surveys and two waves of intervention messages throughout the semester. Figure 1.1 is a detailed timeline for the course. It describes the timing of each exam and the experiment implementation.

As can be seen from the timeline illustration, at the beginning of the semester, all registered students are invited to complete two baseline surveys.[10] In the first survey, I collect background information such as whether they have declared a major and student organization affiliations. I get demographic information and official high school and college GPA data from the Office of the Registrar. Then in the second survey, I ask students about their past statistics course experience and expectations going into the semester. These data are used for the balance check.

The first exam (Exam 1), takes place roughly a month after the semester starts. 10 days before Exam 2, I send out a Pre-E2 Message through *StatsOnline* that contains the intervention to ask students to plan for exam preparation. Within each lecture section, I randomly select 25% of the students to be in the baseline group. The baseline group students receive a Pre-E2 Message which is a reminder of the exam time and date. The remaining 75% of the students are in the treatment group. In addition to the reminder, the treatment group students also receive advice and prompts to make study plans. The purpose of the prompts is to engage students

---

[10]The teaching assistants reserve the last 10 minutes of the first two lab sessions for students to take these surveys.

with active and self-regulated learning which is an important input in education (Duckworth and Seligman, 2005; Pintrich and De Groot, 1990). In particular, a third of the treated students receive advice on the importance of planning time use and see an example of a reasonable time plan. Then students do a time use planning activity - they write down a specific and reasonable time plan for exam preparation. Another third receive advice on the importance of planning material use and see a sample reason of why practice exam questions are an effective study material for exam preparation. Then students do a material use planning activity - they select materials that they believe are most effective for exam preparation and provide justifications for their selections. The remaining one third do both planning activities. Students can review their plans throughout the semester via *StatsOnline*. Both activities prompt students to make plans. Those planning prompts can encourage desirable behaviors in many different contexts.[11] In the course context, these planning prompts may function as a commitment device or a reminder and thus change how students study for exams. I argue that the prompts can help reduce students' uncertainty regarding how to study for exams, by inducing students to commit to a plan and helping students avoid distractions. For example, through planning, students can better foresee and remove logistical barriers. This can help students avoid distractions due to planning fallacy (i.e. underestimate how long studying takes).[12]

Since this paper aims to estimate the spillover effects from peers' exposure to the prompts, I will pool the three arms together. In Figure A.3a and A.3b I show screenshots of the intervention prompts.

I send out the Post-E2 Survey online to elicit the study partner network one day after Exam 2. In particular, I ask students whom within the course they have talked to about this course so far (such as discussing lecture slides and working on homework questions). Figure 1.3 displays the user interface of the study partner elicitation. Students can search by name and the interface displays both exact matches and close matches in the case of typos.[13] Using a

---

[11]These behaviors include getting flu vaccines (Milkman et al., 2011), voting (Greenwald et al., 1987; Nickerson and Rogers, 2010), purchase of a product (Morwitz et al., 1993) and so on.

[12]Other benefits of the prompts can also contribute to reducing the uncertainty. Planning can help strengthen their self-efficacy (i.e. belief in succeeding) and ease exam anxiety, which can reduce the uncertainty in exam preparation process as well. For a detailed discussion on why prompting people to make plans is effective at inducing positive behavioral changes, see Rogers et al. (2015).

[13]I have received emails about not being able to find names from the system. In most situations, students are

Planning

Because planning helps you achieve your goals, please describe how you plan to prepare for Exam 2.

Your plan should be as **specific** as possible, **realistic**, and **useful** in guiding your exam preparation.

For example:
Wednesday, 3/18, 3pm - 5pm, Library: Work on practice exam questions from 1 past exam.
Friday, 3/20, 1pm - 3pm, Home: Read lecture notes for 2 topics.

(a) Prompt to Plan Time Use

Please select the Stats 250 resources that you think will help you prepare for Exam 2 effectively.

- ☐ **Lecture notes**
- ☐ **Video captured lectures (Blue Review or Adobe Connect)**
- ☐ **Past required HW problems**
- ☐ **Past recommended HW problems**
- ☐ **Practice exam questions (from lab workbook)**
- ☐ **Past exam questions through Problem Roulette**
- ☐ **Name That Scenario Applet**
- ☐ **Lab materials (ILPs)**
- ☐ **Textbook readings**
- ☐ **Yellow formula card**
- ☐ **Discussions with other students in the class (e.g. study group)**
- ☐ **Office hours held by a lecture instructor**
- ☐ **Office hours held by GSIs**
- ☐ **Asking questions in class**
- ☐ **Private tutoring**

Usefulness

You have chosen:
(1) For each resource you selected, please describe **_why_** you think it will be useful for your exam preparation.
(2) Be specific about how each resource will help you study effectively.

For example: "Doing practice exam questions will allow me to practice on problems similar to what I will encounter on my exam. While doing the problems, I will look out for how I am expected to apply my knowledge. This will give me an idea of what material I should review in the lecture notes and how to apply the formulas I learn."

(b) Prompt to Plan Material Use

Figure 1.2: Intervention Prompts

Before we conclude the survey, we would like to gain an insight into how students interact.

Please recall all the classmates whom you have talked to about this course such as discussing lecture slides, homework questions, or anything related to the course.

Once you have some in mind, please search for their first and/or last names in the corresponding search boxes. The system will search and display a list if fewer than 100 students are matched.

After you find a match, please click (+add) and he/she will appear under the "I talked to..." heading. Then please select how frequently you two talked. You can remove him/her by clicking (-remove) at any time.

I talked to...   about this course

If you cannot find a name by searching, please email: wjxu@umich.edu

Search: First Name	Search: Last Name
erin	
found: 11	found: 58
Exact matches:	**Close matches:**
(+add)Erin E	(+add) Adrienne

Figure 1.3: Eliciting Study Partners

similar interface, students also select partners within the course whom they study together with for Exam 2 specifically. I choose such an elicitation protocol to make it as easy as possible for students to list names. I do not limit how many study partners one can list. I also provide no incentives for listing partners that are more likely to reciprocate. Other studies have constrained how many friends one can list and/or have provided incentives for mutually listed friends (e.g. Leider et al. (2009)). I do not provide incentives for the study partner elicitation since it does not fit well with the course design.[14]

I choose to elicit the network after the intervention to avoid focusing students' attention on studying with others. If I were to elicit the network before or together with the Pre-E2 Message, I might prime control students to think about study partners since they do not see other alternatives for preparing for exams in the Pre-E2 Message.

One potential concern of eliciting the network after the intervention is that the intervention may change whom students study with. If this is happening, then treated students' partners will be different from those of the untreated. In the results section, I do not find such differences. I

---

trying to find someone who took the course in a previous semester. Sometimes students cannot remember another student whom they happened to talk to during pooled office hours.

[14]The study partner elicitation yields an over 50% agreement rate on the study partner relationship, which indicates that the elicitation is quite successful even without the incentives. For a survey on network elicitation methods, see Brañas-Garza et al. (2013).

also show that the treatment assignment cannot predict study partnerships or changes in study partnerships between the two exams.

For Exam 3, I repeat the same pre-exam message and post-exam survey. I keep the treatment assignment constant between Exam 2 and 3. To incentivize participation, all the pre-exam messages and the post-exam surveys are awarded with 2 extra homework credits out of 100 total homework points.[15] Students have one week to complete the surveys online.

1.2.3 The Applet

The key study behavior I focus on is the use of an online learning applet with practice questions to help students understand applications of statistical concepts and tests. The applet is embedded in *StatsOnline* and is available to students after the first exam. Each question presents students with a scenario and then students choose appropriate concepts or appropriate statistical tests to apply given the scenario. After students submit the answers, the applet will provide explanations on why the chosen answer is right or wrong. Figure 1.4 shows a screenshot of a question. The applet is designed for this statistics course and has been used previously.

The applet has the following features that make it an outcome of interest. First, students find the applet to be helpful for exams because the applet questions are similar to those on the exams.[16] Second, although the applet is rated as helpful, not every student uses the applet yet. In the plans students write, about half mention to use the applet. The applet is the fifth mostly mentioned material in student's plans.[17] These two stylized facts suggest that there is room for the applet usage to be affected by the intervention.

Third, the applet is also a social tool which partners use together when studying together. Figure 1.5 provides descriptive evidence to support this claim. The upper solid line plots out the fraction of partner pairs that have ever used the applet (y-axis) within a window of 10, 20, 30,..., up to 120 minutes (x-axis). About 10% pairs use the applet within the ten-minute window and

---

[15]Homework accounts for about 10% of the final course grade.

[16]Based on two semesters of course feedback data, the median helpfulness rating of the applet is 3 on a scale of 1 to 4.

[17]The five frequently mentioned materials are, in descending order, lecture notes, formula cards, past and practice exams, past homework, and the applet.

Question I

Several runners claim they can make it across the city faster than cars, since they are not restricted to roads and traffic lights. To test the claim, 8 runners and 6 cars attempt to (legally) traverse the city in as little time as possible.

One Mean is NOT the correct answer...

The response variable, time to traverse the city, is quantitative and therefore best summarized by the mean. Note, there is no pairing as we are interested in comparing the means of the two groups rather than a particular runner to a particular car. A two-sample t-test for a difference in population means should be used to infer about the difference in population means.

RETURN TO START   NEXT QUESTION

One Proportion
Two Proportions
One Mean
Paired
Independent T-test

Figure 1.4: A Sample Question from the Online Learning Applet
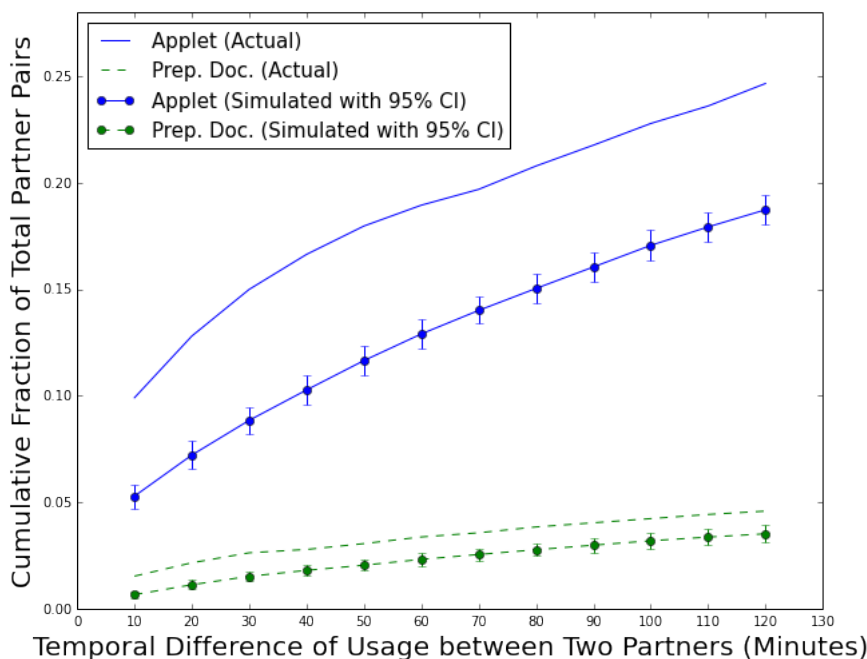
Figure 1.5: Potential Contagious Applet Usage

25% use the applet within two hours.[18] I then contrast the upper solid line to the lower solid one with confidence intervals. This line shows the fraction of randomly paired study partners using the applet within a certain window.[19] The obvious gap between the two solid lines suggests that study partners are more likely to use the applet together. Similarly, I generate two dashed lines for fractions of pairs reviewing the exam preparation summary document through *StatsOnline* within a certain window.[20] The fact that the solid lines are much higher than the dashed lines shows that using the applet is more social than reviewing summary documents. This feature suggests the applet usage might be subject to peer effects.

1.2.4 Datasets

In my analysis, I will use the following four datasets:

1) The dataset obtained from the Registrar and the Office of Greek Life: This dataset has demographic and academic achievement data, and Greek organizations membership.

2) The survey dataset: This dataset has students' self-reported past experience with statistics

---

[18]50% pairs use within a day.

[19]In each simulation, I keep total partners a student has the same while rewiring the links. The confidence intervals are from 100 simulations.

[20]This summary document includes advice for studying for exams, and contents covered by each exam.

courses, expectations about the courses and other background information from the baseline surveys, exam study plans from the pre-exam surveys, the self-reported study partnerships and study behaviors from the post-exam surveys.

3) The applet usage dataset: This dataset has objective measures of applet usage obtained from the log file of *StatsOnline*.

4) The performance dataset: This dataset has detailed course performance measures are available from the instructors.

1.3   The Network and Data Descriptives

In this section, I first provide an overview of the endogenous study partner network. I look at the factors that predict partnerships. Then I describe the final sample for the empirics.

1.3.1   The Study Partner Network Descriptives

A total of 1564 (73%) out the 2156 registered students responded to the network elicitation questions after Exam 2 and 1896 (88%) respond to the Post-Exam 3 Survey. Among the post-exam survey respondents, a third of them do not list any study partners. The remaining two-thirds list on average two study partners, and are listed back as a study partner as well with a 50% probability.[21] Even though my network elicitation protocol is not incentivized, the reciprocation rate of 50% is high compared to previous studies with network elicitations (Chandrasekhar and Lewis, 2011). I then construct an "or" network where two students are considered as study partners as long as one of them lists the other either as a talking partner or an exam study partner. Previous literature has primarily used the "or" relationship to construct network (e.g. Conley and Udry (2010); Mobius et al. (2005)). The advantage of using the "or" relationship is that it reduces measurement errors if students forget to recall every study partner. The disadvantage of using the "or" is that it also captures more weak ties than strong ties. As is discussed in section 4.1, the spillover effects can be downward biased due to the "or" definition. I will exploit the directionality of the partnerships and compare the effects from mutually versus unilaterally listed partners to argue for social learning as a potential mechanism.

---

[21]This is conditional on that a listed partner also takes the survey.

Figure 3 displays a visualization of the elicited "or" network after Exam 2. Dark nodes represent untreated students and white represent treated students. A link connects two study partners as long as one of them lists the other as a study partner. For exposition reasons, in this figure I exclude isolated nodes. The nodes are not densely connected to each other based on the colors, which partially alleviates the concern that the treatment affects network formation.

The elicited network exhibits characteristics that are consistent with the previous literature: students sort into study partnerships based on common characteristics - a phenomenon known as homophily.[22] Table 1.1 examines the predictors for one to list another student from the same lecture section after Exam 2.[23] In column (1), each row is a separate logit regression. The dependent variable is whether one lists another student from the same lecture as his/her study partner, and the predictor is an indicator for sharing some common characteristics or a characteristic of a potential study partner. Each regression also controls for individual gender and high school and college GPA. All the indicators for two people sharing the same attributes are positive and significantly different from zero, indicating homophily in the network - students are more likely to study with someone sharing similar attributes. Three most predictive factors are being in the same lab section, being in the same student organization, and being of the same gender. However black students are less likely to be listed as a study partner ($\beta =$ -0.0006, $p<0.05$). Column (2) is a joint logit regression with all the characteristics included. A student is on average 0.21% more likely to name a study partner from the same lab section than from another section. There is some evidence suggesting that a student selects partners based on their academic performance measures. The estimates on a partner's high school GPA and college GPA are marginally statistically significant.

Columns (3) and (4) examine the predictors for link reciprocity - how likely it is that two students both agree that they are study partners. A student is more likely to be listed back as a partner as well by another student of same gender ($\beta = 0.074$, $p < 0.001$). High school GPA becomes a strong predictor for link reciprocity ($\beta = 0.226$, $p < 0.001$). Being in the same

---

[22] Also consistent with the education literature (Bruun and Brewe, 2013; Calvo-Armengo et al., 2009; Maroulis and Gomez, 2008), network position can predict student achievements. Table A.4 show that the correlation between the centrality measures and grades become even more positive and significant over time.

[23] Due to the large size of the course, I choose to do so at the lecture section level.

Figure 1.6: A Visualization of the Study Partner Network

Notes: This is a part of the "or" network using response from the Post-Exam 2 Survey. I pool the talking and exam study partners together. Isolated students are not shown in the graph. Treated students are white and untreated students are in gray.

student organization does not predict link reciprocation ($\beta = 0.046$, p > 0.1). The coefficient in front of the total number of study partners a student lists is negative ($\beta = -0.093$, p < 0.01). A student is less likely to be listed back the more partners he/she lists. One plausible explanation is that students are listing partners in order based on tie strength.

Table 1.1: Predictors for Partnerships Elicited in Post-Exam 2 Survey

Logit	(1) Unilaterally List (Source→Target)	(2) Unilaterally List (Source→Target)	(3) Mutually List (Source↔Target)	(4) Mutually List (Source↔Target)
Same Gender	0.0012***	0.0009***	0.066**	0.074***
	(0.0001)	(0.0001)	(0.028)	(0.028)
Same Student Org.	0.0020***	0.0013***	0.0423	0.046
	(0.0001)	(0.0001)	(0.029)	(0.030)
Same Cohort	0.0009***	0.0006***	0.023	0.0304
	(0.0001)	(0.0001)	(0.024)	(0.024)
Same Ethnicity	0.0008***	0.0006***	0.014	0.009
	(0.0001)	(0.0001)	(0.024)	(0.024)
Same Major	0.0009***	0.0005***	0.006	0.009
	(0.0001)	(0.0001)	(0.024)	(0.024)
Same Lab Section	0.0025***	0.0021**	0.012	0.001
	(0.0001)	(0.0001)	(0.032)	(0.032)
Target's HS GPA	0.0009***	0.0005*	0.208***	0.226***
	(0.0003)	(0.0003)	(0.076)	(0.079)
Target's College GPA	0.0007***	0.0004*	-0.018	-0.090
	(0.0001)	(0.0002)	(0.031)	(0.070)
Target Being Female	0.0003***		0.114***	
	(0.0001)		(0.027)	
Target Being Black	-0.0006**		-0.059	
	(0.0003)		(0.080)	
Total # Partners Listed			-0.024***	-0.0224***
			(0.006)	(0.006)
Obs.		727,578		1782

Notes: The results are based on responses to the network elicitation after Exam 2. "Same Student Org." equals to 1 if two students are in the same fraternity or sorority organization. "HS GPA" is the high school GPA. "College GPA" is the official cumulative college GPA before the winter 2015 semester. Each regression controls for gender, HS GPA and College GPA. In column (1) and (2), the independent variable of the logit regression is whether a student lists another student from the same lecture ("Target") as his/her study partner. In column (1), each estimate is from a separate univariate regression. Column (2) is a joint regression with all the independent variables. In column (3) and (4), the independent variable of the logit regression is whether a listed student ("Target") also lists the student back. Column (4) is a joint regression with all the independent variables. All the estimates are average marginal effects.

In Table A.2 I compare the two networks, one elicited after Exam 2 and the other elicited after Exam 3, in order to see what predicts the network dynamics. First, a student's treatment status is not predictive of the deletion of the partnership. A treated student is as equally likely as a control student to drop a partner or to be dropped by a partner. Second, new partnership establishment is also unlikely to be driven by the treatment assignment. I consider students from the same lab section as potential study partners since being in the same lab section is

predictive of nominating a partner in Table 1.1. A control student and a treated student are equally likely to add or to be added by someone from the same lab section. This is evidence suggesting that it is unlikely that students are strategically changing partnerships in response to the treatment.[24] At the student level, there are changes in partnerships. At the whole network level, the network property measures are similar between Exam 2 and 3.[25] In the following sections I use the "or" network elicited after Exam 2.

1.3.2 Sample and Balance Check

I include the 1564 students who respond to the network elicitation questions after Exam 2 in my final sample of analysis. I choose not to include the non-respondents. As can be seen in Table A.1, the nonrespondents are more likely to be male, black and upperclassmen (sophomore, junior, and senior students). These nonrespondents are more likely to withdraw from the course and those who stay also have lower exam scores compared to the respondents.

Of the final sample, 407 (26%) are thus in the control condition, and the rest 1157 (74%) are in the treatment condition. Table 1.2 presents the summary statistics of the treated and the untreated students in this class.

First, this table shows that students' demographic variables are balanced between the untreated and the treated group. Students do not differ on their academic background performances measured by high school GPA, college GPA, and Exam 1 score. Course expectations in terms of studying with someone else and the amount of time to be spent on this course also do not differ by the treatment assignment.

Secondly, treated and control students are equally likely to have actually received the pre-exam messages. I do so by tracking whether students view and especially respond to the prompts in the pre-exam messages. As is shown in Table 1.2, both the Pre-Exam 2 Message and the Pre-Exam 3 Message have a response rate over 80% and the response rates do not differ by the treatment assignment. I then qualitatively check treated students' responses to the prompts. Almost all the treated students give reasonable responses to the prompts.

---

[24]See Table A.2.

[25]See Table A.3.

Table 1.2: Balance Table

	Untreated			Treated			T Test		OLS	
	Mean	S.D.	N	Mean	S.D.	N	Diff.	t stat.	Coeff. on #Treated	(S.E.)
Panel A: Demographics										
Age	19.3	1.38	407	19.2	1.26	1157	-0.105	-1.42	-0.003	(0.032)
Exam 1 Score	60.1	9.82	407	60.9	9.61	1157	0.812	1.46	0.554	(0.374)
High School GPA	3.842	0.220	407	3.841	0.196	1157	-0.001	-0.13	-0.001	(0.006)
College GPA	3.421	0.409	401	3.411	0.416	1137	-0.010	-0.43	0.002	(0.012)
Female	0.597	0.491	407	0.571	0.495	1157	-0.026	-0.90	-0.022	(0.018)
White	0.671	0.471	407	0.698	0.459	1157	0.028	1.04	-0.013	(0.019)
Black	0.039	0.195	407	0.034	0.181	1157	-0.006	-0.53	0.006	(0.005)
Asian	0.270	0.445	407	0.251	0.434	1157	-0.020	-0.78	-0.006	(0.019)
Hispanic	0.025	0.155	407	0.041	0.197	1157	0.016	1.49	-0.005	(0.007)
In CAS*	0.825	0.380	406	0.847	0.360	1157	0.022	1.04	0.005	(0.012)
Major Declared	0.319	0.467	407	0.309	0.462	1157	-0.011	-0.41	0.002	(0.015)
Live Off Campus	0.409	0.492	386	0.429	0.495	1092	0.019	0.66	-0.005	(0.017)
Panel B: Course Specific Responses										
#Semesters of Stats.	0.535	0.816	385	0.514	0.786	1083	-0.021	-0.44	0.026	(0.031)
#hr/week Study	19.31	8.233	384	18.77	8.292	1082	-0.217	-1.15	0.354	(0.300)
#hr/week Study Stats.	5.30	3.374	384	5.08	3.108	1083	-0.544	-1.11	0.001	(0.122)
Will Study w. Others	0.802	0.399	384	0.791	0.407	1083	-0.011	-0.45	0.004	(0.013)
Panel C: Pre-Exam Message Take-up										
Pre-Exam 2	0.929	0.258	407	0.935	0.246	1157	-0.015	-0.80	-0.014*	(0.007)
Pre-Exam 3	0.892	0.920	407	0.920	0.272	1157	-0.001	-0.04	-0.012	(0.010)
Panel D: Partner Characteristics										
#Total	2.6	2.06	407	2.8	2.09	1157	0.156	1.30		
#Treated	1.9	1.66	407	2.1	1.72	1157	0.155	1.55	N/A	
Exam 1 Score	60.25	0.345	353	60.42	0.209	1035	0.170	0.41		
Female	0.353	0.612	353	0.577	0.012	1035	-0.036	-1.47		

Notes: All the variables except "Live Off Campus" in Panel A are from the Registrar dataset. CAS stands for College of Arts and Science. The rest of the variables, except partners' scores and gender, are from the survey dataset. The last two columns present the coefficients and the standard errors of the treatment dummy in OLS regressions for balance checks. A joint F-test with the first and the second group of variables has an F-statistic of 1.11 and a p-value of 0.28.
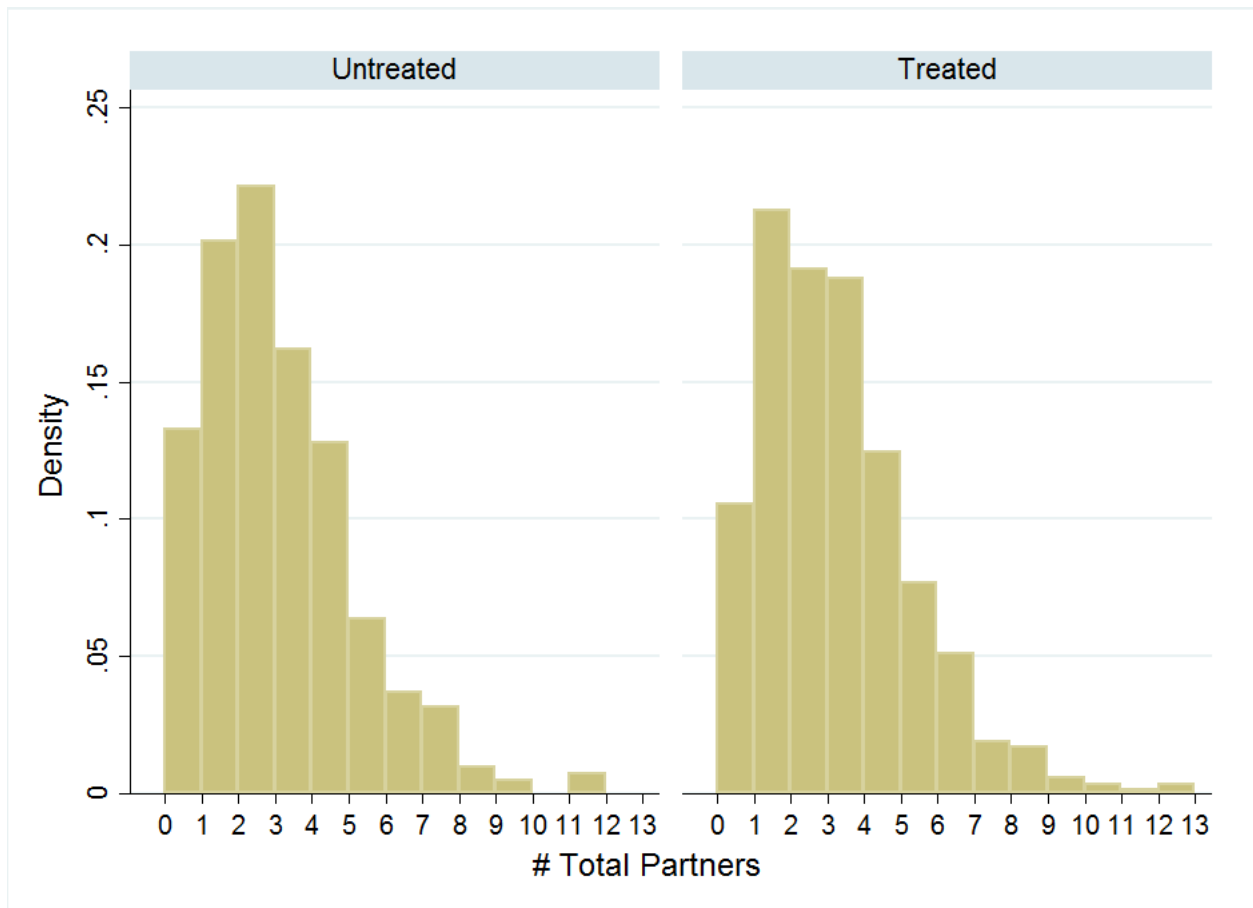
Figure 1.7: Distribution of Total Number of Partners

More importantly, I test the validity of the randomization of treated partners. I begin by checking whether the treated students select partners differently from the control students. If the treatment does cause the network formation process to differ between the students, then there should be differences in how many partners one has or partners' characteristics. Panel D in the balance table rules this out. On average, a student has about three partners in total and two of them are treated. The realized likelihood of partners being assigned as treated (1.9/2.6 = 73% for the control and 2.1/2.8 = 75% for the treated) is close to the treatment assignment rate at 75%. Total partners, the number of treated partners, partners' Exam 1 score, and gender are also balanced. Figure 1.7 plots out the distribution of the number of total partners a student has by treatment. A Kolmogorov-Smirnov test cannot reject that the distributions are the same.

I also regress each baseline characteristics of interest on the number of treated partners and indicators of total friends, in order to test for the central exogeneity assumption that conditional on the total number of partners, the number of treated partners is randomly assigned.

Coefficients on the number of treated partners from these regressions are shown in the last column. In most cases, these coefficients are small and statistically insignificant, suggesting the randomization is successful.

These checks warrants the empirical strategy to estimate peer effects based on the elicited network after Exam 2.

1.4 Results

1.4.1 Empirical Strategy

Recall that the intervention is randomized at the individual level. Although students choose whom to study with, the number of partners getting the intervention is random conditional on the total number of partners. By exploring the exogenous variation in the number of partners treated, I can identify both the treatment effect and the spillover effects from treated partners, on outcomes of interest (Y). In particular, I run the following regression:

$$Y_i = \alpha + \beta^* T_i + \gamma^* \text{FriTreated}_i + \rho^* \{T_i^* \text{FriTreated}_i\} + \sum_{k=1} \theta_k^* 1(\text{Total}_i = k) + \sigma X_i + \epsilon_i \quad (1.1)$$

where "T_i" is the treatment indicator. "FriTreated_i" is the number of study partners receiving the treatment information. "Total_i" is the total number of study partners. I use a dummy indicator for the number of total partners to allow for more flexibility instead of imposing a functional form. X_i includes individual control variables such as gender, lecture section, cohort, Exam 1 score, and high school and official cumulative college GPA. I include an interaction term between "FriTreated_i" and the student's own treatment indicator. This allows me to capture differential effects of peer exposure depending on own treatment status. β is the direct treatment effect. Both γ and ρ are my key estimates of spillover effects. γ is the effect of having an additional treated partner when i self is *untreated*. $\gamma + \rho$ is the effect of having an additional treated partner when i self is *treated*. [26]

Now I discuss how network measurement error can potentially bias my results from the

---

[26]The estimation strategy follows Miguel and Kremer (2004) and Oster and Thornton (2012).

above regression. If students randomly forget to report some partners, then both γ and ρ are overestimated. This is because for each additional treated partner I observe, the true number of additionally treated partners is larger. In other words, under-reporting scales the number of treated partners down. On the flip side, if students report too many partner whom they do not interact with, then both γ and ρ are underestimated. By using the "or" network construction, I am more likely to have over-reporting than under-reporting and hence underestimate the spillover effects.[27] If misreporting is not random, for example, if students are more likely to report treated than untreated partners, then both γ and ρ are overestimated if an unobservable that positively correlates with the number of total friends is also positively correlated with the outcomes.

One way to circumvent potential measurement errors is to use the fraction, instead of the count, of partners treated. Because of random assignment, the fraction of partners treated is exogenous, even unconditional on the total number of partners. The data confirm that the number of treated friends indeed is proportional to the number of total friends. For a more accessible interpretation, I choose to present results from using the count in the results section. I include the results from the alternative specification with fraction treated as a robustness check and the main results are qualitatively unchanged (See Appendix Table A.5).[28]

1.4.2 Applet Usage as an Outcome

The primary objective measure of study behaviors is applet usage: whether a student takes up the applet and how many questions a student answers. Figure 1.8 provides descriptive evidence of spillover effects on applet usage among the untreated. For exposition purposes, I only show the results for untreated students with fewer than five partners (which accounts for about 70% of the total sample). The y-axis corresponds to the fraction of students ever using the applet in Figure 1.8a and corresponds to the number of questions answered through the applet in Figure 1.8b. The x-axis in both figures corresponds to the number of treated partners one has, conditional on the total number of partners a student has. Both figures show an upward slope

---

[27]Oster and Thornton (2012) find that peer influence is smaller under the "or" network compared to the "and" network.

[28]This alternative estimation strategy has been used in Cai et al. (2015). Oster and Thornton (2012) find qualitatively the same results using both estimation strategies.

in applet usage with more treated partners.

To quantify the spillover effect, Table 1.3 shows the estimated effect sizes. In Panel A, the binary outcome variable is whether a student has ever used the applet for exam study. I stack the regressions for Exam 2 and 3 together and add in exam fixed effects in addition to the independent variables described before in section 1.4.1. On average, 55.5% of the untreated students with no treated partner use the applet. Column (1) uses a linear probability model. Column (2) is a logit regression and the estimates are average marginal effects. The estimates in column (3) are the average marginal effects on answering at least one question, estimated by a Tobit model (used in Panel B). Those three specifications yield estimates of similar magnitudes. The direct treatment effect, β, is about 0.073 and is statistically significant at the 10% level. Thus the treatment increases a student's likelihood of using the applet by about 7.3 percentage points, or 15% relative to the control mean. The estimated γ is around 0.02. This means that having an additional treated partner makes an untreated student 2 percentage point more likely to use the applet. Although γ is not statistically significantly different from 0, its magnitude is worth noting. The estimated peer effect on an untreated student is about 30% of the direct treatment effect ($= \frac{0.020}{0.072}$). The interaction term, ρ, has a negative sign, and is statistically significantly different from 0. This means that the spillover effects on a treated student are much smaller than that on an untreated student. The effect of having a treated partner on a treated student is $\gamma + \rho$. To test for whether treated students benefit from being with an additional treated partners, I report the p-value for the test under the null hypothesis that $\gamma + \rho$ equals to zero for each column. The p-values are large and the point estimates of $\gamma + \rho$ are negatively signed, which means that the treated students do not increase their take-up of the applet significantly from being with other treated students around them.

In Panel B, I look at the intensive margin — how many questions students answer using the applet. Column (1) is an OLS regression. The estimates reported in column (2) are the average marginal effects on the censored outcome from a Tobit model. The last two columns use count models: a Poisson regression and a negative binomial regression (in case of a potentially over-dispersed range of the number questions answered). The magnitudes of the estimates are similar across different functional forms in the four columns. Receiving the treatment makes a
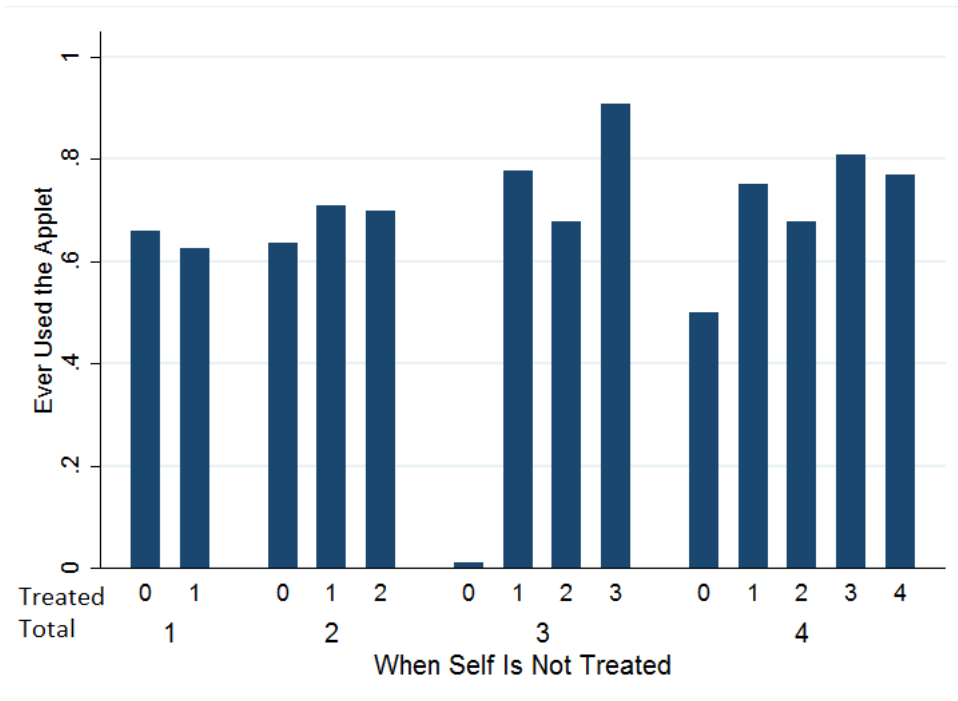
(a) Likelihood of Ever Using the Applet
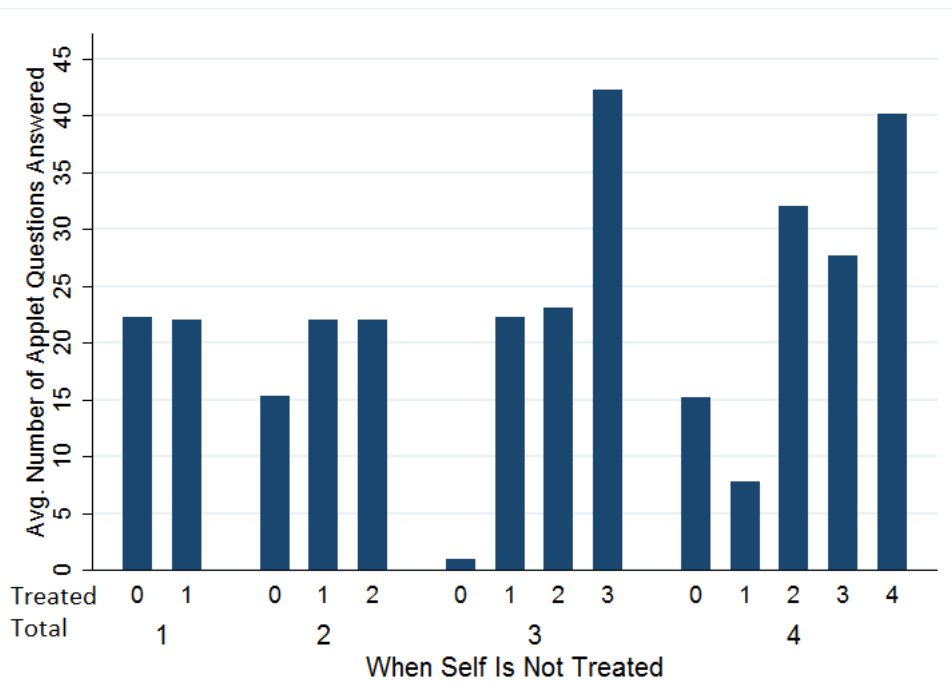
(b) Number of Questions Answered

Figure 1.8: Descriptive Evidence of Spillover Effects on Applet Usage of Untreated

Notes: The y-axis corresponds to the fraction of students ever using the applet in figure (a) and corresponds to the number of questions answered through the applet in figure (b). The x-axis in both figures corresponds to the number of treated partners one has. I plot out graphs by the total number of partners one has.

student answer 6 more questions, or 35% more questions relative to the control mean of 16.83. An additional treated partner makes an untreated student answer two more questions on average. The estimated peer effect is also about 30% of the direct treatment effect ($= \frac{2.320}{5.990}$), similar to what I find on the extensive margin. The interaction term is statistically significantly different from 0 as well. A Wald test shows that on the intensive margin, the treated students also do not increase their intensity of use from studying with other treated students (p>0.7).

Table 1.3: Treatment and Spillover Effects on Applet Use

Panel A: Ever Used	(1) LPM	(2) Logit	(3) Tobit	
T_i (β)	0.072*	0.072**	0.073**	
	(0.038)	(0.036)	(0.029)	
FriTreated$_i$ (γ)	0.020	0.022	0.025*	
	(0.014)	(0.015)	(0.013)	
T_i*FriTreated$_i$ (ρ)	-0.028**	-0.030**	-0.029***	
	(0.011)	(0.012)	(0.010)	
Control Mean	0.555	0.555	0.555	
Control S.D.	(0.498)	(0.498)	(0.498)	
p-value: $\gamma + \rho = 0$	0.506	0.561	0.720	
Obs.	3054	3054	3054	
Panel B: #Questions Answered	(1) OLS	(2) Tobit	(3) Poisson	(4) Negative Binomial
T_i (β)	5.990**	5.983**	6.779**	6.792**
	(2.358)	(2.445)	(2.739)	(2.902)
FriTreated$_i$ (γ)	2.320	2.043*	2.375*	2.026*
	(1.402)	(1.073)	(1.261)	(1.082)
T_i*FriTreated$_i$ (ρ)	-2.486**	-2.352***	-2.610***	-2.272**
	(1.042)	(0.868)	(1.007)	(0.997)
Control Mean	16.83	16.83	16.83	16.83
Control S.D.	(23.17)	(23.17)	(23.17)	(23.17)
p-value: $\gamma + \rho = 0$	0.874	0.720	0.809	0.776
Obs.	3054	3054	3054	3054

Notes: All the columns stack results from Exam 2 and 3 together. "Ever Used" in Panel A equals to 1 if a student has used the applet for at least once for exam studying, and 0 otherwise. "#Questions" in Panel B counts the number of questions a student answers through the applet. The estimates are the mean marginal effects. Standard errors are clustered at the lab section level (68 clusters). The network is constructed based on the network elicited after Exam 2, using the "or" definition. The individual controls included, but not shown here, are gender, official college cumulative and high school GPA, cohort and lecture section dummies, and Exam 1 test score. Control mean is the mean outcome for an untreated student with no treated study buddies. The p-value shown is under the null hypothesis that there are no peer effects on the treated student.

The above results show that the intervention of asking students to make study plans has positive impacts on applet usage. It is interesting to see whether the effectiveness of the nudge varies with students' baseline test scores (Exam 1 score). On one hand, students with higher academic ability measured by their Exam 1 score may respond less to the intervention since they might already have good study habits. On the other hand, students with higher test scores may respond more to the intervention since they might in general pay more attention to the intervention messages. In Table 1.4 column (1) to (4), I break the sample by students' baseline Exam 1 test score into four quartiles. In general the direct treatment effect is larger for students with a higher than median Exam 1 score. Focusing on the take-up in Panel A, I find that the direct treatment effect, β, is smaller on students with below median Exam 1 scores compared to other students. Moreover, students with lower Exam 1 scores are also less likely to be affected by treated peers as reflected by the magnitudes of the γ estimates. In Panel B, the treatment effect is in general larger for students with higher Exam 1 scores. Students in the bottom quartile only answer on average about 1.5 more questions. The heterogeneous impacts by initial exam scores suggest an interesting interaction between test performance and the intervention. Notice that both in Panel A and in Panel B, the control means of students with lower Exam 1 scores are much higher than those of students with higher scores. This might be another reason for a larger intervention impact on the above median students. In general, from Figure 1.4 and Table 1.4, the spillover effects on untreated students are consistently positive, while social interactions do not appear to reinforce the treatment effect among the already treated students.

1.4.3 Spillovers and Tie Strength

So far, I have estimated an average spillover effect from additional treated partner without distinguishing tie strength. A large volume of literature in social psychology and social networks has shown that stronger ties are more likely to exert influence on behavior (Aral and Walker, 2014; Bakshy et al., 2012; Bond et al., 2012). The network data allow me to proximate tie strength. I separately define strong ties as: Partners who study together for exams are more likely to be strong ties than partners who only talk about the course materials; Pairs of students who mutually list each other as study partners are more likely to be strong ties than pairs of

26

Table 1.4: Spillover Effects by Exam 1 Test Scores

Panel A: OLS Ever Used	(1) Bottom Quartile	(2) 2nd Quartile	(3) 3rd Quartile	(4) Top Quartile
T$_i$ (β)	-0.020	-0.001	0.192***	0.108*
	(0.068)	(0.073)	(0.056)	(0.055)
FriTreated$_i$ (γ)	-0.010	0.005	0.056**	0.012
	(0.028)	(0.025)	(0.028)	(0.032)
T$_i$*FriTreated$_i$ (ρ)	0.007	-0.033	-0.049**	-0.014
	(0.024)	(0.023)	(0.021)	(0.023)
Control Mean	0.625	0.630	0.423	0.536
Control S.D.	(0.489)	(0.488)	(0.499)	(0.508)
p-value: $\gamma + \rho = 0$	0.931	0.156	0.742	0.933
Obs.	772	850	826	606

Panel B: Poisson # Questions	(1) Bottom Quartile	(2) 2nd Quartile	(3) 3rd Quartile	(4) Top Quartile
T$_i$ (β)	1.446	6.099	12.830**	7.755*
	(4.765)	(5.056)	(6.159)	(4.782)
FriTreated$_i$ (γ)	2.207	0.972	4.168	1.358
	(2.462)	(2.112)	(2.983)	(2.673)
T$_i$*FriTreated$_i$ (ρ)	0.289	-3.268	-3.560	-2.759
	(1.988)	(2.042)	(2.275)	(1.777)
Control Mean	22.84	19.04	10.19	13.50
Control S.D.	(26.57)	(21.45)	(15.83)	(27.23)
p-value: $\gamma + \rho = 0$	0.239	0.160	0.802	0.419
Obs.	772	850	826	606

Notes: All the columns stack results from Exam 2 and 3 together. "Ever Used" in Panel A equals to 1 if a student has used the applet for at least once for exam studying, and 0 otherwise. "#Questions" in Panel B counts the number of questions a student answers through the applet, and the estimates are the mean marginal effects. Standard errors are clustered at the lab section level (68 clusters). The network is constructed based on the network elicited after Exam 2, using the "or" definition. The individual controls included but not shown here are the same as before. Control mean in each column is the mean outcome for an untreated student with no treated study buddies and with an Exam 1 test score corresponding to each quartile. The p-value shown is under the null hypothesis that there are no peer effects on the treated student.

students where one student unilaterally lists the anothe; Partners who have fewer other study partners to study with are also more likely to be strong ties since the fewer partners one has, the more time one can afford to study with each partner.

To separately estimate spillover effects from strong versus weak ties, I split the number of treated peers and run the following regression:

$$\begin{aligned}
Y_i = {} & \alpha + \beta^* T_i + \gamma_1^* \text{FriTreated}_i^{strong=1} + \gamma_0^* \text{FriTreated}_i^{strong=0} \\
& + \rho_1^* T_i^* \text{FriTreated}_i^{strong=1} + \rho_0^* T_i^* \text{FriTreated}_i^{strong=0} \quad\quad (1.2) \\
& + \theta_k^* 1(\text{Total}_i = k) + \sigma X_i + \epsilon_i
\end{aligned}$$

"$\text{FriTreated}_i^{strong=1}$" counts the number of treated partners who are strong ties. The count of remaining partners is given by "$\text{FriTreated}_i^{strong=0}$". I also add two interaction terms between these two counts and the treatment dummy.

Results in Table 1.5 show that strong ties are associated with stronger spillovers. Column (1) separates exam study partners from general course study partners. Column (2) splits the treated partners by whether their degree is above the median degree. Column (3) distinguishes treated partners who mutually list each other as a study partner from other treated partners. Consistent with the hypotheses, the point estimates of γ_1 is larger than that of γ_0. Especially on the extensive margin in Panel A, the null hypotheses that $\gamma_0 = \gamma_1$ can be rejected at the 10% level.[29]

The above finding that γ_1 is larger than γ_0 is suggestive of a social learning mechanism. Strong ties might be more likely to share information about study strategies such as the benefit of making study plans or using the applet. Without detailed data on conversations among study partners or how they study together, it is difficult to pin down the exact mechanism. In order to test this mechanism, I will next propose a theoretical model for this social learning process and show evidence to support this mechanism.

---

[29]Oster and Thornton (2012) find statistically larger peer effects from strong ties than from weak ties. Cai et al. (2015) find qualitatively the same result but do not test the significance of the difference.

Table 1.5: Spillovers Depend on Tie Strength

Panel A: OLS	Ever Used			
	(1)	(2)	(3)	(4)
strong =1 if	Exam Partner	Low Degree	Mutually List	Plan to Use Applet
T (β)	0.072*	0.086**	0.074*	0.074*
	(0.038)	(0.038)	(0.038)	(0.038)
FriTreated$^{c=1}$ (γ_1)	0.040**	0.063**	0.032*	0.049***
	(0.017)	(0.024)	(0.017)	(0.018)
FriTreated$^{c=0}$ (γ_0)	0.005	0.005	0.013	-0.008
	(0.017)	(0.016)	(0.017)	(0.018)
Fri w.o. a Plan				0.071
				(0.045)
H$_0$: $\gamma_1 = \gamma_0$				
p-value:	0.085	0.041	0.354	0.017
Obs.	3054	3054	3054	3054

Panel B: Poisson	#Questions Answered			
	(1)	(2)	(3)	(4)
strong =1 if	Exam Partner	Low Degree	Mutually List	Plan to Use Applet
T (β)	6.757**	7.695***	6.902**	7.034**
	(2.750)	(2.845)	(2.758)	(2.776)
FriTreated$^{c=1}$ (γ_1)	3.704*	5.103**	3.940*	4.738***
	(1.939)	(2.160)	(2.023)	(1.807)
FriTreated$^{c=0}$ (γ_0)	1.370	1.351	1.474	0.349
	(1.454)	(1.451)	(1.269)	(1.644)
Fri w.o. a Plan				3.632
				(3.565)
H$_0$: $\gamma_1 = \gamma_0$				
p-value:	0.310	0.140	0.231	0.038
Obs.	3054	3054	3054	3054

Notes: All the columns stack results from Exam 2 and 3 together. "Ever Used" in Panel A equals to 1 if a student has used the applet for at least once for exam studying, and 0 otherwise. "#Questions" in Panel B counts the number of questions a student answers through the applet, and the estimates are the mean marginal effects. Standard errors are clustered at the lab section level (68 clusters). The network is constructed based on the network elicited after Exam 2, using the "or" definition. The individual controls included but not shown here are the same as before. The p-value shown is under the null hypothesis that treated partners who are more likely to share plans have the same magnitude of influence as those who less likely to share plans.

1.5 A Simple Social Learning Model

In this model, I assume students choose applet usage to maximize the expected utility from using the applet, which can increase their grades.[30] To form expectations, students can learn from their partners about the return of the applet usage. The key assumption is that the intervention reduces uncertainty in the return on applet usage. I allow for the strength of social learning to differ based on the tie strength. I first derive two model predictions that map to the main results from Section 4. I then test an additional model prediction and show supporting evidence.

1.5.1 A Model for Applet Usage

A student i chooses, e_i, how effort she should spend on the applet. The resulting utility is a function of the grade, G_i and the cost of time spent, ke_i^2. I assume a quadratic cost function where $k > 0$ for increasing marginal costs.

$$U_i(e_i) = f(G_i(e_i), e_i) = G_i(e_i) - mG_i^2(e_i) - ke_i^2 \tag{1.3}$$

$$G_i(e_i) = G_{0i} + e_i v \tag{1.4}$$

If m is positive (negative), then the marginal return of a higher grade is increasing (decreasing). I only assume that $m < \frac{1}{2G_{0i}}$. G_{0i} is the grade that i gets without using the applet. The return from using the applet is v, unknown to students. For exposition purposes, I assume v to be the same for all the students. When making study decisions, students face uncertainty in v. Such uncertainty can arise from not knowing how helpful the applet is in general for everyone. It can also rise from not knowing how helpful the applet is in particular for oneself. Students thus have a belief over v. I use b_i to denote the belief over v.

$$b_i \sim N(\mu_i, \sigma_i^2)$$

---

[30]I can arrive at the same comparative statics under a social learning model used by development economists (e.g. Foster and Rosenzweig (1995) and Bandiera and Rasul (2006)). See Appendix A for this alternative model.

The belief of v, b_i, is an independently identically distributed normal random variable with a mean of μ_i and a variance of σ_i^2. The key assumption in my model is that the intervention reduces σ_i^2. This is because the prompts can induce students to plan ahead, to preemptively think about the effectiveness of different study activities and to have a commitment device to avoid distractions.[31] Following this argument, I assume that the returns to applet usage are more predictable for treated students: $\sigma^2(T_i = 1) < \sigma^2(T_i = 0)$.[32] I follow the literature on social learning (Bandiera and Rasul, 2006; Foster and Rosenzweig, 1995) and assume that σ^2 is known for both T=0 and T=1.

Students choose e_i to maximize the utility specified in (3). Plugging (4) into (3) and taking the first order condition in the Appendix, I derive the optimal level of applet use:

$$e_i^* = \frac{\mu_i - 2mG_{0i}\mu_i}{2\big(m(\mu_i + \sigma_i^2) + k\big)} \tag{1.5}$$

It is straightforward to see that $\frac{\partial e_i^*}{\partial \sigma_i^2} < 0$. Under the condition that m is not too big, i.e. $m < \frac{1}{2G_{0i}}$, I show that $\frac{\partial e_i^*}{\partial \mu_i} > 0$ in the Appendix. In other words, the optimal applet usage is increasing with a more precise belief (i.e. a smaller σ_i^2) and a higher mean belief about the return (i.e. a larger μ_i). It is also straightforward to see that under the assumption that treated students have a smaller σ^2, treated students use the applet more. This maps to the positive treatment effect estimate, β.[33]

When students learn about v from partners, social learning will reduce the variance in b_i and affect applet usage choice. To see that I now integrate social learning into the choice model.

---

[31]For example, students who do not plan ahead might be crunched for time. As a consequence, they might experience a higher level of mental anxiety or more difficulties to smooth out the preparation process. Thus they might face a higher level of uncertainty about the return to applet usage.

[32]It is possible that the treated students increases μ_i, especially those students who choose to use the applet. I will include this extension later in the model and show that treated partners who plan to use the applet can influence a student's propensity to use the applet differently from those who do not mention to use the applet in the plan. Allowing for treated students to have higher μ_i does not change comparative statics and the main model predictions.

[33]Another comparative static is that $\frac{\partial e_i^*}{\partial G_i} < 0$. This may explain why students with higher initial Exam 1 scores use the applet less in subsequent exam preparation.

1.5.2   Social Learning

With social learning, students learn about partners' beliefs about the return of the applet from talking or observing partners' applet usage. Each new signal from a partner j, b_j, is distributed as $N(\mu_j, \sigma_j^2)$. I allow the strength of the information, $a_{ij} \in (0, 1)$, to differ based on the tie strength. A higher a_{ij} means a higher likelihood that i learns about j's belief. Students update their beliefs according to Bayes' rule and the updated belief has the following mean and variance:

$$\mu_i' = \frac{\frac{\mu_i}{\sigma_i^2} + \sum_j a_{ij} \frac{\mu_j}{\sigma^2(T_j)}}{\frac{1}{\sigma_i^2} + \sum_j a_{ij} \frac{1}{\sigma^2(T_j)}} \tag{1.6}$$

$$\sigma_i' = \frac{1}{\frac{1}{\sigma_i^2} + \sum_j a_{ij} \frac{1}{\sigma^2(T_j)}} \tag{1.7}$$

Two main predictions thus follow and can explain the results in section 1.4. First, if a student learns from every study partner, then the model predicts that information from a treated partner decreases the variance in beliefs more than the information from an untreated partner because $\sigma^2(T_j = 1) < \sigma^2(T_j = 0)$. This predicts a positive spillover effect estimate, ρ, on the applet usage of the untreated students.[34]

Second, the effect of learning from more treated partners is smaller for students with a more precise belief on the return. This predicts that the effect of having a treated partner is smaller on a treated student than on an untreated student. In my analysis in section 1.4, this is captured as a negative estimate, γ, on the interaction term.

In addition, if I allow treated partners to differ in their mean beliefs, μ_j, then a treated partner who thinks that the return is high makes a student more likely to use the applet. This is because a higher μ_j leads to a higher μ_i'. This is an additional prediction that can be tested by the data. Next I will show evidence that further supports this prediction.

---

[34]This also predicts a positive correlation between the number of partners one has and the applet take up. This is what I observed in 1.5. See Figure AA.3 in the Appendix.

1.5.3 Testing the Model

Although I do not elicit beliefs directly, treated students' study plans provide a way to approximate beliefs. Students who explicitly mention to use the applet in their plans are more likely to have higher beliefs than those who do not mention the applet. Therefore, following equation (2), I split the number of treated partners a student i has by whether the partners (denoted by j) mention to use the applet in the exam plan (counted by $\text{FriTreated}^{a=1}$):

$$
\begin{aligned}
Y_i = \alpha + \beta^* T_i &+ \gamma_1{}^* \text{FriTreated}_i^{a=1} + \gamma_0{}^* \text{FriTreated}_i^{a=0} \\
&+ \rho_1^* T_i^* \text{FriTreated}_i^{a=1} + \rho_0^* T_i^* \text{FriTreated}_i^{a=0} \\
&+ \theta_k^* 1(\text{Total}_i = k) + \sigma X_i + \epsilon_i
\end{aligned} \tag{1.8}
$$

The model predicts that γ_1 and $\gamma_1 + \rho_1$ should be positive and γ_1 and $\gamma_0 + \rho_0$ should be negative. Table 1.6 column (1) and (2) show estimates separately for γ and $\gamma + \rho$. Panel A column (1) shows that a treated partner with a willingness to use the applet increases a student's likelihood to use the applet by 4.9 percentage points and the estimate is strongly statistically significant. In contrast, the spillover estimate from a treated partner without a plan to use the applet is statistically insignificant and negative (-0.007, s.e. = 0.018). I notice the same pattern on the intensive margin. Panel B column (1) shows that a treated partner with a willingness to use the applet makes a student answer 4.616 more questions. The finding that the positive spillover effects on untreated students mostly come from treated partners who also planned to use the applet implies that exposure to treated partners is not a sufficient condition for spillovers to happen. Rather exposure to the "right" individuals – those who also have a willingness to use the applet is a key for leveraging social networks to scale up policy effects.

Interestingly column (2) in Panel A and B shows that a negative influence from partners who did not plan to use the applet. The estimates become statistically significant in column (3) when I restrict to the sample of treated students with a willingness to use the applet. These students are 5.3 percentage points less likely to use the applet and answer 7 fewer questions with an additional treated partner who did not plan to use the applet. Recall from Table 1.3, the direct treatment effect on the extensive and intensive margin is about 0.073 and 6 respectively.

Table 1.6: Evidence for Social Learning

Panel A: OLS	Ever Used			
	(1)	(2)	(3)	(4)
Partner pair ij, When i is	untreated	treated	treated and planned to use	treated but did not plan to use
j Planned to Use Applet	0.049***	0.005	-0.003	-0.011
	(0.018)	(0.015)	(0.021)	(0.019)
j Did Not Plan to Use Applet	-0.007	-0.015	-0.053***	-0.006
	(0.018)	(0.014)	(0.018)	(0.022)
j Did Not Make a Plan	0.071	-0.038*	-0.028	-0.058
	(0.045)	(0.021)	(0.033)	(0.037)
Controls	Y	Y	Y	Y
Obs.	3054	3054	882	1118

Panel B: Poisson	#Questions Answered			
	(1)	(2)	(3)	(4)
Partner pair ij, When i is	untreated	treated	treated and planned to use	treated but did not plan to use
Planned to Use Applet	4.616**	1.504	0.138	1.040
	(2.770)	(1.180)	(1.966)	(4.278)
Did Not Plan to Use Applet	0.295	-2.000*	-7.014***	-0.043
	(1.658)	(1.024)	(2.157)	(2.221)
Did Not Make a Plan	3.607	1.075	-1.823	2.450
	(3.549)	(2.510)	(4.551)	(100.82)
Controls	Y	Y	Y	Y
Obs.	3054	3054	882	1118

Notes: The results stack Exam 2 and 3 together. "Ever Used" is a binary dependent variable. "#Questions Answered" equals to the number of applet questions a student answers. All standard errors are clustered at the lab section level shown in the brackets (67 clusters). The network is constructed based on the network elicited after Exam 2, using the "or" definition. The individual controls included but not shown here are the same as before.

The large negative influence from the low belief treated partners is large enough to offset the direct treatment effect. This highlights circumstances under which peer interactions mitigate the treatment effect and alerts policy makers to be careful about treatment assignment. Even in a setting where everyone can be treated, it might not be optimal to treat everyone.

The results in Table 1.6 provide in general consistent support for the social learning model where high belief partners encourage and low belief partners discourage the applet usage.

1.5.4 Other Predictions for Future Testing

Additionally, the model generates predictions ripe for testing in future work. For example, the model predicts that while learning from partners increases the belief precision, the posterior

mean can decrease if the ratio of the signal mean to the signal variance is small. In other words, a student may lower her mean belief when her partner's signal is too noisy even when the partner has a high usage. For example, if the partner changes study plan frequently, then the partner's high use of the applet can be a mere chance. Another example of such scenario is in financial markets where information that signals a high return associates with a high volatility as well. Future lab or field experiments can help understand the relative effect of signal mean and variance.

1.6 Discussion

1.6.1 Other Potential Mechanisms

There are several other channels through which peer exposure to the intervention could also affect a student's applet usage. For example, students might be learning about the intervention message advocating for efficient exam preparation and it is the spread of the intervention message that leads to the spillover effects on behavior. However, I would argue that this mechanism cannot explain why social interactions can mitigate treatment among a subset of the treated students.

Another potential mechanism is the spread of the applet take-up happens when there are economies of scale or complementarity in using the applet. This channel predicts that a high usage of partners also causes a student to use more; Manksi (1993) refers to this as the endogenous peer effect. To test this, in the Appendix, I use a linear in the mean peer effects model to causally estimate the endogenous peer effect using the instrumental variable approach proposed by Bramoullé et al. (2009). While limited by the weak instruments to estimate the endogenous peer effect credibly, I argue that this alternative explanation is unlikely to explain the differential spillover effects from different ties I find in the data.

1.6.2 Scores as Outcomes

Recall that the treatment makes a student use the applet. it is natural to ask whether this increase in exam preparation leads to better performance outcomes. On the other hand, since scores are

noisy measures and there are other study efforts, performance might not change significantly.[35]

To estimate the effects of treatment and peer exposure on scores as outcomes, I run regressions based on Equation (1) using individual preparatory lab assignment scores, exam scores, and course grades as outcomes.

I find that the treatment affects how well students prepare for lab sections. Before each lab section, students complete a pre-lab assignment in order to prepare for the upcoming lab assignments. In Table 1.3 column (3), I use the accumulated pre-lab assignment scores since the Pre-Exam 2 message. The treatment increases the score by 0.649. This is equivalent to a 4% increase or a 0.15 standard deviation increase relative to the control mean. Having an additional treated partner also increases a control student's score by 0.172. The relative peer effect is about 27%, similar to the estimates from the applet usage data. The negative and statistically significant estimate for the interaction term is consistent with the previous conclusions from Table 1.3 column (2) and (3) that peer exposure does not have a significant effect on treated students.

Using exam scores and course grades outcomes, I find that the estimate of the treatment dummy in Table 1.7 is largely small and insignificant. Averaging over the two exams, a treated student without any treated partners scores 0.0009 standard deviations lower than a control student without any treated partners. For reference, the control mean is 0.047, with a standard deviation of 0.888. For aggregate course performance, I show results using different measures. The course has two methods to calculate the final course grade and the larger one becomes the final course GPA. Exposure to treated peers affects course performance very little, whether I use the two grade calculation methods separately in column (3) and (4) or the official course GPA in column (5).

These zero results on performance, despite the positive effects on efforts through the applet usage and lab section preparation, can be due to multiple reasons. One reason is that the increase in effort inputs does not map directly to gains in exam scores. The pre-lab assignments

---

[35]I find a correlation of 0.15 between study partners' scores. The magnitude is similar to that in Sacerdote (2001) who finds that roommates' first year college GPA has a correlation of 0.11. See Figure A.2 in the Appendix.

Table 1.7: Spillover Effects on Performance

OLS	(1)	(2)	(3)	(4)	(5)
				Course Grade	
	Pre-Lab	Exam Score	Method1	Method2	Final Grade
T_i (β)	0.649**	-0.009	0.218	-0.006	-0.009
	(0.323)	(0.054)	(0.700)	(0.724)	(0.052)
FriTreated$_i$ (γ)	0.172*	0.001	0.032	-0.374	-0.026
	(0.094)	(0.029)	(0.339)	(0.325)	(0.026)
T_i*FriTreated$_i$ (ρ)	-0.238**	0.034	0.252	0.278	0.020
	(0.098)	(0.021)	(0.252)	(0.253)	(0.020)
Control Mean	15.066	0.047	80.325	83.173	3.132
Control S.D.	(4.689)	(0.888)	(10.408)	(9.181)	(0.741)
p-value: $\gamma + \rho = 0$	0.380	0.126	0.314	0.715	0.737
Obs.	1527	3016	1527	1527	1527

Notes: Each column is an OLS regression. "Pre-lab" is the points (out of 19) received from the pre-lab assignments since Pre-E2 Message. Column 2 stacks Exam 2 and 3 together as an independent variable. Column 3 to 5 use the aggregate course grade calculated according to two formulas, as well as the final course grade (which should be the maximum of the two methods) as the outcome. All the standard errors clustered at the lab section level and are shown in the brackets (68 clusters). The network is constructed based on the network elicited after Exam 2, using the "or" definition. The individual controls included but not shown here are the same as before. Control mean is the mean outcome for an untreated student with no treated study buddies. The p-value shown is under the null hypothesis that there are no peer effects on the treated student.

are graded based on effort.[36] Students can start doing the assignments earlier, manage their time better, or put more effort into them. But for a test, more effort might not help. Table 1.4 shows that the treatment effects are larger on students with higher Exam 1 scores. For these students with higher initial performance, the marginal return from doing additional online practice questions might be low.

Another potential reason is that the increase in the applet usage crowds out the usage of other materials. Although I do not have behavioral data on how students utilize different materials, students list what they have used for exam preparation in the post-exam surveys. In Table 1.8 column (1), I examine whether the treatment affects how many resources a student reports to use for exam preparation. Treated students, on average, use slightly fewer materials than the control students ($\beta = -0.724$, $p < 0.001$). The magnitude is small given that a student generally uses 5 different materials for studying.[37]

Results in column (2) show that treated students are 13% less likely to claim to use the for-

---

[36]I say so because students can get credits if they list steps, even their answers are wrong.

[37]The five most popularly used materials, in a descending order, are lecture notes, formula cards, practice exam questions, the applet and past required homework questions. These choices are consistent with what students plan to use in the pre-exam messages.

mula flash cards compared to control students. Since students can use the formula cards during exams, one possible explanation is that treated students shift their time away from memorizing formulas, a seemingly less efficient activity.[38] In fact, although about half of the students wrote in their plan that they would use the formula cards for exam preparation, a third of them claimed that they did not actually use the formula cards. This might indicate that students find the formula cards less effective over time.

Column (3) uses students' self-reported usage of the applet. The estimated γ and ρ are similar to the results from Table 1.3. However β is much smaller. Hence, although the self-reported data suggests that students use fewer materials and might substitute away from using the formula cards, these findings are not conclusive. It would be useful to measure the effectiveness of the materials, to further support why students would substitute between the materials.

Table 1.8: Self-reported Exam Preparation

OLS	(1) #Materials Used	(2) Used Formula Card	(3) Used the Applet
T_i (β)	-0.724***	-0.126***	0.005
	(0.194)	(0.038)	(0.037)
FriTreated$_i$ (γ)	0.001	-0.001	0.023
	(0.081)	(0.015)	(0.016)
T_i^*FriTreated$_i$ (ρ)	-0.076	0.007	-0.022*
	(0.076)	(0.015)	(0.012)
Controls	Y	Y	Y
Control Mean	5.32	0.627	0.463
Control S.D.	2.616	0.485	0.500
p-value: $\gamma + \rho = 0$	0.305	0.617	0.892
Obs.	2969	2969	2969

Notes: All the standard errors clustered at the lab section level shown in the brackets (68 clusters). The network is constructed based on the network elicited after Exam 2, using the "or" definition. The individual controls included but not shown here are gender, college cumulative GPA until the semester starts, cohort dummies and high school GPA. Control mean is the mean outcome for an untreated student with no treated study buddies. The p-value shown is under the null hypothesis that there are no peer effects on the treated student.

1.7 Conclusion

This paper uses a field experiment to causally estimate spillover effects on study behaviors across self-selected peer groups in a large college statistics course. To overcome the challenge for defining relevant peers, I collect detailed information to measure a social network of study

---

[38]Treated and control students self-report to spend a similar amount of time on studying. Hence the treatment does not induce students to relax their time constraint.

partners. I then randomly assign an educational intervention, based on insights from behavioral economics, at the student level. Treated students receive advice on exam preparation and write down study plans for upcoming exams. To identify causal spillover effects, I rely on exogenous variations in students' exposure to the treatment via study partners.

I examine how the study partner network affects student adoption of an educational technology. To overcome the difficulty in obtaining detailed and objective study behavior, I use high-frequency applet usage data. Combining detailed network and usage data, I find that each additional treated study partner increases one's usage of the applet as well as how prepared one is for the discussion sections. The estimated magnitude of the spillovers on an untreated student is about 30% of the direct treatment effect.

To explore potential mechanisms, I present a simple social learning model where students update beliefs about the return of the applet usage and then choose how much to use the applet for better grades. Social learning increases belief precisions and expected utilities. My model allows the level of social learning to vary with tie strength. Additional prediction of the model is also corroborated by my empirical results. My findings imply that spillovers do not happen just because of exposure to other treated individuals. Instead, an untreated student uses the applet more when he/she is exposed to other treated individuals who also have a willingness to use the applet. This highlights that policy makers, when facing constraints on how many to treat, should target particular individuals to leverage social influence to scale up policy impacts. Second, I find circumstances under which peer interactions mitigate the treatment effect. This alerts policy makers to be careful about whom to treat even there is no constraints on how many can be treated. A future extension of this paper is to structurally model spillover effects in a social network for counter-factual policy experiments where I can change who receives the treatment.

This study is among the first few to provide experimental evidence of how peer effects generate spillover effects and induce behavioral changes in higher education, with a particular focus on self-selected peers. The ability to measure study behaviors with better precision adds to the few empirical studies exploring the mechanisms behind peer influence. From a policy

perspective, my study shows that policy makers can leverage social interactions to spread intervention impacts through social learning, but they need to be careful when choosing whom to target.

Chapter 2

Peer Effects in Randomly Assigned Study Groups

2.1 Introduction

It is widely believed that peer interactions can have profound impacts on student academic outcomes. Relying on random assignment into peer groups such as roommates and classmates, past literature leverages the exogenous variations in peers' characteristics to identify peer effects, circumventing biases due to selection into peer groups. Peer interactions in these groups can be academic, social or both. Nonetheless, studies have not yet agreed on whether peers' past academic performances affect students current academic performances.[1] Understanding how one can engineer peer groups can help teachers leverage peer effects to improve student academic outcomes. The key goal of this paper is to provide direct and clean evidence of peer effects on learning outcomes at the study group level, where student interactions are mostly academically related.

To do that, I exploit a unique dataset in a naturally occurring environment with random assignment of study groups. Students in my sample are college students from a large STEM (Science, Technology, Engineer and Mathematics) field introductory course. Students sign up

---

[1] Sacerdote (2001) and Zimmerman (2003) find positive peer effects on college GPA between freshman roommates. Having a roommate with high baseline performance (measured by e.g. SAT/ACT scores) increases a student's first year GPA relative to otherwise. Furthermore, Ding and Lehrer (2007) further claim that reducing the variation in baseline performance of one's cohort mates increases one's achievement outcomes. In contrast, other studies using either randomly assigned roommates or seating neighbors find evidence for peer effects from peers' baseline performance (Foster, 2006; Han and Li, 2009; Zimmerman, 2003). Two other papers using regression discontinuity in admission scores also find little evidence of having high achieving peers affecting college enrollment, graduation or quality (Abdulkadiroglu et al., 2014; Dobbie and Fryer, 2014).

for study groups and choose a preferred time slot. Within each time slot, there are several parallel groups and students are assigned to those groups by a computer algorithm that I later argue is plausibly random. Then I use exogenous variations in group members' background characteristics such as background academic performances and demographics to causally identify peer effects on the course grade.

I find little evidence that peers' background academic performances affect student course grade, using both linear and nonlinear specifications of peer performance measures. The estimated peer effects are in the expected directions but not statistically significant. This finding is consistent with the results from Foster (2006) and Lu and Anderson (2015) but in contrast with Sacerdote (2001) and Zimmerman (2003).[2]

Meanwhile, I provide evidence advocating for having more female students in the study groups. I find a gender mix effect. Female students get higher course grades when they are in a more female group, a finding consistent with Hoxby (2000) and Lu and Anderson (2015). Moreover, I find that male students also benefit from studying with more females, a finding in contrast with Hill (2015). Hill (2015) uses self-reported friendship data and a novel instrumental variable approach to conclude that the share of female school friends negatively affects male students' high school GPA. One key difference between his paper and mine is that I focus primarily on academic interactions whereas he looks at potentially both academic and social interactions. The contrasting findings suggest that peers may play a different role depending on the nature of the interactions, which raises the necessity for future research to separately study peer effects in academic versus social interactions.

A primary contribution of this paper is to shed light on possible mechanisms to explain the gender mix effect. To do so, I gather detailed online course website usage data. The log data show that being in a more female group makes one more likely to download exam solutions and lecture handouts. This might happen due to peer influence or norm pressure since female students tend to visit and download the materials more often than male students. Regressions show a positive correlation between the course grade and how students download these materials and thus a potential mechanism of peer effects through which the gender mix

---

[2]Foster (2006) also does not find such causal effects through friends' background performance measures.

42

affects the course grade is through changing particular study behavior.

As a secondary result, I provide evidence to advocate for mixing students based on their lecture sections. Peers in different sections are exposed to different instructors and peers. Therefore even when the course materials are standardized across the sections, study groups may still act as a place for students to exchange and share information. I show that being in a group with peers from different lecture sections increases one's final course grade. Interestingly, the relationship between how mixed a group is and the course grade is not linear. Having peers from another lecture section is better than having none, but the benefit does not linearly increase. Per the estimates, the optimal number of peers from a different section is around 9 for an average size group with 12 students.

This paper advances the previous literature on peer effects in the following way. First, previous literature suffers from justifying the definitions of peer groups. For example, Stinebrickner and Stinebrickner (2006) doubt whether one should use freshmen roommates, who are merely randomly assigned to live together, to look for evidence of peer effects on academic outcomes. This study defines peer groups in a collaborative learning environment that has not been studied before. Second, this paper focuses on peer effects in academic interactions while previous studies have mixed both academic and social interactions together. Lastly, earlier literature except Stinebrickner and Stinebrickner (2006) does not offer clear insights into how peer effects operate. They attribute peer effects to some unobservable peer quality that can manifest through past academic performances. I unpack a possible mechanism of peer effects which is that peers can change particular study behavior. With more readily available data measuring how students study (through online course platforms and learning technologies), researchers can acquire detailed behavioral data to better understand how peer effects operate. On a practical level, my paper provides practical suggestions for assigning students into study groups.[3]

---

[3]Peer groups or cooperative learning are popular teaching tools. In a national survey (Puma et al., 1993), 79% of elementary school teachers and 62% of middle school teachers reported that they asked students to form study groups and assist each other's learning process.

2.2 Study Group Administration

In this section, I go over the institutional background and the set up of the study groups. The Science Learning Center (SLC) at the University of Michigan organizes study groups for a variety of STEM courses each semester.[4] The study groups work in the following way. Each study group meets roughly once a week for two hours.[5] During these meetings, each group reviews course materials, solves problems, and discusses concepts together with a peer facilitator who has completed the course in previous semesters. The meetings are designed to engage all group members, and are not meant to be tutoring or review sessions led solely by the group facilitator. The facilitators' job includes administration (such as taking attendance and contacting absent group members) and supporting members to teach and learn from each other. The facilitators are screened for qualifications and all have standardized formal training before working as facilitators. They are paid for the time preparing for and being in the study group meetings. The facilitators are randomly assigned to groups based on their time schedule. Students do not know who the facilitators are until the first study group meeting.

Students sign up for a study group by choosing a preferred time slot via the SLC website. See Figure 2.1 for a user interface of the registration website. Within each slot, there are usually several seemingly identical groups. They only differ from each other by having different group numbers attached. For example, if a time slot has two groups, then one is labeled as "xxx-001" and the other is labeled as "xxx-002". These two group labels are almost identical except the group number. There is no information revealed to the students about who have signed up, who the peer facilitator is or where the study groups meet.[6] When a student clicks to join a group, the system assigns him/her to the smallest group of that time slot. In other words, a student can end up in a different group from his/her group of choice. The system processes students' signing up requests in batches. About every other two minutes, the system processes all the requests received in the past two minutes. This design adds noise to the information students

---

[4]These courses are Bio 171, 172 and Chem 130, 210, 215. SLC is the only official university office that holds study groups for these courses on campus.

[5]They usually meet once a week and a couple of extra times for exams. The facilitators sometime move the meeting time for special events such as religious holidays and fall break. A typical study group meets for 15-17 times a semester.

[6]Most of the groups meet on central campus.

Figure 2.1: Signing up for a Study Group Online

see, because a group that is open when a student signs up may no longer be open when the system processes his/her request. Students do not know the system's algorithm at all. Hence this serves as a natural experiment where students are randomly to study groups given their time slot choices. Another feature that makes the assignment random is that the study group time slots are filled up very quickly within a couple of hours after the registration process is open. Students have little time to coordinate with their friends' schedules and preferences in order to pick a specific group to join together.[7]

All the groups start with about 12 assigned students. If no groups in a student's desired time slot are open, he/she can choose to be on the wait-list for that time slot or choose another time slot. If the SLC staffs see excessive demand for particular time slots (such as a long waiting list or emails from students asking for additional groups), they increase the size limit for the study groups within those time slots. Students register for popular time slots therefore will have more peers in their groups. In my regressions, I control for time fixed effects.

Attendance is voluntary. Facilitators have the right to drop students from the study group

---

[7]In a typical system, the SLC has over thousands of students signing up for peer study groups but only six full time staffs. This system algorithm is the best randomization process given the administrative constraints. In later years, the SLC uses some non-random grouping mechanism such as having students in special education programs to be in the same group.

if they miss more than two meetings without excuses. Once a student is dropped, a slot opens up and wait-listed students have priority to join the group. Students who want to switch groups need to go through the SLC website to drop themselves from the current groups and sign up for another one. This rule is strictly enforced.[8] The facilitators cannot manually add in students. The SLC staffs do not handle students' email request to switch groups. Although switching groups may lead to sorting in study groups, I show in section 2.3.4 that this is not a concern.

2.3   Basic Data Description

2.3.1   Sample and Key Variables

My sample consists of undergraduate students taking Chemistry 210 (Chem 210, Introduction to Organic Chemistry) during the Fall 2013 semester. Chem 210 is the first course to introduce students to organic chemistry. It is one of the largest undergraduate STEM gatekeeper courses for the science majors in the College of Literature, Science and Art. Students from this course have always had a high demand for study groups. Each semester, on average more than half of the course takers participate in a Chem 210 study group organized by the SLC. This participation rate is much higher than that of other courses for which the SLC provides the study group program.[9]

In my analysis, I only include study group participants who attend at least two meetings over the whole semester. See Appendix B for the full distribution of attendance. Those who are assigned to groups but rarely show up are unlikely to cast a significant amount of peer effects on others or be influenced by peers. 61 students (about 7%) of the students who initially sign up for a study group successfully are dropped because of this.[10] Six students switched to another group after attending a couple of sessions in one group. For these six students, I treat the same student in different groups as two different students.[11]

The SLC provides detailed group and facilitator assignment, weekly attendance (including

---

[8]In only one special case, a student attends two different study groups throughout the semester.

[9]Those courses have a participation rate around 25%.

[10]I check to make sure that the student characteristics that I use in the analysis are not predictors for a student being dropped by me.

[11]The choice of dropping data and the choice of treating the six students appearing in different observations do not affect my main results.

excused absence) and study group weekly agenda recorded by the facilitators. The university registrar provides detailed information on all students enrolled in Chem 210 during the fall 2013 semester. The information includes gender, ethnicity, residency, high school GPA, SAT and ACT scores in each subject, class standing, credit units taken, academic level, cumulative GPA, special education program affiliation and Chem 210 course grade.[12] Two students from the SLC data set are not found in the registrar data set.[13]

For ACT and SAT scores, I convert them into percentile scores to measure past academic performance. I use the SAT percentile scores only when the ACT scores are missing.[14] I use the math test scores to approximate past math performance. To quantify past English performance I can either use the ACT English percentile score or the ACT reading percentile score. In most of the results shown below, using these two different English performance measures yields very similar results. For succinctness, I only present results using the ACT English percentile scores. When the ACT percentile scores are missing, I use the SAT verbal percentile scores. For the outcome measure, I use the course GPA by converting the letter grade using the University's conversion rubric.[15]

2.3.2 Descriptive Statistics by Study Group Participation

Since students decide whether to sign up for the study group program, I first compare the characteristics of students who are in the study groups to those who are not, in order to understand how different the study group participants are from those who do not participate.

Table 2.1 compares the background characteristics such as gender, being an in state student, past academic performances, course completion rate and grades by study group participation. Of all the 1437 students enrolled in Chem 210, 830 students participate in a study group at least three times a semester and I classify them as study group participants. Comparing the

---

[12]The special education program is called the Comprehensive Studies Program (CSP). According to the official CSP website description - "many CSP students are the first in their family to go to college, come from populations historically underrepresented at the University, or attended under-resourced high schools".

[13]It is possible that these two students are not officially registered for the class. SLC does not require students to be enrolled in courses to participate in the study groups.

[14]9 students have neither scores in the registrar data.

[15]Please refer to http://www.lsa.umich.edu/students/academicsrequirements/lsadegreesrequirements/ creditsandgradepointaverage.

Table 2.1: Individual Characteristics

	Study Group = 0		Study Group = 1		*(Study Group = 1)-*
					(Study Group = 0)
	Mean	N	Mean	N	*Mean*
	(Std.)		(Std.)		*(S.E.)*
Female	0.384	607	0.577	830	0.193***
	(0.486)		(0.494)		(0.02)
In State (MI)	0.702	607	0.666	830	-0.036
	(0.458)		(0.472)		(0.025)
High School GPA	3.860	551	3.866	786	0.007
	(0.166)		(0.146)		(0.009)
Math Performance	93.323	588	91.776	821	-1.547***
	(8.833)		(10.138)		(0.508)
Eng. Performance	90.850	588	89.801	821	-1.049*
	(10.907)		(11.291)		(0.291)
Withdraw	0.089	607	0.034	830	-0.055***
	(0.285)		(0.181)		(0.013)
Course Grade	2.716	553	2.908	802	0.191***
	(1.061)		(0.884)		(0.0549)

Notes: Math Performance equals one's ACT math percentile score (or SAT math percentile score if ACT scores are missing). Eng. Performance equals one's ACT English percentile score (or SAT verbal percentile scores if ACT scores are missing). The standard deviations of the variables and the standard errors for the t-tests are shown in brackets. The null hypothesis for the t-test is that the two sample means are not statistically different, assuming equal variances. * $< 10\%$, ** $< 5\%$, *** $< 1\%$.

participants and the non-participants, I see more female students (*diff.* $= 0.193$, $p < 0.01$) and students with slightly worse past math (*diff.* $= -1.547$, $p < 0.01$) and English performances (*diff.* $= -1.049$, $p < 0.1$) select into study groups.

The study group participants nonetheless are less likely to withdraw from the course (*diff.* $= -0.055$, $p < 0.01$) and end up getting higher final course grade (*diff.* $= 0.191$, $p < 0.01$) than the non-participants, despite the fact that students with slightly worse math or English preparations select into the study groups at the beginning. The distribution of the final course grades also differ between the participants and the non-participants. Figure 2.2 shows that the participants are less likely to have extremely low grades. This might suggest the effectiveness of the study groups since past academic preparations are expected to positively predict the course grade. Table B.1 presents various regression estimates for the relationship between study group participation and the course final grade. Column (3) and (4) fail to reject that the grade difference between study group participants and non participants is the same for males and females.
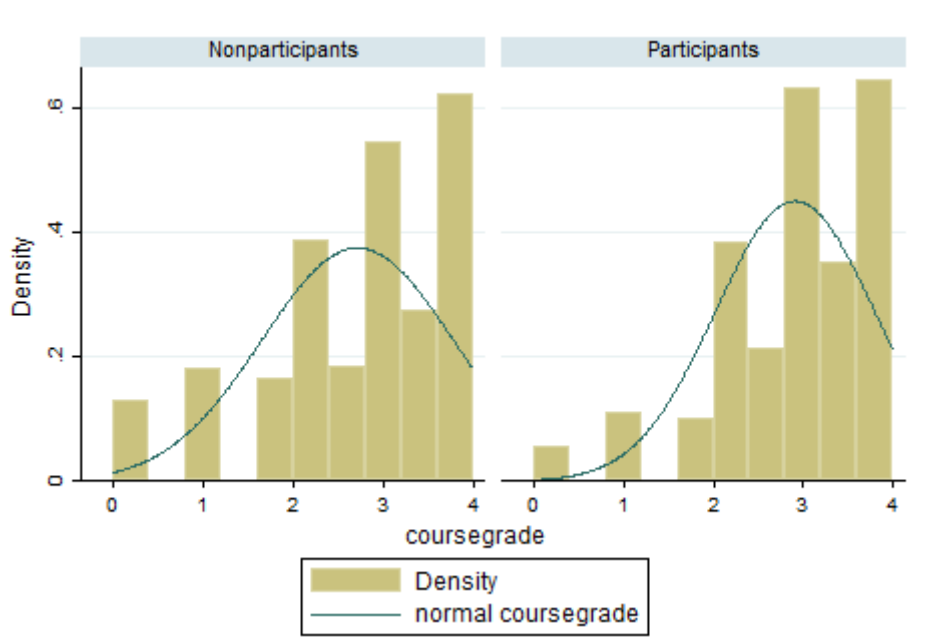
Figure 2.2: Course Grade (GPA) Distribution

It is noted that the sample that select into the study groups are quite different from the rest along several dimensions.

2.3.3 Peer Group Descriptive Statistics

The study groups on average have 12 students with a maximum size of 15 and a minimum size of 9. Table 2.2 presents the summary statistics of three key academic background characteristics for the peers in a total of 69 study groups.

Table 2.2: Descriptive Statistics

	Obs.	Mean	SD	Min	25th	Percentile 50th	75th	Max
Group Size	69	12	1.5	9	11	13	13	15
Avg. Peer High School GPA	791	3.87	0.05	3.72	3.84	3.87	3.90	3.98
Avg. Peer Math Performance	826	91.75	2.94	83.82	89.92	92.18	93.89	97.00
Avg. Peer Eng. Performance	826	86.80	3.83	79.45	87.87	90.15	91.83	97.18

Notes: The total number of study group participants is 836.

49

Table 2.3: Endogeneity Checks

	Own X =		
	High School GPA	Math Performance	Eng. Performance
Avg. X of	-0.020	0.001	-0.007
Own Group Peers	(0.083)	(0.043)	(0.072)
Avg. X of Peers	-38.86***	-43.07***	-40.26***
from Concurrent Groups	(4.765)	(3.354)	(3.855)
Study Group Time Slot Fixed Effect	Y	Y	Y
R^2	0.81	0.85	0.82
N	791	826	826

Notes: Each column is a separate regression. The dependent variable is one's own prior academic performance measure. "Avg. X of Group Peers" is the mean of the group peers (excluding oneself). "Avg. X of Peers from Concurrent Groups" is the mean of everyone from study groups that are in the same time slot. Standard errors in all regressions are clustered at the study group level, resulting in 69 clusters.

2.3.4 Endogeneity Checks

Before turning into examining peer effects, I first check for the exogeneity of the group assignment within each time slot. That is, I test whether there is a statistically significant relationship between one's own background characteristics and those of the group members.

I use a modified test first used by Guryan et al. (2009) to avoid the mechanical negative biases when one follows previous studies (e.g. Sacerdote (2001) and Foster (2006)) to check for exogeneity. For each background characteristic, I regress one's own on the average of that of the peers in one's own study group and the average of that of students from concurrent study groups.[16] If the study group assignment is not random, then the coefficients on peers' past academic performances should be statistically different from zero.

The results from the modified test are shown in Table 2.3. The standard errors are clustered at the study group level. The correlations between all of these three measures are small in magnitude and statistically insignificant ($\beta_{\text{High School GPA}} = -0.020, \beta_{\text{Math Performance}} = 0.001, \beta_{\text{Eng. Performance}} = -0.007$, p > 0.1). Hence, I cannot reject that students are randomly assigned into study groups within a time slot.

---

[16]If I were to use the tests in Sacerdote (2001) and Foster (2006), I would not include the average background characteristics of students from concurrent study groups as an independent variable. For a discussion on why the modified test is preferred to the typical test, please see Guryan et al. (2009).

2.4   Results

Since there is no evidence suggesting that the students are assigned nonrandomly I now examine how peers' background academic performances affect one's own course grade. In this section, I first examine peer effects using a typical linear in the mean model. I then run alternative regressions where I use different statistics of peer characteristics to test for other peer effect models in the literature.

2.4.1   Effects of Peers' Background Abilities

I first use a simple model where one's course grade is linear in one's own and peers' past academic performance measures. Specifically, for a student 'i' from study group G_i:

$$y_i = \alpha + \beta P_i + \gamma \overline{P_{-i,G_i}} + \eta X_i + \delta T_{G_i} + \epsilon_i \tag{2.1}$$

where y_i is i's course grade. P_i includes own past academic performance measures (math and English ACT/SAT test scores and high school GPA). $\overline{P_{-i,G_i}}$ includes group peers' average academic performances (excluding oneself). ϵ is the conditional mean-zero error term. T_{G_i} is the study group meeting time dummy. β measures the effects of own past academic performances on the course grade, while γ measures the effects of the average performances of the peer group. γ is the key peer effect estimate of interest. The ratio, $\frac{\gamma}{\beta}$, equals to the relative effect of an increase in peers' past performance compared to one's own. Note that with measurement errors, both α and β are subject to attenuation biases. When there is an equal amount of measurement error in both P_i and $\overline{P_{-i,G_i}}$, $\frac{\gamma}{\beta}$ is a measurement-error-corrected estimate of the relative effect.

Column (1)-(4) in Table 2.4 use the average peer past performance measures as the independent variables. Column (4) jointly estimates the effects of all peer past performances. As one would expect, the coefficients on own past performance are always positive and strongly statistically significant. The magnitude of these coefficients are not subtle. One standard deviation increase in one's own math performance measure leads to more than 0.3 GPA (more than a whole grade tier) increase. The effect of past English performance is much smaller in comparison.

The estimated effects of peers' past performances (γ) are not statistically significant. The magnitudes of the peer effect estimates are about a third of the effects from one's own past performances. Using the median statistics of peers' past performance in column (5)-(8), I also do not find significant peer effects.[17] These results show that peers' background academic performances have little impact on the course grade, even in a context where interactions are primarily academic. Such a conclusion highlights that studying with, on average, peers of "high performances" measured by past test scores is not necessarily effective at increasing course performance.

The above regressions assume that it is peers' average performances that affect course grades. Now I relax this assumption and turn to examine alternative peer effect models taken from Hoxby and Weingarth (2005). These models hypothesize different forms of peer effect structures. The Shining Light model claims that the maximum past peer performances have a positive effect on one's achievement. The Bad Apple model claims that the minimum past peer performances negatively affect achievement outcomes. The Focus model claims that peer homogeneity improves academic outcomes. Having some peers with high performances and some with low performances is better than only having peers of high performances.[18]

In order to check these models, I replace $\overline{P_{-i,G_i}}$ with alternative measures corresponding to these three models. To test for the Shining Light model, column (1) in Table 2.5 replicates the same regression in Table 2.4 but uses the maximum peer performances as $\overline{P_{-i,G_i}}$. The point estimates of own performance measures are largely unchanged to compared to those in Table 2.4. However, having peers with higher math performances has a negative and significant impact on the course grade. A one standard deviation increase in the maximum math ACT/SAT score of the peers is associated with about a 0.5 grade GPA (half a standard deviation) decrease. This finding is against the Shining Light model. Foster (2006) and Guryan et al. (2009) also find a similar negative effect, although their estimates are smaller than mine. Column (2) tests the Bad Apple model. I then use the minimum peer performances to test whether the lowest peer background performance measures matter. The estimates in front of the peer measures

---

[17]When I run column (4) and (8) specifications separately for female and male students, I still do not find evidence for peer effects.

[18]For a nice summary table of different models, please refer to Sacerdote (2011), p.255.

Table 2.4: Peer Effects: Peers' Performances on Course Grade

Course Grade	Using Avg. Peers' Ability				Using Med. Peers' Ability			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High School GPA (Peers')	-0.105			-0.324	0.046			0.486
	(0.558)			(0.604)	(0.916)			(0.743)
Math Performance (Peers')		0.011		0.009		0.002		-0.006
		(0.009)		(0.013)		(0.014)		(0.013)
Eng. Performance (Peers')			0.005	-0.001			0.011	0.011
			(0.011)	(0.012)			(0.012)	(0.011)
High School GPA	1.906***			1.565***	1.914***			1.585***
	(0.243)			(0.238)	(0.241)			(0.240)
Math Performance		0.034***		0.028**		0.034***		0.027***
		(0.00356)		(0.004)		(0.004)		(0.004)
Eng. Performance			0.014***	0.003			0.015***	0.004
			(0.003)	(0.003)			(0.003)	(0.003)
Control	Y	Y	Y	Y	Y	Y	Y	Y
Time Slot Fixed Effect	Y	Y	Y	Y	Y	Y	Y	Y
R^2	0.26	0.27	0.20	0.32	0.26	0.26	0.20	0.32
N	808	808	808	808	808	808	808	808

Notes: OLS regressions: Control variables include when one shows up for the first study group meeting, dummies for cohort, female gender, hispanic, asian and black ethnicity, lecture sections and special education programs, and dummies for missing achievement measures. The standard errors are clustered at the study group level (69 clusters) and are shown in the brackets.

Table 2.5: Results with Alternative Measures of Peers'
Background Performances

	(1)	(2)	(3)
	Measure of Peers' Performance		
	Maximum	Minimum	Std. Dev.
High School GPA	-3.129*	-0.109	0.711
(Peers')	(1.623)	(0.156)	(0.263)
Math Performance	-0.162***	0.003	-0.011*
(Peers')	(0.054)	(0.002)	(0.006)
Eng. Performance	-0.052	-0.002	0.009
(Peers')	(0.032)	(0.001)	(0.006)
High School GPA	1.544***	1.567***	1.560***
	(0.235)	(0.235)	(0.233)
Math Performance	0.027***	0.028***	0.028***
	(0.004)	(0.004)	(0.004)
Eng. Performance	0.003	0.003	0.003
	(0.003)	(0.003)	(0.003)
Control	Y	Y	Y
Time Slot Fixed Effect	Y	Y	Y
R^2	0.33	0.32	0.32
N	808	808	808

Notes: This is an OLS specification. Standard errors are shown in brackets. Control variables are the same as in the previous table. The standard errors are clustered at the study group level (69 clusters) and are shown in the brackets.

are small and statistically insignificant. To test the Focus model, I calculate the standard deviations of pees' performances. Column (3) shows weak evidence that a bigger spread in peers' math performance leads to a decrease in grades and the estimates are marginally statistically significant ($\beta = -0.011$, $p < 0.1$). Hence, there is also no supporting evidence for the Focus Model.

2.4.2 Effects from Peers' Non-performance Measures

In the previous section, I find little evidence for peer effects through background performances. Other dimensions of group composition can also play a crucial role. In this section, I explore other channels of peer effects relying on the variations in the non-performance measures.

These non-performance measures include the gender mix of the group, as well as the academic diversity of the group measured by the mix of different lecture sections. Previous findings show that being surrounded by female classmates positively impacts learning (Hoxby, 2000; Lu and Anderson, 2015). Gender composition is a topic that has received wide attention

in classroom and school organization. I examine the gender mix effect in an environment with direct academic interactions amongst the peers.

Since the course is a gatekeeper course and has multiple lecturers, students from different lecture sections might be exposed to different contents or different ways to understand the concepts even the contents are the same. Therefore one might benefit from studying with others from other sections by exchanging course related information. Therefore I will also examine the effect of mixing students by lecture section on the course grade.

In Figure 2.3, a student's course grade is plotted against the number of female peers and the number of peers from other lecture sections. In both Figure 2.3 (a) and (b), the non-performance measures are positively correlated with the course grade.

To quantify the effect size, I run the following specification,

$$y_i = \alpha + \beta P_i + \gamma Z_i + \eta X_i + \delta T_{G_i} + \epsilon_i \tag{2.2}$$

where Z_i measures the group mixture based on gender and lecture sections.

I use # Female Peers as an independent variable in column (1) of Table 2.6. The point estimate of 0.031 implies that a standard deviation in crease in the number of female peers, holding the group size constant, leads to about a 0.01 standard deviation increase in the course grade. In column (2), I use # OtherLec Peers as an independent variable. This variable equals to the number of group members from a different lecture section. The point estimate is a third of the effect of # Female Peers but it is not precisely estimated. In column (3) I include the square term to capture the nonlinear relationship.[19] Interestingly, the relationship between the lecture section mix and the course grade is not linear. Having peers from another lecture section is better than having none, but the benefit does not linearly increase. The results show that the optimal number of # OtherLec Peers is around 9 ($= \frac{0.116}{2*0.008}$) for a group an average size of 12 students. In the last column, I combine the two measures and the results are largely unchanged.

---

[19]The second order term for # Female Peers has a small and statistically insignificant estimate, and is therefore not included in the regressions.
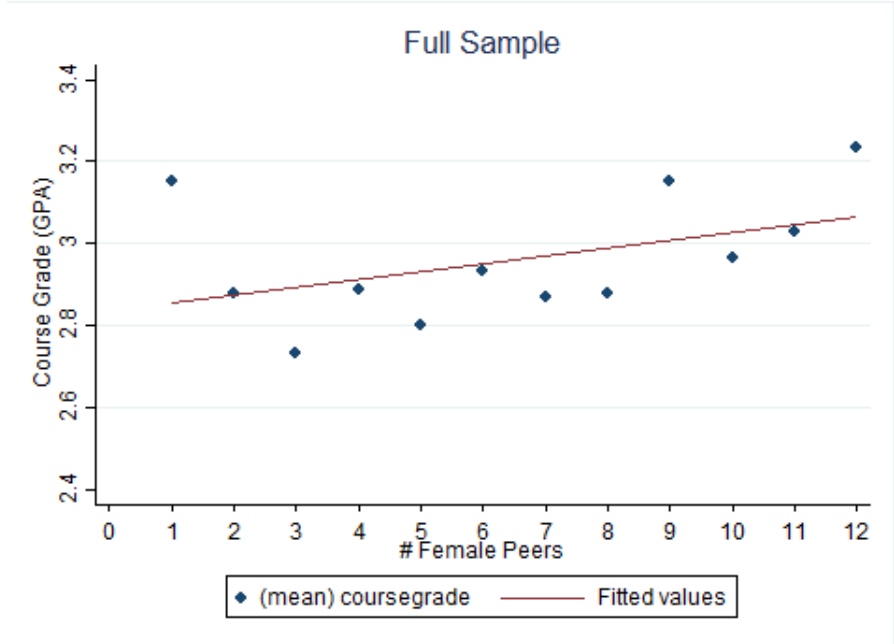
(a) By Gender Mix
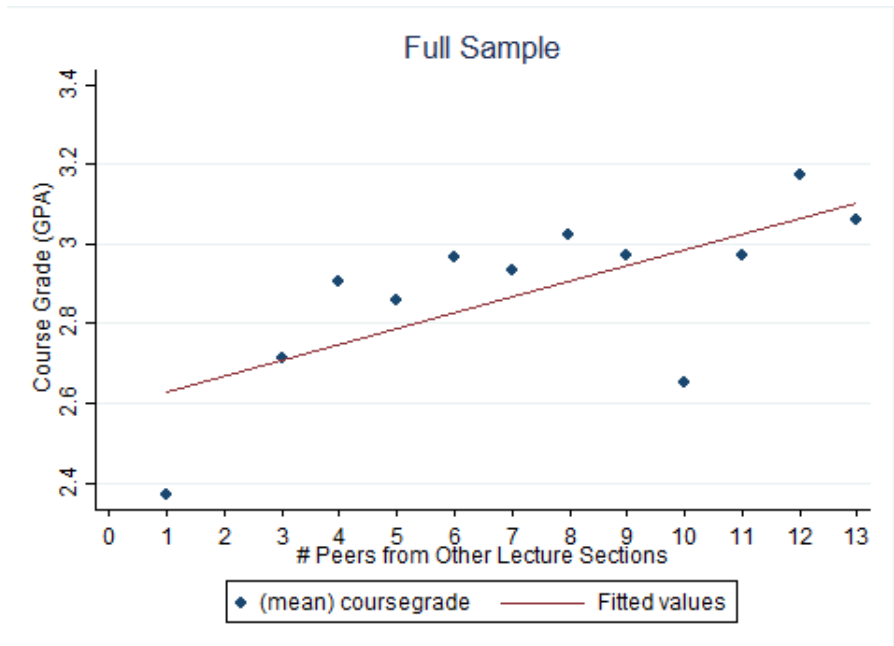
(b) By Lecture Section Mix

Figure 2.3: Peers' Non-performance Measures on Grades

Table 2.6: Peer Effects: From Non-performance Measures

DV: Course Grade	(1)	(2)	(3)	(4)
# Female Peers	0.031**			0.029**
	(0.013)			(0.012)
# OtherLec Peers		0.012	0.116*	0.109*
		(0.018)	(0.064)	(0.062)
# OtherLec Peers^2			-0.008*	-0.007*
			(0.004)	(0.004)
High School GPA	1.591***	1.581***	1.598***	1.613***
	(0.233)	(0.236)	(0.235)	(0.236)
Math Performance	0.027***	0.0275***	0.027***	0.027***
	(0.004)	(0.004)	(0.004)	(0.004)
Eng. Performance	0.003	0.003	0.004	0.004
	(0.003)	(0.003)	(0.003)	(0.003)
Female	-0.274***	-0.276***	-0.279***	-0.276***
	(0.055)	(0.054)	(0.054)	(0.054)
Controls	Y	Y	Y	Y
R^2	0.32	0.32	0.32	0.33
N	808	808	808	808

Notes: This is an OLS specification. Standard errors are shown in brackets. Standard errors are shown in brackets. Control variables include those from the previous table, plus group size. # Female Peers is the number of female group peers, with a mean of 6.6 and a standard deviation of 2. # OtherLec Peers is the number of group members taught by a different lecturer, with a mean of 6.5 and a standard deviation of 2.5.

I further cut the sample by gender and do not find that these effects differ by gender.[20] My findings support the policies to further increase the female presence in particularly these introductory STEM classes as both female and male students benefit from being surrounded with other female students. In addition, the results suggest that is an optimal way to mix students based on their lecture sections.

2.5   Mechanisms

Previous literature finding that student are affected by the number of female in one's environment do not offer concrete evidence to explain potential mechanisms. In this section, I explore mechanisms to explain why a student is better off when there are more female in the study group. One potential mechanism is that female students study in a different way from male students and the differences in study habits lead to differences in learning. Psychologists recently

---

[20]Results not shown but available from the author upon requests.

find that female students are more self-disciplined (Duckworth and Seligman, 2005) and being self-disciplined positively predicts student course performances after controlling for cognitive measures such as IQ (Almlund et al., 2011; Duckworth and Seligman, 2006).

Motivated by the psychology literature, I explore the usage of course website, especially how students access the course materials online.[21] Such behavior might be affected by personal traits as self-disciplined students are more likely to view in a browser tab or download these materials for learning purposes. I first categorize the online documents posted before the final exam (a total of 17) into the following five categories: general documents on learning guidance (a total of 4), practice questions (a total of 4), exam keys (a total of 3), topic handouts (a total of 3) and grading policies (a total of 3).[22] Because the grading policies should not affect students' efforts to achieve grades, I focus my analysis only on the downloads of the first four categories (a total of 14 materials). I ask whether being in more female group makes one more likely to assess these materials as well and whether more assess leads to a higher grade grade.

Table 2.7 Panel A is based on the following estimation equation to test the effect of assessing materials on the course grade:

$$Grade_i = \alpha + \beta * P_i + \gamma * usage_i + \eta X_i + \epsilon_i \tag{2.3}$$

For each category of materials, I set $usage_i$ to one if one has accessed all the materials of that category. I also have an overall measure of $usage_i$ which equals to one when one has downloaded all the 14 files. P_i and X_i control for individual academic performances and demographics. γ is expected to be positive and significant.

Panel B of the same table tests for the correlation between the group gender mix and the downloading behavior:

$$usage_i = \alpha + \beta * P_i + \gamma * \{\#\text{Female Peers}_i\} + \eta X_i + \epsilon_i \tag{2.4}$$

γ is expected to be positive and significant.

---

[21] Due to data limitation, I cannot observe how students utilize the materials.
[22] For a detailed list of all the materials, please see the Appendix.

Table 2.7: Course Website Usage as a Mechanism

Panel A DV: Grade	(1) All	(2) General Docs	(3) Practice Q	(4) Exam Keys	(5) Handouts
Usage	0.162	0.021	0.127**	0.130**	0.126**
	(0.120)	(0.079)	(0.060)	(0.063)	(0.059)
Controls	Y	Y	Y	Y	Y
N	808	808	808	808	808
Panel B DV: Usage	(1) All	(2) General Docs	(3) Practice Q	(4) Exam Keys	(5) Handouts
# Female Peers	0.012**	0.007	0.006	0.024**	0.021**
	(0.003)	(0.006)	(0.008)	(0.011)	(0.010)
Controls	Y	Y	Y	Y	Y
N	808	808	808	808	808

Notes: Controls include past high school gpa, math and English performances, plus dummies for study group meeting time dummies. Standard error adjusted for 69 clusters at the study group level.

The estimates in Table 2.7 Panel A show that the action of downloading exam keys and topic handouts is positively correlated with the course grade. A student who has downloaded all the exam keys on average gain a 0.127 course GPA increase than a student who does not do so. The magnitude of the estimates in Column (4) and (5) is very similar. Panel B suggests that the number female peers in groups positively impacts downloading exam keys and handouts. Taking these two observations together, I conclude that female group members help a student learn better making him/her more likely to download course materials which are helpful for getting a higher course grade.

2.6 Conclusion

It is well known that peers are an important factor in learning. Study groups are popular teaching tools but very few have studied how study group composition affects student academic outcomes. This paper closes this gap and provides direct and clean evidence of peer effects on learning outcomes at the study group level. I exploit a unique dataset from a natural experiment. The Science Learning Center at the University of Michigan organizes study groups for main STEM field introductory classes and the sign up process creates an exogenous variation in peer group composition.

Although students with slightly worse past academic performances chose to join the study

groups, their average course grade is almost a letter grade higher than those who did not participate. After presenting evidence that the study group assignment is plausibly exogenous, I test various peer effect models.

I find little evidence that peers' background academic performance measures have effects on learning outcomes using the linear in the mean model. Non-linear models such as the Bad Apple, Shining Light, and Focus Model. Consistent with previous studies, I find that gender mix has an economically and statistically significant impact on learning. Specifically, an increase in the number of female peers leads to an increase in the course GPA. I also find that studying with peers from another lecture section improves one's course grade and there is an optimal mixture.

More importantly, I show evidence for a potential mechanism underlying peer effect in study groups. I show that the gender mix can affect the course grade through changing how students study.

My results overall suggest that educators should assign students into groups based on particular background characteristics such as gender instead of SAT/ACT test scores. Also, mixing students from different sections may foster knowledge sharing and thus promotes better learning outcomes.

One future extension of this work is to gather intermediate learning outcome measures such as homework scores in order to unpack how peers affect the learning process.

Chapter 3

Laboratory on the Social Network: Homophily and Peer Influence for Economic Preferences

with Erin L. Krupka and Steve Leider

3.1   Introduction

Ample cross-sectional data find that the behaviors of individuals in one's social network are positively correlated with one's own behavior. This correlation is observed for behaviors such as academic performance (Burke and Sass, 2013; Calvo-Armengo et al., 2009; Sacerdote, 2001), technology uptake (Conley and Udry, 2010), health outcomes (Christakis and Fowler, 2007, 2009; Stevens and Prinstein, 2005), labor market participation (Calvo-Armengol and Jackson, 2004; Topa, 2001), crime participation (Glaeser and Sacerdote, 1999), new product diffusion (Aral and Walker, 2011a,b), and the exchange of goods and information (Bramoullé et al., 2009; Granovetter, 1985). Another stream of cross-sectional studies has focused on behavior and diagnostic tasks that are more closely tied to fundamental economic preferences. These papers find suggestive evidence that preferences are also correlated within a social network.[1]

However, these cross-sectional studies have only looked at correlations at a particular point

---

[1]Dohmen et al. (2012) find that an individual's risk preference and levels of trust are positively correlated both with their parents' preferences, and with the average preferences of their region. Bettinger and Slonim (2006) and Bettinger and Slonim (2007), by contrast, find no correlation between the altruism or patience of a parent and a child.

in time, and are not able to speak to why and how such correlations evolve. The detection and measurement of network effects that produce the observed correlations in behavior and preference correlations in individual's network is a difficult exercise but a critical one for understanding how institutions and social contexts shape behavior (Fehr et al., 2013; Fehr and Hoff, 2011; Frank et al., 2013). This correlation can occur because we tend to select friends because they are similar to us (referred to as "homophily"), or because we become more similar to our friends over time (referred to as "influence").[2] Distinguishing influence from homophily (while eliminating confounds) requires dynamic, longitudinal network information about the emergence of ties between people in a network and also separate measures of behaviors, aspirations and preferences of the individuals in the network (i.e., repeated measures on preference constructs, on behaviors such as academic course selection or study group attendance and on attitudes such as aspirations as major or graduate). Ahern et al. (2014) look at how the average preferences in an MBA student's randomly assigned class section affects an individual's own preferences relative to their initial preferences before enrollment. The authors find a positive peer effect for risk attitudes, and a negative peer effect of altruism. By using the pre- and post-enrollment measures of preferences, the authors are able to provide causal evidence for peer influence on preferences. But they cannot look at both homophily and influence at the same time. In this paper, we use a longitudinal survey design to track both homophily and influence over time.

Another feature in most of the cross-sectional studies is that they construct social networks based on random peer group assignment (e.g. random assignment into lecture sections in Ahern et al. (2014)) or geographic proximity (e.g. regions of residence in Dohmen et al. (2012)) to identify the influence of the social and institutional environment on preferences. Random assignment into peer groups eliminates the confound that arises when individuals can pick who is in or out of their network. However, earlier work in psychology and sociology suggests that individuals with whom one associates or with whom one shares a similar identity are more influential (Cialdini and Garde, 1987; Granovetter, 1985; Tajfel and Turner, 1979). Networks

---

[2]Quite many papers show evidence for homophily (Currarini et al., 2010; Fisher and Bauman, 1988; Kandel, 1978; Mayer and Puller, 2008; McPherson et al., 2001). Another well-established stream of literature shows evidence for influence (Bauman and Ennett, 1996; Fehr and Hoff, 2011; Meyer and Waller, 2001; Oster and Thornton, 2012). Very few examine these two processes at the same time.

based on geographic proximity may not be a relevant transmission channel for social influence. We directly elicit a student's social network to define *relevant* peers to circumvent these issues.

We test for the effects of friend behavior on social network formation as well as behavior and economic preference change in the context of newly arriving undergraduate students.[3] The preferences we focus on are time, risk and altruism. Understanding individual decision-making under risk and over time, as well as an individual's inclination to exhibit generosity toward others, are three foundations of economic analysis and are correlated with important life decisions.[4] These preferences also influence many academic outcomes. For example, risk preferences describe an individual's willingness to accept more or less risky choices and are correlated with decisions such as whether to enter competitive environments such as high stakes testing or entrance exams to competitive schools (Ors et al., 2013). Time preferences characterize how an individual is willing to trade off amongst costs and benefits that occur at different times and have been shown to predict cumulative GPA of college graduates and whether or not they complete college within four years (Burks et al., 2015). Altruism preferences describe a person's concern for the welfare of others and are correlated with college students' willingness to donate to a charitable fund offering low interest loans to financially challenged students (Benz and Meier, 2008).

We focus on testing the effect of social networks on risk, time and generosity preferences as well as testing the dynamics of adding and dropping friendship ties. To do this we create a laboratory on the social network in which we measure the emerging social network of 399 incoming freshman at the University of Michigan. Our data collection strategy consists of (1) recruiting voluntary participants, (2) mapping students' social network intensely in their first year and (3) using surveys and economic experiments to measure key variables of interest during the

---

[3]We use the freshman cohort as our sample because when college students arrive on campus for the first time, they are also arriving to a whole new social environment from which they will build new friendship, mentoring, studying, and employment networks. During their first year away from home, they will develop new work and personal habits and they will make choices about college courses and majors that will impact them for years to come. The newly formed social networks, and associated social capital (Coleman, 1988), can have profound effects on their experiences at college - from which major they choose to where they get a job on campus, to whether they experience mental health issues during their college years.

[4]Social preferences have been linked to macro economic phenomena such as cross county variation in Gross Domestic Product, and poverty disparities (Karlan et al., 2009; Knack and Keefer, 1997). Risk preferences have been correlated with investment, retirement health related behavior and career choices, and time preferences have been linked to smoking, obesity, educational and savings behavior (Cardenas and Carpenter, 2008).

63

course of their first year. Unlike other studies that provided only cross-sectional evidence, or looked only at one mechanism (influence or homophily), our design allows us to observe and distinguish between two mechanisms for correlated behavior on a network: network dynamics and peer influences. Additionally, we can provide direct evidence for homophily based on behavioral measures of economic preferences (rather than behaviors correlated with economic preferences), and evidence for influence of social peers on preferences (rather than an influence of family members such as parents on preferences).

To measure students' social networks we use an incentivized elicitation method taken from Leider et al. (2009). In each of three phases (October, January and April), participating students are asked to name ten other freshman as friends. Students receive a monetary lottery payment for each named student that names them back. For each phase, after the friendship elicitation stage, subjects are invited to also complete an incentivized preference elicitation survey. Because we want to have a measure of fundamental economics preferences, we use diagnostic tasks that are closely tied to our preferences of interest (risk tolerance, patience and altruism). We use multiple price lists between safe and risky outcomes to measure risk attitudes, multiple price lists between sooner and distant payments to measure patience, and dictator allocation decisions to measure altruism. We also ask subjects to self-report their tolerance for risk and their patience. Using this data we can ask two broad sets of questions: (1) are changes in a student's network (e.g. adding or dropping a friendship link) between phases driven by the similarity between the two students on some economic preference? (2) Are a students economic preferences influenced by the corresponding preferences of the students current and/or past social ties?

We find in the affirmative for both questions. First, we find that an individual is significantly more likely to add a friendship tie with another student, and significantly less likely to drop an existing friendship tie, if that student has similar level of generosity to the focal student. The magnitude of homophily is substantial a one standard deviation increase in similarity changes the likelihood of adding or dropping by an amount equal to 20-50% of the base rate. It is also large compared to other factors that one would expect might drive network changes a one standard deviation change in altruism similarity has an effect half as large as participating in

the same activity, and is about 2.5 times as large as a one standard deviation increase in the network centrality of the other student.

Additionally, we find evidence for peer influence on risk preferences, as well as some evidence for peer influence on time preferences. We find that an individual's self-reported tolerance for risk is significantly correlated with the average rating for both the individual's friends, and their broader network community. The correlation is also robust to looking at both the average behavior of others in the same phase as well as the lagged behavior from previous phases. A one standard deviation increase in the average choices of an individual's friends or social community is associated with an increase of 1/8th to 1/10th of a standard deviation for that individual. Similarly, we find that an individual's incentivized patience choices are correlated with their friends and community. A one standard deviation increase in friends' or community's patience increases the patience of the focal individual by 1/7th to 1/12th of a standard deviation. The results for risk preferences, but not time preferences, are robust to the inclusion of lagged dependent variables, although with our short panel (and the incomplete participation of some subjects) this substantially reduces our sample size.

We additionally find robust evidence for a negative peer influence effect on generosity. That is, individuals become significantly less generous when they have friends or a social community that are particularly altruistic. This result is consistent with free-riding behavior, and aligns with recent evidence from a related study by Ahern et al. (2014) on peer influence among MBA students.

Our contribution is to provide evidence for two separate mechanisms (homophily and influence) that can be the source of the often observed contemporaneous correlation between an individual's behavior or preferences and their social network's behavior and social preferences. Additionally, we demonstrate these effects for the subtle (but broadly important) underlying economic preferences, rather than the observable but potentially domain-specific behaviors previously studied. Further this work advances both the understanding of how a student's social network and preferences evolve upon entering college, and how a student's social network affects one's economic preferences.
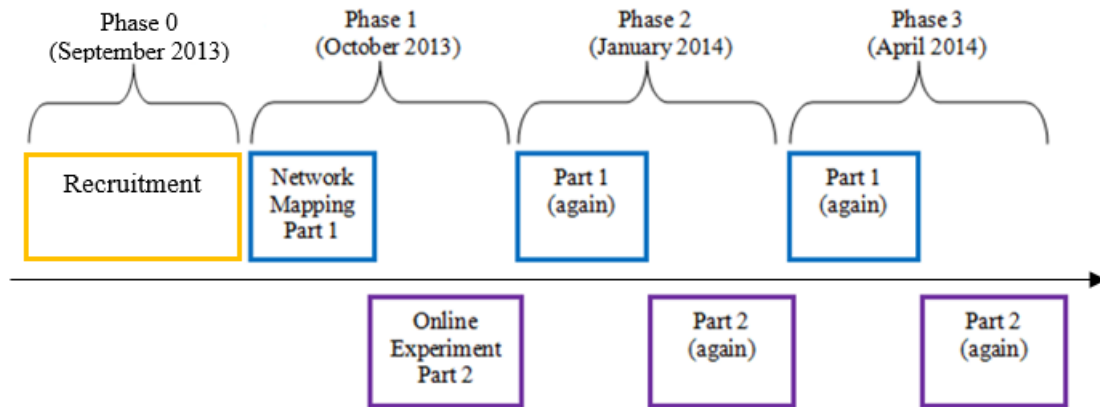
Figure 3.1: Research Structure and Data Collection Timeline

3.2 Research Design

Figure 1 shows an overview of the experimental timeframe. Our target population were the incoming freshman at the University of Michigan. To this end, we needed to create our own freshman subject pool from which to recruit subjects into our experiment. Thus, in Phase 0 we recruited freshman using (the traditional method of) flyers and we also developed a Facebook application. We created the Facebook application as both a method of gathering general personal data (including friend lists) and a recruitment tool for the next phases of the research. All students who were eligible to participate had to provide a valid university email address when enabling the application; this allowed us to verify that they were in fact incoming University of Michigan freshman (this was crossed checked with a database of freshman email addresses which we obtained in cooperation with the Registrar).[5] We concluded recruiting by late September of 2013; all subjects who had consented during the month of September were eligible for receiving emails which invited them to participate in our online experiments in Phase 1 to 3.

Having concluded recruiting, we moved on to Network Mapping stage of Phase 1. Consented freshman were contacted via email and invited to participate in an online survey. Subjects were told that the online study would consist of two parts and that they would be paid for

---

[5]We designed the Facebook application to pilot a mechanism for reaching and identifying densely connected subsets of students. However, for this study we ended up using everyone who agreed to participate.

their participation after completing both parts. In part one we mapped subjects' social network. Subjects were sent a link to part one and told that they could complete part one (and subsequently part two) within a 3 day window. However, once they opened the link to part one, they would be required to finish part one before receiving a separate link to part two. They were told that they had 30 minutes to complete part one, once they had moved to the next screen they would not be able to return to the previous question to change their answers and that they would not able to come back to Part 1 of the study once they logged out of Part 1 or once the 30 minutes are up.[6]

To elicit the network we used a protocol developed by Leider et al. (2009).[7] The protocol asks a participant to list their 10 best friends on campus and pays subjects a bonus for any friend who also participated and listed them as well. If the subject listed a friend who also completed the survey and listed her as well, she received a 50 percent chance of getting a prize of $0.50. Otherwise, she was be paid nothing. If both also agreed on the amount of time spent together each week the winning probability was increased to 75 percent. However, if the subjects names a person who does not name them, then they will received nothing. Thus, subjects had ten independent chances to win $0.50 because they could name up to ten friends. Following Leider et al. (2009) we use a lottery-based incentive so that subjects could not definitively know if the person they named had also named them. This avoids the problem that a subject may choose not to list someone because they do not want to know if the person listed them back, or that a subject would feel obliged to list someone to avoid giving offense.

Figure C1 in Appendix C shows how a subject reports a freshman friend. She first selects the last initial of the friend in the first drop down menu. Then the system automatically filters the second drop down menu so that all the last names in the drop down menu start with the chosen last initial. Then the system again filters the third drop down menu so that it only contains the names of freshmen with the chosen last name. After a subject chooses the first name, the fourth drop down menu displays the university email address of this selected friend

---

[6]We timed how long it would take and determined that for most it would take about 10 minutes to answer all of the questions.

[7]There are multiple ways to collect social network data. Advantages and disadvantages of those methods are discussed in Carrington et al. (2005) and Brañas-Garza et al. (2013).

for confirmation.

After selecting a friend, a subject also indicated how much time they spent with the friend (the scale ranged from "0-30 minutes" to "more than 8 hours" per week) and also whether the person was a roommate or not.[8] We also incentivize truthful time reporting by telling a subject that she can increase the probability to win the $0.50 prize from 50 percent to 75 percent for each friend who lists her and also agrees on the amount of time spent. Thus, when subjects completed part one, we obtained a list of whom they consider to be their ten best friends (as well as a description of how much they interact with each friend). When a subject's friend completed the survey, we also see whether the friend lists her.

Once subjects completed part one, they were sent a link to part two of the study, the online experiment. In this part of the study they were asked to make a series of financial decisions; they had 30 minutes to complete the study. The online experiment included questions to elicit the subjects level of altruism (using a dictator game), risk aversion, and time preferences. We use diagnostic tasks to elicit risk, time and social preferences, as well as subjects' guesses about the behaviors of the ten friends from the friendship elicitation survey. For each preference we use a diagnostic choice task that is well established in the literature to identify and measure preferences along each of our dimensions. All the games and guesses are incentivized.

To elicit risk attitudes, we employ a 15 question multiple price list, where subjects make choices between a 50-50 lottery with prizes of $200 or $0, and a fixed payoff that ranges from $0 to $140. For time preferences, subjects make choices between an $80 payment in three months, or a payment in two weeks that ranges from $5 to $75. See Figures C2 and C3 for a picture of the two tasks.

For the social preference elicitation task (shown in Figure C4), subjects make a "dictator allocation", where they divide 100 tokens between themselves and another randomly selected anonymous participant. Tokens are worth $0.75 to the person making the decision about how to divide the tokens, and $1.50 for the anonymous recipient.

After subjects completed these parts, they were also asked to guess about the average

---

[8]We asked them to report the number of hours with the friend but not to include class time.

choices (among participants) of the ten friends from the friendship elicitation survey. They were asked to guess if the average choices of their friends are similar to their own choice (within +/- 1 for risk and time, +/- 10 for social preferences), or are above or below their own choice. Subjects could also answer that they are not sure. Subjects were rewarded $1.00 for correct guesses (and receive $0.50 if they say they are not sure).

Finally, subjects made un-incentivized self-reports about risk tolerance and patience (shown in Figures C5 and C6). Subjects rated themselves on a 10 point scale. The risk scale ranges from "risk averse" to "fully prepared to take risks", while the patience scale ranges from "very impatient" to "very patient". Finally, subjects also reported how frequently they engage in various activities (such as volunteering, attending religious activities, eating out and so on).

For payment of the online experiment, one out of every seven subjects were randomly selected for payment. For selected subjects, we randomly selected one incentivized task, and then for the risk and time preference tasks randomly selected one choice to implemented for payment. All subjects were informed of this payment mechanism. Subjects could choose to be paid via electronic or physical gift card, or check.

The two parts of Phase 1 were repeated three times in total (in Phase 1, 2 and 3). All phases were identical to the procedures described above for Phase 1.

This design generates a panel dataset where for each subject (conditional on participation) we observe their friendship linkages, and their preferences, at three points in time. If they have friends who also participate we can also observe the preference of their friends. This allows us to observe the two main dynamics of interest: changes in the social network over time, and changes in preferences over time. We can then test for our two mechanisms of interest: homophily based on economic preferences, and peer influence on economic preferences. The results section below describes our analysis strategy and predictions.

3.3 Results

3.3.1 Participation and Network Information

A total of 399 freshman participated in our experiment. However, subjects are most useful to use when they participate in both the friendship elicitation and behavioral measurement survey. Furthermore, to answer many of our questions we need subjects to participate in multiple phases of the experiment. Table 3.1 reports the number of students participating in each survey of each wave. We have approximately 200 students who participated fully in at least two waves. Finally, for many of our research questions we need to have the friend of a subject also participate. In total we have 199 subjects who fully participated in, and had a friend also participate in, at least one phase.

Table 3.1: Student Participation

Phase	Friendship	Behavior	Both
Phase 1	399	287	287
Phase 2	198	189	177
Phase 3	298	211	202

On average subjects had 1.2-1.5 friends who also participated (depending on the phase), and for those friends the average amount of reported time spent together was "2 to 4 hours per week". Housing arrangements played an important role in friendships. In Phase 1 every student listed their roommate as a friend, while in Phase 3 80% of students listed their roommate. Similarly, 64% of friends in Phase 1, and 68% of friends in Phase 3, were from the same dorm. Shared academics and activities were also relevant, with approximately 20% of friends having the same major and approximately 50% of friends participating in the same club.

Figure 3.2 shows the measured social network from Phase 1 and 3. In Phase 1 there was one large connected component of 242 students, and then a number of smaller components. The average clustering coefficient was 0.181, while the average eigenvector centrality was 0.141. Phase 3 had a somewhat more fracture network the largest component had 49 students, however there were three other moderately large components. The average clustering coefficient decreased slightly to 0.154, while the average eigenvector centrality decreased to 0.123.
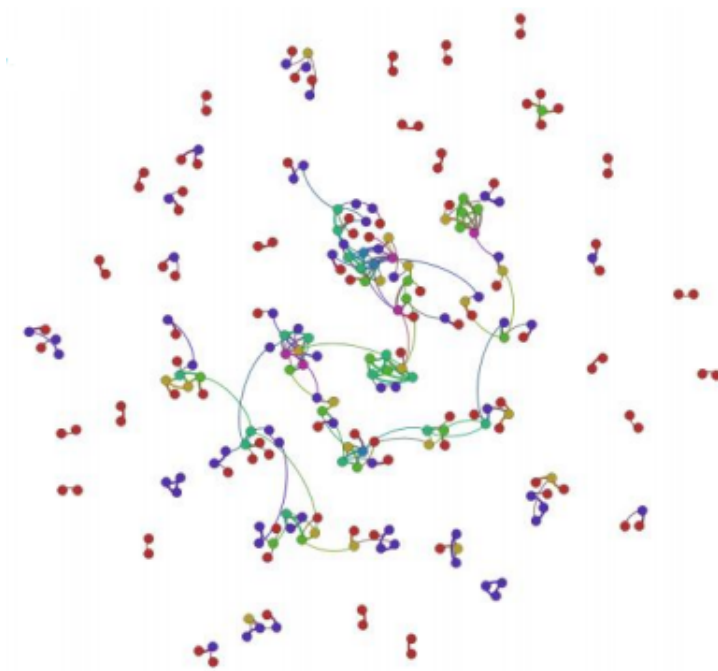
(a) Phase 1
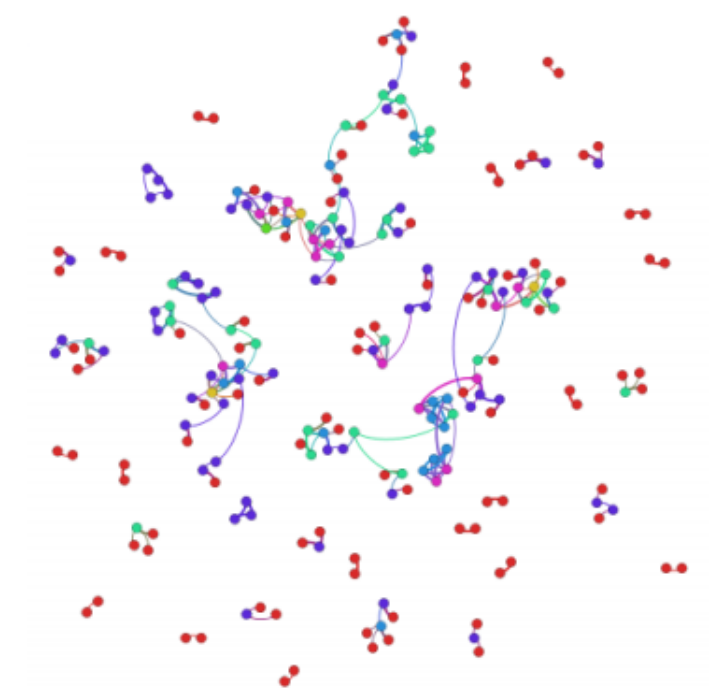
(b) Phase 3

Figure 3.2: Peers' Non-ability Measures on Grades

3.3.2   Behavioral Measurements

We next turn to the preferences and choices elicited by our behavioral survey. Table 3.2 reports the average choices in our incentivized choice tasks, as well as the self-reported measures. For both Risk and Time preferences almost every participant made monotonic choices. Subjects were on average moderately risk averse, with the average indifference point being $75 (compared to the lottery with an expected value of $100). As expected, the incentivized behavioral measure is correlated with the self-reported measure of risk tolerance subjects who rated themselves higher on willingness to take risks chose the safe option fewer times (β = -.307, s.e. = .081). Subjects were also moderately impatient, with the average indifference point being a sooner payment of $64 (compared to a delayed payment of $80). The self-reported measure for patience is also correlated with the behavioral measure subjects who rated themselves as more patient chose the sooner payment less often (β = .260, s.e. = . 087). Allocations from our social preference task indicates a moderate level of altruism, with the average allocation being $45 for the decision-maker and $60 for the recipient.

While we are interested in how a student's social context influences their preferences, and hence we expect some amount of change in the observed preferences, we do want to make sure that our behavioral measures are capturing true fundamental preferences. Therefore, we should expect that the observed measures have in general a fair amount of stability. Fortunately, all of our behavioral measures do appear to be fairly stable. Each measure is quite highly correlated with itself in each pair of phases.[9]

3.3.3   Network Dynamics

We now examine in greater detail how the social network changed over the course of the year. We can first observe that there is a fair amount of turnover among our subjects' friendships.

---

[9]In order to make sure that there is no selection bias between those freshmen who agree to participate in our survey and those who do not, in February 2016 we conduct the surveys again to compare a sample of subjects recruited through these methods with subjects from the standard lab pool. To do so, we send out the network survey to all current freshman 2016 cohort. Among those finish the network elicitation survey, we invite 80 of them to take the behavior survey. Then we also contacted 80 randomly selected freshman students and invite them to the same behavior survey as well. Both groups have the same behavior survey completion rate (= $\frac{44}{80}$). We see no significant differences in survey responses between these two groups for any of our outcome measures. See Appendix C for more details.

Table 3.2: Behavioral Measures Summary Statistics

	Behavioral Measures			Self-Reported Measures	
	Risk	**Time**	**Social**	**Risk**	**Time**
Phase 1	8.53	3.21	59.5	5.76	5.88
Phase 2	8.58	3.21	59.86	5.79	6.16
Phase 3	8.17	3.51	60	5.83	6.04
Correlation (1 to 2)	0.532	0.549	0.528	0.614	0.601
Correlation (2 to 3)	0.637	0.609	0.603	0.775	0.522
Correlation (1 to 3)	0.534	0.616	0.475	0.633	0.524

Notes: The Risk behavioral measure reports the average number of safer choices. The Time behavioral measure reports the average number of sooner choices. The Social behavioral measure reports the average allocation to the other recipient. The self-reported measures are 10-point scales with larger values denoting a greater willingness to take risks/be patient.

Between Phase 1 and Phase 2 subjects kept on average 6.3 friends, and changed 3.7 friends. Between Phase 1 and Phase 3 subjects kept 5.5 friends and changed 4.5 friends. Our primary question, then, is what factors affect an individual's decision to add or drop friends. Specifically, are individuals more like to add friendships with individuals who are similar to themselves on certain dimensions and/or more likely to drop friendships with those who are dissimilar. Specifically, for each characteristic X, we want to see whether the absolute difference $\Delta_X = \|X_i - X_j\|$ in that characteristic between an individual i and a (potential) friend j is predictive of a change in their friendship. We will use a probit specification as follows:

$$\text{prob}(\text{Change}_{ij}) = \Phi(\alpha + \beta\Delta_X + \epsilon) \tag{3.1}$$

If students expressed homophily along a particular behavioral dimension we would expect a negative β for additions (decreased likelihood of forming friendships with dissimilar others) and a positive β for drops (increased likelihood of ending friendships with dissimilar others).

In analyzing added friendships we need to identify a pool of possible new friends. One natural set of potential new friends is the other students from the same dorm. For another measure, we use Clauset et al. (2004)'s community detection algorithm to identify the close social environment of the student. The community detection algorithm partitions the social network so that the number of links within communities is as large as possible, and the number of links between communities is as small as possible. We can then use anyone in the student's community that is not currently their friend as a socially individual who is a candidate to be

Table 3.3: Network Dynamics for Risk Measures

Panel A: Behavioral Measures

	Add Friend				Drop Relationship			
	From Dorm		From Community		Friend		Community	
	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ_{Risk}	-0.0306	-0.0004	0.025	0.0012	0.033	0.0121	-0.0308**	-0.0117**
	(0.0198)	(0.0003)	(0.021)	(0.0010)	(0.0412)	(0.015)	(0.0141)	(0.0054)
Const.	-2.515***		-2.144***		-0.535***		0.427***	
	(0.0714)		(0.105)		(0.191)		(0.0831)	
# Obs	11,854		1,957		180		1,374	

Panel B: Self-Reported Measures

	Add Friend				Drop Relationship			
	From Dorm		From Community		Friend		Community	
	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ_{Risk}	-0.0333	-0.0004	-0.0103	-0.0005	0.0482	0.0176	0.0176	0.0067
	(0.0319)	(0.0004)	(0.0382)	(0.0019)	(0.0625)	(0.0229)	(0.025)	(0.0095)
Const.	-2.538***		-2.032***		-0.518***		0.278***	
	(0.083)		(0.116)		(0.155)		(0.085)	
# Obs	11,854		1,957		180		1,374	

Notes: All specifications are probit regressions with standard errors clustered at the subject level. Odd numbered columns report coefficients, while even numbered columns report marginal effects. Dependent variables are: adding a friendship (columns 1-4), dropping a friendship (columns 5-6), or ceasing to be in the same network community (columns 7-8). For columns 1-2 the sample includes all subjects who are in the same dorm (but not currently a friend). For columns 3-4 the sample includes all subjects who are in the same network community (but not currently a friend). For Panel A the independent variable is the absolute difference in the number of sure choices, while for Panel B it is the absolute difference in the self-reported measure of risk tolerance.

a new friend. Similarly, we can look at changes in the network such that an individual is no longer a friend, or no longer in the same community.

Table 3.3 reports the results of regressing the likelihood of adding a friend from the dorm or network community and dropping a friendship or shared community, on the absolute difference in risk measures. Panel A uses the incentivized behavioral measure of risk, while Panel B uses the self-reported measure.

Overall we find relatively little evidence that risk attitudes matter for the dynamics of our subjects' social network. There is no effect of a pair's similarity in either our incentivized risk measure or the self-reported on the likelihood of adding or dropping direct friendship linkages. We do see a significant effect for the likelihood of no longer being in the same social community however the sign of the effect is the opposite of what one would expect from a homophily dynamic. Our results suggest that a pair of individuals who have more dissimilar risk attitudes are less likely to change social communities. Table 3.4 presents the

same regression specifications for the absolute difference in patience. In line with our results for risk attitudes, we find no effect of the difference in any measure of time preference on any measure of network dynamics.

Table 3.4: Network Dynamics for Time Preference Measures

Panel A: Behavioral Measures

	Add Friend				Drop Relationship			
	From Dorm		From Community		Friend		Community	
	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ_{Time}	-0.0112	-0.0001	0.0072	0.0003	0.0396	0.0145	-0.0179	-0.0068
	(0.0181)	(0.0002)	(0.0239)	(0.0011)	(0.0386)	(0.0141)	(0.0152)	(0.0058)
Const.	-2.570***		-2.083***		-0.534***		0.383***	
	(0.0795)		(0.123)		(0.152)		(0.081)	
# Obs	11,854		1,957		180		1,374	

Panel B: Self-Reported Measures

	Add Friend				Drop Relationship			
	From Dorm		From Community		Friend		Community	
	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ_{Time}	-0.0346	-0.0005	-0.0407	-0.0019	-0.0362	-0.0131	0.0093	0.0035
	(0.0287)	(0.0004)	(0.0436)	(0.0021)	(0.0557)	(0.0201)	(0.0208)	(0.0079)
Const.	-2.522***		-1.960***		-0.341**		0.301***	
	(0.090)		(0.141)		(0.157)		(0.087)	
# Obs	11,854		1,957		180		1,374	

Notes: All specifications are probit regressions with standard errors clustered at the subject level. Odd numbered columns report coefficients, while even numbered columns report marginal effects. Dependent variables are: adding a friendship (columns 1-4), dropping a friendship (columns 5-6), or ceasing to be in the same network community (columns 7-8). For columns 1-2 the sample includes all subjects who are in the same dorm (but not currently a friend). For columns 3-4 the sample includes all subjects who are in the same network community (but not currently a friend). For Panel A the independent variable is the absolute difference in the number of sooner choices, while for Panel B it is the absolute difference in the self-reported measure of patience.

By contrast, we find evidence for differences in generosity influencing several dimensions of change within the social network. Table 3.5 reports the results of regressing relationship changes on the absolute difference in the number of tokens kept in the allocation game. We find that an individual is significantly less likely to add a potential friend (either from the set of students in the same dorm, or in the same social community) the more dissimilar they are in generosity. Specifically, a one standard deviation increase in the difference in tokens kept would lead to a 0.19 percentage point decrease in the likelihood of adding someone in the same dorm as a friend, approximately half the base rate probability of 0.38% of adding them as a friend. We also find that greater dissimilarity increases the likelihood of dropping a friendship with someone. A one standard deviation increase the dissimilarity would increase the likelihood of

dropping someone as a friend by 7.3 percentage points, one fifth of the base rate probability of 33.9%. These results are consistent with students exhibiting homophily with respect to generosity. Our subjects prefer to form friendships with those who are similarly generous, and drop friendships with those who are differently generous.

Table 3.5: Network Dynamics for Generosity

	Add Friend				Drop Relationship			
	From Dorm		From Community		Friend		Community	
	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ_{Kept}	-0.0108***	-0.0001***	-0.0110***	-0.0005***	0.0107**	0.0039**	0.0021	0.0008
	(0.0032)	(0.000)	(0.0042)	(0.0002)	(0.0051)	(0.0019)	(0.0023)	(0.0009)
Const.	-2.429***		-1.886***		-0.655***		0.277***	
	(0.0639)		(0.0835)		(0.147)		(0.0836)	
# Obs	11,854		1,957		180		1,374	

Notes: All specifications are probit regressions with standard errors clustered at the subject level. Odd numbered columns report coefficients, while even numbered columns report marginal effects. Dependent variables are: adding a friendship (columns 1-4), dropping a friendship (columns 5-6), or ceasing to be in the same network community (columns 7-8). For columns 1-2 the sample includes all subjects who are in the same dorm (but not currently a friend). For columns 3-4 the sample includes all subjects who are in the same network community (but not currently a friend). For Panel A the independent variable is the absolute difference in the number of sooner choices, while for Panel B it is the absolute difference in the self-reported measure of patience.

Risk, time and social preferences are potentially subtle and hard to observe aspects of an individuals behavior. While the effects of differences in generosity seem to be large compared to the base rates for relationship changes, it is difficult to say whether the effect is as large as one might expect. In order to provide an alternative benchmark, we can also examine the effect on friendships of two clear markers of shared interests: sharing the same academic major, and participating in the same extracurricular club. Intuition suggests that these traits should have a strong effect on friendship formation and maintenance. For each pair of subjects we construct an indicator variable I_X that equals 1 if the two subjects share the same major, or participate in the same club (respectively). We then estimate the previous regressions using the indicator variable in place of the absolute difference measures:

$$\text{prob}(\text{Change}_{ij}) = \Phi(\alpha + \beta I_X + \epsilon) \tag{3.2}$$

In these regressions homophily would predict a positive β for additions and a negative β for drops. The results are reported in Table 3.6.

For shared major we find significant effects for reducing the likelihood of dropping a friend-

ship, but no effect for adding a friendship. By contrast, we find a significant effect of shared clubs for both adding and dropping friendships. If we compare these effects to the previously estimated effect of differences in generosity, we see that the effect of shared activities is about twice the magnitude of a one standard deviation difference in generosity. This provides an alternative demonstration that homophily with respect to generosity plays a substantial role in the changes in the social network.

Table 3.6: Network Dynamics for Shared Interests and Activities

Panel A: Same Major

	Add Friend				Drop Relationship			
	From Dorm		From Community		Friend		Community	
	Coeff. (1)	M. Eff. (2)	Coeff. (3)	M. Eff. (4)	Coeff. (5)	M. Eff. (6)	Coeff. (7)	M. Eff. (8)
$I_{SameMajor}$	0.0957	0.0012	-0.0152	-0.0004	-0.507*	-0.171**	-0.280**	-0.109**
	(0.0949)	(0.00129)	(0.0689)	(0.00196)	(0.270)	(0.0825)	(0.135)	(0.0531)
Const.	-2.681***		-2.287***		-0.316***		0.362***	
	(0.0393)		(0.0349)		(0.115)		(0.0685)	
# Obs	23,684		21,176		180		1,374	

Panel B: Same Club

	Add Friend				Drop Relationship			
	From Dorm		From Community		Friend		Community	
	Coeff. (1)	M. Eff. (2)	Coeff. (3)	M. Eff. (4)	Coeff. (5)	M. Eff. (6)	Coeff. (7)	M. Eff. (8)
$I_{SameClub}$	0.288***	0.00417***	0.392***	0.0172***	-0.388*	-0.144*	-0.285***	-0.109***
	(0.0796)	(0.0015)	(0.1070)	(0.0066)	(0.2080)	(0.0779)	(0.0890)	(0.0341)
Const.	-2.727***		-2.308***		-0.174		0.433***	
	(0.0420)	(0.0329)	(0.164)	(0.0741)				
# Obs	11,854		1,957		180		1,374	

Notes: All specifications are probit regressions with standard errors clustered at the subject level. Odd numbered columns report coefficients, while even numbered columns report marginal effects. Dependent variables are: adding a friendship (columns 1-4), dropping a friendship (columns 5-6), or ceasing to be in the same network community (columns 7-8). For columns 1-2 the sample includes all subjects who are in the same dorm (but not currently a friend). For columns 3-4 the sample includes all subjects who are in the same network community (but not currently a friend). For Panel A the independent variable is an indicator variable that equals one if the two students share the same major, while for Panel B it is an indicator variable that equals one if the two students participate in the same club.

One potential reason why behavioral and preference differences do not play a larger role in friendship dynamics is that subjects had quite inaccurate beliefs about their friends' behavior. Subjects accurately predicted the average choices of their friends 39% of the time for lottery choices, 43% of the time for time preference choices, and 46% of the time for token allocations. Inaccurate beliefs were primarily driven by subjects overestimating how similar their choices were to their friends. The bias was largest for risk preferences, where 70% of subjects guessed that they made the same lottery choices as their friends, while only 27% of them did. Similarly,

74% of subjects guessed that they made the same payment timing choices as their friends, compared to 37% that actually did so. In both cases belief accuracy was significantly lower for those who guessed their friends were the same ($p < 0.01$ in both cases). Subjects also believed they made similar generosity choices as their friends, but they were more correct in doing so: 74% guessed they made the same choices, compared to 62% who actually did so.

This may suggest that the observed homophily for social preferences is primarily an unconscious behavior. Subjects believe that their friends are similar on all three preference dimensions, and do not recognize the differences with their friends for risk tolerance and patience. Explicitly sorting based on preferences would require individuals to be able to accurately infer the preferences of others, which our data suggests they cannot do very well. We note that Leider et al (2010) also found that subjects had somewhat inaccurate beliefs about their friends. They found that although subjects were approximately accurate in predicting how much more their friends would allocate to them in a dictator game, they were completely unable to identify which friends would be relatively more or less generous.

However, we can show that the sorting behavior we observe for generosity is not simply a side effect of sharing the same major or extracurricular activity. Table 3.7 reports the results from regressing network changes on differences in generosity while also controlling for shared major or club. Our results are largely unchanged, which suggests that other factors, such as direct interpersonal kindness, must be the driver of sorting on generosity.

Network centrality is another natural characteristic to explain the changes in the observed social network throughout the study. The network formation literature (see de Solla Price (1976) and Barabasi and Albert (1999)) suggests that if individuals choose whether to form and maintain friendships based on the instrumentality of the relationship then nodes that are central in the network will be particularly desirable.[10] We examine two measures of network centrality. Eigenvector centrality (C_{Eigen}) uses the eigenvector for the principle eigenvalue of

---

[10]de Solla Price (1976) proposed this mechanism as cumulative advantage, known today more commonly as preferential attachment (a term introduced by Barabasi and Albert (1999)). The preferential attachment mechanism assumes that a new node prefers to connect to existing nodes with more links (a larger degree). This generates a rich-get-richer effect as existing nodes with high degrees gain more links faster than nodes with low degrees. The advantage of the preferential attachment model is that it can reproduce networks with commonly observed power-law degree distributions.

Table 3.7: Network Dynamics for Generosity controlling for Major/Club

Panel A: Same Major

	Add Friend				Drop Relationship			
	From Dorm		From Community		Friend		Community	
	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ_{Kept}	-0.0117***	-0.0001***	-0.0119***	-0.0005***	0.0132**	0.0048**	0.0036	0.0010
	(0.0031)	(0.0000)	(0.0042)	(0.0002)	(0.0051)	(0.0018)	(0.0024)	(0.0009)
$I_{\text{SameMajor}}$	0.362***		0.369***		-0.605**		-0.291**	
	(0.108)		(0.149)		(0.294)		(0.135)	
Const.	-2.480***		-1.934***		-0.591***		0.315***	
	(0.0731)		(0.0940)		(0.153)		(0.0842)	
# Obs	11,854		1,957		180		1,374	

Panel B: Same Club

	Add Friend				Drop Relationship			
	From Dorm		From Community		Friend		Community	
	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ_{Kept}	-0.0112***	-0.0001***	-0.0114***	-0.0005***	0.0121**	0.0044**	0.0023	0.0009
	(0.0032)	(0.0000)	(0.0042)	(0.0002)	(0.0052)	(0.0019)	(0.0023)	(0.0009)
I_{SameClub}	0.298***		0.239*		-0.438**		-0.287***	
	(0.095)		(0.131)		(0.211)		(0.089)	
Const.	-2.537***		-1.967***		-0.412**		0.391***	
	(0.0420)	(0.0329)	(0.164)	(0.0741)				
# Obs	11,854		1,957		180		1,374	

Notes: All specifications are probit regressions with standard errors clustered at the subject level. Odd numbered columns report coefficients, while even numbered columns report marginal effects. Dependent variables are: adding a friendship (columns 1-4), dropping a friendship (columns 5-6), or ceasing to be in the same network community (columns 7-8). For columns 1-2 the sample includes all subjects who are in the same dorm (but not currently a friend). For columns 3-4 the sample includes all subjects who are in the same network community (but not currently a friend). The first independent variable is the absolute difference in the number of tokens kept in the allocation decision. For Panel A the second independent variable is an indicator variable that equals one if the two students share the same major, while for Panel B it is an indicator variable that equals one if the two students participate in the same club.

the adjacency matrix as the measure of centrality. With this measure, an individual is assigned more centrality if they are connected to others who are themselves highly central. Betweenness centrality (C_{Between}) counts the number of shortest paths that pass through the individual. For betweenness centrality, an individual is assigned more centrality if they help connect many other people together.

We use the observed social network in Period 1, calculate the centrality for each individual, and then examine whether the changes in the network between Period 1 and Period 3 are driven by the centrality of the individuals. We estimate the previous relationship change regressions using the centrality measures as the independent variables:

$$\text{prob}(\text{Change}_{ij}) = \Phi(\alpha + \beta C_X + \epsilon) \tag{3.3}$$

In these regressions, if individuals prefer to form and maintain friendships with highly central individuals, we would expect a positive β for additions and a negative for drops. The results are reported in Table 3.8.

Our results are broadly consistent with a preference to form and maintain relationships with central individual, with similar results for both measures. In particular, the preference for centrality seems to be more important for maintaining relationships than in forming new ones. This difference could be because it is hard to know the network position of people you are not already socially close with. Difficulty knowing the network position of potential friends could help explain why we find an effect for adding from the community compared to adding from the dorm. Subjects may be more likely to know the network position of people who are already in their social environment.

The effect of network centrality on friendship dynamics appears to be somewhat smaller than the effect of homophily. A one standard deviation increase in the network centrality of a potential friend from the network community increases the likelihood of forming a friendship by .014 percentage points (approximately one-tenth of the base rate) for eigenvector centrality, and .024 percentage points (approximately one-fifth of the base rate) for betweenness centrality. Recall that the effect of decreased similarity in generosity was about half the base rate.

Table 3.8: Network Dynamics for Network Centrality

Panel A: Eigenvector Centrality

	Add Friend				Drop Relationship			
	From Dorm		From Community		Friend		Community	
	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
C_{Eigen}	0.203	0.002	1.578***	0.006***	-3.782***	-1.470***	-3.786***	-0.679***
	(0.499)	(0.004)	(0.480)	(0.002)	(1.002)	(0.389)	(0.446)	(0.084)
Const.	-2.767***		-3.054***		0.252***		1.280***	
	(0.0371)		(0.0308)		(0.0376)		(0.0162)	
# Obs	31,887		200,168		3,975		201,610	

Panel B: Betweenness Centrality

	Add Friend				Drop Relationship			
	From Dorm		From Community		Friend		Community	
	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.	Coeff.	M. Eff.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
C_{Between}	4.244	0.0370	11.68***	0.0444***	-3.976	-1.545	-21.13***	-3.780***
	(5.142)	(0.045)	(2.856)	(0.011)	(3.664)	(1.423)	(0.792)	(0.187)
Const.	-2.810***		-3.080***		0.241***		1.319***	
	(0.0693)		(0.0327)		(0.0399)		(0.0159)	
# Obs	31,887		200,168		3,975		201,610	

Notes: All specifications are probit regressions with standard errors clustered at the subject level. Odd numbered columns report coefficients, while even numbered columns report marginal effects. Dependent variables are: adding a friendship (columns 1-4), dropping a friendship (columns 5-6), or ceasing to be in the same network community (columns 7-8). For columns 1-2 the sample includes all subjects who are in the same dorm (but not currently a friend). For columns 3-4 the sample includes all subjects who are in the same network community (but not currently a friend). For Panel A the independent variable is an indicator variable that equals one if the two students share the same major, while for Panel B it is an indicator variable that equals one if the two students participate in the same club.

Similarly, a one standard deviation increase in eigenvector centrality decreases the likelihood of dropping a friendship by 5 percentage points (approximately one-twelfth the base rate), and a one standard deviation increase in betweenness centrality decreases the likelihood by 0.9 percentage points (1.5% of the base rate). By contrast, the effect of decreasing similarity was one-fifth the base rate.

3.3.4   Influence Effects on Preferences

We now turn to our second research question: Are an individual's economic preferences influenced by the social context the individual experiences? Specifically, is an individual's risk, time or social preferences correlated with the average preferences of his or her friends (or broader social community). To measure this effect, we use the following specification:

$$\text{Preference}_{i,t} = \alpha + \beta\text{Preference}_{-i,t-\Delta t} + \epsilon \tag{3.4}$$

Preference$_{i,t}$ denotes the choice in the preference elicitation task (or the self-reported preference measure) for individual i in Phase t. Preference$_{-i,t-\Delta t}$ is the average choice (or self-report) for the individuals friends (or network community) in Phase t-Δt. The error term ϵ is clustered at the subject level, allowing for arbitrary correlations in the choices/reports of the focal individual. As before we also look at the network community as a broader measure of the individual's social context. The network community specifications also have a somewhat larger sample size, as some subjects did not have direct friends who also took the behavioral survey.

We look at both the contemporaneous correlation ($\Delta t = 0$; e.g. April preferences as a function of April friends' preferences), as well as the correlation with the individual's friends from earlier in the year ($\Delta t = 1$ or 2; e.g. April preferences as a function of January or October friends' preferences). We are interested in looking at the effect of friends at various different time lags for several reasons. First, if we focused on the contemporaneous correlation we might worry that it was due simply to common shocks or contextual factors. However, it is unlikely that I would experience the same shock today that my friends experienced three months ago. Second, it is not immediately clear over what time scale we should expect to see influence

effects. An individual's current social context may be the most prominent, on the other hand it may take an extended amount of time and/or sustained exposure for influence to occur. On the other hand, the shared experiences with friends from too long ago may have faded in memory or influence. Third, seeing effects at multiple lags would actually be the most encouraging result as it would be the clearest and most robust evidence for influence. Seeing a significant effect only with a one period lag, for example, could represent real influence that is highly time sensitive, or it could be a spurious correlation. Consistent correlations across multiple time periods are less likely to be spurious.

As an additional robustness check, we also consider specifications using the focal individual's previous choices and reports:

$$\text{Preference}_{i,t} = \alpha + \beta\text{Preference}_{-i,t-\Delta t} + \gamma\text{Preference}_{i,t-1} + \epsilon \tag{3.5}$$

$$\text{Preference}_{i,t} = \alpha + \beta\text{Preference}_{-i,t-\Delta t} + \gamma\text{Preference}_{i,t-2} + \epsilon \tag{3.6}$$

This provides a more direct control of the individual's "initial" preference than just the clustered standard errors. However, the structure of our data does give us a smaller sample size for many specifications.

One caveat to keep in mind is that we are using the individual's measured social network in the reference period this means that in the contemporaneous influence specification, for example, the set of friends that make up Preference-i is changing over the course of the experiment. If there are strong homophily effects in the social network dynamics this could drive positive correlation in later periods. We have several ways of addressing this problem. First, we cluster the errors at the subject level, so if the subject's preferences are primarily fixed (or evolving for non-influence reasons) the repeated measures should help account for this. Additionally, for risk and time preferences the changing social network is unlikely to be a problem in this respect, as we have already demonstrated that there is no significant homophily on these dimensions. For generosity this could be a problem for $\Delta t = 0$ or 1. However, this is unlikely to be driving results for $\Delta t = 2$, which uses the friends from the social network measured in

October to predict behavior in April. At this point there has been very little time for homophily to shape the social network, so a homophily-driven correlation is unlikely. Finally, if the observed correlation is primarily driven by homophily changing the network, we would expect the correlation to strengthen as the lag shrinks from 2 to 0 and there is more time for homophily to shape the network.

Table 3.9 reports the results of regressing our measures of risk preferences on the average measure for friends and social communities. We find essentially zero correlation for the incentivized elicitation task, however we do find a significant positive correlation for five of the six specifications using the self-reported measure (and a positive but insignificant result for the sixth). Additionally, we do not find that the magnitude of the effect is systematically growing as we go from $\Delta t = 2$ to $\Delta t = 0$, suggesting that this is not simply a reflection of underlying homophily. As a demonstration of the magnitude of the influence effect, a one standard deviation increase in the average self-reported risk measure for an individual's friends is predicted to increase that individual's self report in the same time period by 0.25 categories (1/8th of a standard deviation). We find similar sized effects for the influence of the broader social community, with a one standard deviation increase the average self-report for an individual's community corresponding to increase in the individual's report of 0.21 categories (1/10th of a standard deviation). This suggests that both an individual's immediate friends and the larger social context can have a significant influence on risk attitudes.

Table 3.10 reports the corresponding specifications with once- and twice-lagged dependent variables included as additional controls. We find results that are largely consistent with our main specification, albeit with a reduction in power, likely due to a smaller sample size. All of the estimated coefficients for the self-reported measure are positive, with many of them remaining statistically significant.

Table 3.11 reports the results of the time preference regressions. For this measure we find similar, albeit slightly weaker evidence, for social influence using the incentivized behavioral measure. We find a strongly significant result for the contemporaneous friends regression, and similar magnitudes but weaker significance for the lagged regressions. For the commu-

Table 3.9: Influence Effects for Risk Measures

Panel A: Behavioral Measures

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Risk}_{-i,t-\Delta t}$	0.0083	-0.0506	0.0298	-0.0853	0.11	0.0403
	(0.0668)	(0.115)	(0.123)	(0.088)	(0.122)	(0.198)
Constant	8.381***	8.794***	7.925***	9.195***	7.503***	7.860***
	(0.629)	(1.081)	(1.099)	(0.768)	(1.084)	(1.736)
# Obs	380	139	97	648	306	197

Panel B: Self-Reported Measures

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Risk}_{-i,t-\Delta t}$	0.148**	0.322***	0.151	0.202**	0.212*	0.298**
	(0.067)	(0.113)	(0.125)	(0.087)	(0.124)	(0.133)
Constant	4.874***	4.123***	4.918***	4.592***	4.499***	4.096***
	(0.431)	(0.745)	(0.780)	(0.531)	(0.773)	(0.807)
# Obs	380	139	97	648	306	197

Notes: All specifications are OLS regressions with standard errors clustered at the subject level. The dependent measure is the number of safe choices (Panel A) or self-reported risk measure (Panel B) of the focal individual. The independent measure is the average choice/report of the focal individual's friends (columns 1-3) or network community (columns 4-6). Columns 1 and 4 report the contemporaneous correlation ($\Delta t = 0$), columns 2 and 5 use the one-period lagged friend/community average ($\Delta t = 1$), while Columns 3 and 6 use the two-period lagged average ($\Delta t = 2$).

Table 3.10: Influence Effects for Risk Measures with Lagged Choices

Panel A: Behavioral Measures, Once-Lagged Choices

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Risk}_{-i,t-\Delta t}$	0.0252	-0.0206	0.182	-0.029	0.0985	0.285*
	(0.092)	(0.085)	(0.181)	(0.098)	(0.095)	(0.163)
$\text{Risk}_{i,t-1}$	0.485***	0.484***	0.588***	0.588***	0.587***	0.633***
	(0.111)	(0.112)	(0.151)	(0.063)	(0.063)	(0.068)
Constant	3.969***	4.371***	1.423	3.551***	2.465**	0.371
	(1.427)	(1.340)	(2.093)	(1.117)	(0.978)	(1.651)
# Obs	139	139	55	306	306	141

Panel B: Behavioral Measures, Twice-Lagged Choices

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Risk}_{-i,t-\Delta t}$	-0.186	0.257	0.011	-0.324***	0.252**	0.148
	(0.125)	(0.163)	(0.101)	(0.110)	(0.121)	(0.176)
$\text{Risk}_{i,t-2}$	0.385***	0.388***	0.388***	0.470***	0.534***	0.489***
	(0.121)	(0.143)	(0.122)	(0.077)	(0.086)	(0.077)
Constant	6.429***	2.421	4.754***	6.795***	1.406	2.691
	(1.729)	(1.683)	(1.455)	(1.189)	(1.227)	(1.669)
# Obs	97	55	97	197	141	197

Panel C: Self-Reported Measures, Once-Lagged Choices

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Risk}_{-i,t-\Delta t}$	0.071	0.213**	0.128	0.041	0.089	0.374***
	(0.097)	(0.082)	(0.178)	(0.101)	(0.095)	(0.133)
$\text{Risk}_{i,t-1}$	0.673***	0.658***	0.638***	0.665***	0.663***	0.698***
	(0.084)	(0.080)	(0.109)	(0.055)	(0.054)	(0.057)
Constant	1.596**	0.874	1.177	1.688**	1.425**	-0.467
	(0.719)	(0.661)	(1.425)	(0.671)	(0.673)	(0.747)
# Obs	139	139	55	306	306	141

Panel D: Self-Reported Measures, Twice-Lagged Choices

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Risk}_{-i,t-\Delta t}$	0.198**	0.252*	0.0290	0.209**	0.163	0.0850
	(0.0978)	(0.144)	(0.106)	(0.103)	(0.131)	(0.111)
$\text{Risk}_{i,t-2}$	0.553***	0.557***	0.557***	0.617***	0.701***	0.613***
	(0.0829)	(0.149)	(0.0904)	(0.0594)	(0.0712)	(0.0621)
Constant	1.407*	1.211	2.397***	1.045*	0.807	1.790**
	(0.774)	(0.962)	(0.785)	(0.586)	(0.822)	(0.688)
# Obs	97	55	97	197	141	197

Notes: All specifications are OLS regressions with standard errors clustered at the subject level. The dependent measure is the number of safe choices (Panels A and B) or self-reported risk measure (Panels C and D) of the focal individual. The first independent measure is the average choice/report of the focal individual's friends (columns 1-3) or network community (columns 4-6). Columns 1 and 4 report the contemporaneous correlation ($\Delta t = 0$), columns 2 and 5 use the one-period lagged friend/community average ($\Delta t = 1$), while Columns 3 and 6 use the two-period lagged average ($\Delta t = 2$). Panels A and C also include the focal individuals' choice/report from the previous period, while Panels B and D include the focal individuals' choice/report from two periods ago.

nity regressions we see marginal significantly only in the contemporaneous regression, and no significance for the lagged regressions. The magnitude of the contemporaneous effect is similar to the risk preference effect: a one standard deviation increase in the average number of sooner choice by an individual's friends increases the average number of the individual's sooner choices by 0.48 (an increase in the indifference point of $2.39, equal to 1/7th a standard deviation). A one standard deviation increase in the community average would increase the number of sooner choices by 0.27 (a $1.34 increase in the indifference point, equal to 1/12th a standard deviation). We find no corresponding influence effect for the self-reported measure of patience.

Table 3.11: Influence Effects for Patience Measures

Panel A: Behavioral Measures

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
Time$_{-i,t-\Delta t}$	0.168***	0.242*	0.202*	0.159*	0.035	0.089
	(0.062)	(0.125)	(0.105)	(0.084)	(0.113)	(0.126)
Constant	2.623***	2.758***	2.495***	2.733***	2.954***	3.161***
	(0.281)	(0.499)	(0.437)	(0.313)	(0.401)	(0.456)
# Obs	380	139	97	648	306	197

Panel B: Self-Reported Measures

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
Time$_{-i,t-\Delta t}$	0.099	0.127	0.116	-0.163**	0.147	-0.048
	(0.065)	(0.108)	(0.126)	(0.081)	(0.134)	(0.166)
Constant	5.406***	5.359***	5.365***	6.972***	5.220***	6.332***
	(0.414)	(0.708)	(0.767)	(0.489)	(0.797)	(0.996)
# Obs	362	137	95	631	306	197

Notes: All specifications are OLS regressions with standard errors clustered at the subject level. The dependent measure is the number of sooner choices (Panel A) or self-reported patience measure (Panel B) of the focal individual. The independent measure is the average choice/report of the focal individual's friends (columns 1-3) or network community (columns 4-6). Columns 1 and 4 report the contemporaneous correlation ($\Delta t = 0$), columns 2 and 5 use the one-period lagged friend/community average ($\Delta t = 1$), while Columns 3 and 6 use the two-period lagged average ($\Delta t = 2$).

As before, we also consider specifications with lagged dependent variables as a robustness check. These results are reported in Table 3.12. Our previous results for patience appear to be less robust to this alternate specification than the results for risk tolerance. None of the coefficients for the self-reported measure remain significant, and many of them are smaller

in magnitude. However, we again note that this robustness check comes with a significant reduction in our sample size.

We report the results for generosity in Table 3.13. Here we actually see a negative influence effect, with individuals that are part of generous social communities becoming significantly more selfish. Specifically, if the average generosity of an individual's social community increases by one standard deviation, the results predict the individual's own generosity to decrease by 1.9 tokens (or $2.85 for the recipient, equal to 1/11 a standard deviation). The results for friends also have a negative sign, however the effect size is small and the coefficients are not close to significance. While we did not anticipate this reverse-influence effect, we do note that Ahern et al. (2013) also found negative peer effects for generosity. It is possible that this is a form of free-riding, where individuals attempt to benefit from the generosity of their social context. We also note that this free-riding has limits, since if an individual becomes too dissimilar from their friends and social community the previously demonstrated homophily effect will increase the chances that they are cut off from the network.

We report the robustness specifications with lagged choices in Table 3.14 For both alternate specifications we again find evidence for the negative influence effect of an individual's community, with similar magnitudes and levels of significance. We also now find some significant negative effects for the regressions using friends' choices. While we had not anticipated finding this free-riding result, the effect appears to be quite robust.

3.4   Conclusion

Using a longitudinal design to follow freshman during their first year at university, we test for and demonstrate selection based on preferences and dynamic preference formation for three key and fundamental economic preferences: social risk and time preferences. We use a longitudinal design, in which we follow incoming freshman through their first academic year at their university. Subjects participate in three waves of an online experiment where we elicit their social network using an incentive compatible mechanism and then measure participants' levels of risk attitudes, time preferences and altruism using economic games.

Table 3.12: Influence Effects for Patience Measures with Lagged Choices

Panel A: Behavioral Measures, Once-Lagged Choices

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Time}_{-i,t-\Delta t}$	0.0929	0.127	0.130	0.0481	-0.0608	0.0864
	(0.121)	(0.131)	(0.126)	(0.0877)	(0.0843)	(0.115)
$\text{Time}_{i,t-1}$	0.481***	0.470***	0.519***	0.528***	0.533***	0.626***
	(0.105)	(0.110)	(0.129)	(0.0695)	(0.0698)	(0.0788)
Constant	1.680***	1.636***	0.958	1.347***	1.678***	0.881*
	(0.456)	(0.419)	(0.745)	(0.347)	(0.310)	(0.461)
# Obs	139	139	55	306	306	141

Panel B: Behavioral Measures, Twice-Lagged Choices

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Time}_{-i,t-\Delta t}$	0.101	0.0178	0.0644	-0.0311	-0.0100	-0.101
	(0.0907)	(0.180)	(0.106)	(0.109)	(0.129)	(0.0911)
$\text{Time}_{i,t-2}$	0.467***	0.441***	0.459***	0.609***	0.570***	0.614***
	(0.106)	(0.147)	(0.114)	(0.0711)	(0.0851)	(0.0700)
Constant	1.454***	1.808**	1.629***	1.692***	1.529***	1.895***
	(0.514)	(0.716)	(0.452)	(0.436)	(0.488)	(0.404)
# Obs	97	55	97	197	141	197

Panel C: Self-Reported Measures, Once-Lagged Choices

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Time}_{-i,t-\Delta t}$	0.0002	-0.0294	0.144	-0.125	0.258**	-0.109
	(0.0702)	(0.0799)	(0.117)	(0.102)	(0.115)	(0.181)
$\text{Time}_{i,t-1}$	0.647***	0.653***	0.594***	0.535***	0.554***	0.518***
	(0.0686)	(0.0670)	(0.120)	(0.0652)	(0.0605)	(0.0790)
Constant	2.249***	2.415***	1.653	3.671***	1.246*	3.552***
	(0.615)	(0.689)	(0.993)	(0.865)	(0.645)	(1.148)
# Obs	137	135	54	303	303	141

Panel D: Self-Reported Measures, Twice-Lagged Choices

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Time}_{-i,t-\Delta t}$	0.0132	0.120	0.0144	-0.0724	0.158	0.00307
	(0.0925)	(0.147)	(0.125)	(0.128)	(0.107)	(0.135)
$\text{Time}_{i,t-2}$	0.406***	0.377**	0.408***	0.483***	0.524***	0.484***
	(0.100)	(0.147)	(0.102)	(0.067)	(0.075)	(0.067)
Constant	3.628***	3.132**	3.648***	3.603***	2.032**	3.139***
	(0.797)	(1.229)	(0.879)	(0.892)	(0.867)	(0.859)
# Obs	95	55	93	191	139	191

Notes: All specifications are OLS regressions with standard errors clustered at the subject level. The dependent measure is the number of sooner choices (Panels A and B) or self-reported patience measure (Panels C and D) of the focal individual. The first independent measure is the average choice/report of the focal individual's friends (columns 1-3) or network community (columns 4-6). Columns 1 and 4 report the contemporaneous correlation ($\Delta t = 0$), columns 2 and 5 use the one-period lagged friend/community average ($\Delta t = 1$), while Columns 3 and 6 use the two-period lagged average ($\Delta t = 2$). Panels A and C also include the focal individuals' choice/report from the previous period, while Panels B and D include the focal individuals' choice/report from two periods ago.

Table 3.13: Influence Effects for Generosity Measures

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Kept}_{-i,t-\Delta t}$	-0.0553	-0.0295	-0.0497	-0.172**	-0.0611	-0.463**
	(0.0617)	(0.103)	(0.121)	(0.0794)	(0.149)	(0.222)
Constant	63.41***	62.63***	63.53***	69.92***	63.92***	87.10***
	(3.854)	(6.100)	(7.708)	(4.644)	(9.032)	(13.47)
# Obs	380	139	97	648	306	197

Notes: All specifications are OLS regressions with standard errors clustered at the subject level. The dependent measure is the number of tokens kept by the focal individual. The independent measure is the average choice of the focal individual's friends (columns 1-3) or network community (columns 4-6). Columns 1 and 4 report the contemporaneous correlation ($\Delta t = 0$), columns 2 and 5 use the one-period lagged friend/community average ($\Delta t = 1$), while Columns 3 and 6 use the two-period lagged average ($\Delta t = 2$).

Table 3.14: Influence Effects for Generosity Measures with Lagged Choices

Panel A: Once-Lagged Choices

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Kept}_{-i,t-\Delta t}$	-0.253***	-0.0562	-0.173	-0.182**	-0.0625	-0.499**
	(0.0942)	(0.101)	(0.137)	(0.0799)	(0.112)	(0.193)
$\text{Kept}_{i,t-1}$	0.480***	0.469***	0.541***	0.604***	0.613***	0.654***
	(0.114)	(0.118)	(0.165)	(0.081)	(0.082)	(0.111)
Constant	45.94***	35.61***	41.07***	35.02***	27.39***	50.90***
	(9.939)	(10.25)	(14.09)	(7.336)	(9.141)	(14.57)
# Obs	139	139	55	306	306	141

Panel B: Twice-Lagged Choices

	-i = Avg of Friends			-i = Avg of Community		
	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 0$	$\Delta t = 1$	$\Delta t = 2$
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Kept}_{-i,t-\Delta t}$	-0.0673	-0.367*	-0.0243	-0.256**	-0.148	-0.438**
	(0.120)	(0.200)	(0.133)	(0.117)	(0.165)	(0.203)
$\text{Kept}_{i,t-2}$	0.432***	0.373**	0.437***	0.513***	0.582***	0.517***
	(0.139)	(0.142)	(0.140)	(0.0949)	(0.109)	(0.0978)
Constant	38.92***	64.36***	36.09**	45.25***	35.70***	55.77***
	(13.06)	(15.95)	(14.37)	(9.945)	(13.56)	(13.83)
# Obs	97	55	97	197	141	197

Notes: All specifications are OLS regressions with standard errors clustered at the subject level. The dependent measure is the number of tokens kept by the focal individual. The independent measure is the average choice of the focal individual's friends (columns 1-3) or network community (columns 4-6). Columns 1 and 4 report the contemporaneous correlation ($\Delta t = 0$), columns 2 and 5 use the one-period lagged friend/community average ($\Delta t = 1$), while Columns 3 and 6 use the two-period lagged average ($\Delta t = 2$). Panel A also includes the focal individuals' choice from the previous period, while Panels B includes the focal individuals' choice from two periods ago.

We find evidence for each mechanism on at least one important preference. We show that changes in subject's social networks are significantly influenced the similarity or dissimilarity in generosity (as measured by a standard lab-style dictator task). Students who are similarly generous are more likely to become friends, and less likely to end a pre-existing friendship. The effects of homophily for altruism are of a similar magnitude to other key predictors of network changes, such as participating in the same activities, or the network centrality of the (potential) friend. Additionally, we find evidence for peer influence for both self-reported tolerance for risk and our incentivized measure of patience. The influence effect is robust to looking at both contemporaneous and lagged behavior in the network, and is of substantial magnitude: a one standard deviation change in the average behavior of others is associated with a 1/7th to 1/12th standard deviation change in the corresponding behavior of the focal individual. We additionally find surprising evidence for a negative peer effect on generosity, consistent with free-riding behavior.

Our work advances both the understanding of how a student's social network and preferences evolve upon entering college, and how a student's social network affects one's economic preferences. We focused on studying students during their freshman year, as it is a time of substantial personal and social change. This gave us the best chance to observe homophily and peer influence on preferences. However, one potential concern is that this might represent the high point for these mechanisms. Whether these mechanisms continue to be important later in life is an open question for future research.

Bibliography

Abdulkadiroglu, A., Angrist, J., and Pathak, P. (2014). The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica*, 82(1):137–196.

Ahern, K. R., Duchin, R., and Shumway, T. (2014). Peer effects in risk aversion and trust. *Review of Financial Studies*, page hhu042.

Akerlof, G. A. and Kranton, R. E. (2005). Identity and the economics of organizations. *The Journal of Economic Perspectives*, 19(1):9–32.

Almlund, M., Duckworth, A. L., Heckman, J. J., and Kautz, T. D. (2011). Personality psychology and economics. Technical report, National Bureau of Economic Research.

Andreoni, J. and Scholz, J. K. (1998). An econometric analysis of charitable giving with interdependent preferences. *Economic inquiry*, 36(3):410–428.

Aral, S. and Walker, D. (2011a). Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Management science*, 57(9):1623–1639.

Aral, S. and Walker, D. (2011b). Identifying social influence in networks using randomized experiments. *IEEE Intelligent Systems*, 26(5):91–96.

Aral, S. and Walker, D. (2014). Tie strength, embeddedness, and social influence: A large-scale networked experiment. *Management Science*, 60(6):1352–1370.

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

Backstrom, L., Huttenlocher, D., Kleinberg, J., and Lan, X. (2006). Group formation in large social networks: membership, growth, and evolution. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 44–54. ACM.

Bakshy, E., Rosenn, I., Marlow, C., and Adamic, L. (2012). The role of social networks in information diffusion. In *Proceedings of the 21st international conference on World Wide Web*, pages 519–528. ACM.

Bandiera, O., Barankay, I., and Rasul, I. (2010). Social incentives in the workplace. *The Review of Economic Studies*, 77(2):417–458.

Bandiera, O. and Rasul, I. (2006). Social networks and technology adoption in northern Mozambique. *The Economic Journal*, 116(514):869–902.

Bauman, K. E. and Ennett, S. T. (1996). On the importance of peer influence for adolescent drug use: Commonly neglected considerations. *Addiction*, 91(2):185–198.

Benz, M. and Meier, S. (2008). Do people behave in experiments as in the field?evidence from donations. *Experimental economics*, 11(3):268–281.

Bettinger, E. and Slonim, R. (2006). Using experimental economics to measure the effects of a natural educational experiment on altruism. *Journal of Public Economics*, 90(8):1625–1648.

Bettinger, E. and Slonim, R. (2007). Patience among children. *Journal of Public Economics*, 91(1):343–363.

Betts, J. R. and Zau, A. (2004). Peer groups and academic achievement: Panel evidence from administrative data. *Unpunlished Manuscript*.

Bobba, M. and Gignoux, J. (2016). Spillover effects and take-up of transfers in integrated social policies: Evidence from Progresa. *Working Paper*.

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E., and Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415):295–298.

Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, 150(1):41–55.

Brañas-Garza, P., Cobo-Reyes, R., Jiménez, N., and Ponti, G. (2013). A guided tour to (real-life) social network elicitation. *Working Paper*.

Bruun, J. and Brewe, E. (2013). Talking and learning physics: Predicting future grades from network measures and Force Concept Inventory pretest scores. *Physical Review Special Topics-Physics Education Research*, 9(2):020109.

Burke, M. A. and Sass, T. R. (2013). Classroom Peer Effects and Student Achievement. *Journal of Labor Economics*, 31(1):51–82.

Burks, S. V., Lewis, C., Kivi, P. A., Wiener, A., Anderson, J. E., Götte, L., DeYoung, C. G., and Rustichini, A. (2015). Cognitive skills, personality, and economic preferences in collegiate success. *Journal of Economic Behavior & Organization*, 115:30–44.

Cai, J., De Janvry, A., and Sadoulet, E. (2015). Social networks and the decision to insure. *American Economic Journal: Applied Economics*, 7(2):81–108.

Calvo-Armengo, A., Patacchini, E., and Zenou, Y. (2009). Peer effects and social networks in education. *Review of Economic Studies*, 76(4):1239–1267.

Calvo-Armengol, A. and Jackson, M. O. (2004). The effects of social networks on employment and inequality. *The American Economic Review*, 94(3):426–454.

Cardenas, J. C. and Carpenter, J. (2008). Behavioural development economics: Lessons from field labs in the developing world. *The Journal of Development Studies*, 44(3):311–338.

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

Carrell, S. E., Sacerdote, B., and West, J. E. (2013). From natural variation to optimal Policy? The Lucas Critique meets peer effects. *Econometrica*, 81(3):855–882.

Carrington, P. J., Scott, J., and Wasserman, S. (2005). *Models and methods in social network analysis*, volume 28. Cambridge university press.

Chandrasekhar, A. G. and Lewis, R. (2011). Economics of sampled networks. *Working Paper*.

Christakis, N. A. and Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England journal of medicine*, 357(4):370–379.

Christakis, N. A. and Fowler, J. H. (2009). *Connected: The surprising power of our social networks and how they shape our lives*. Little, Brown.

Cialdini, R. B. and Garde, N. (1987). *Influence*, volume 3. A. Michel.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American journal of sociology*, 94:S95–S120.

Conley, T. G. and Udry, C. R. (2010). Learning about a new technology: Pineapple in Ghana. *The American Economic Review*, 100(1):35–69.

Currarini, S., Jackson, M. O., and Pin, P. (2010). Identifying the roles of race-based choice and chance in high school friendship network formation. *Proceedings of the National Academy of Sciences*, 107(11):4857–4861.

Ding, W. and Lehrer, S. F. (2007). Do peers affect student achievement in china's secondary schools? *The Review of Economics and Statistics*, 89(2):300–312.

Dobbie, W. and Fryer, R. G. (2014). The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools. *American Economic Journal: Applied Economics*, 6(3):58–75.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2012). The intergenerational transmission of risk and trust attitudes. *The Review of Economic Studies*, 79(2):645–677.

Duckworth, A. L. and Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, 16(12):939–944.

Duckworth, A. L. and Seligman, M. E. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of educational psychology*, 98(1):198.

Duflo, E. and Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *The Quarterly Journal of Economics*, 118(3):815–842.

Durlauf, S. N. (2004). Neighborhood effects. *Handbook of regional and urban economics*, 4:2173–2242.

Fafchamps, M., Vaz, A., and Vicente, P. C. (2013). Voting and peer effects: Experimental evidence from Mozambique. *NOVAFRICA Working Paper*.

Fafchamps, M. and Vicente, P. C. (2013). Political violence and social networks: Experimental evidence from a Nigerian election. *Journal of Development Economics*, 101:27–48.

Fehr, E., Glätzle-Rützler, D., and Sutter, M. (2013). The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence. *European Economic Review*, 64:369–383.

Fehr, E. and Hoff, K. (2011). Introduction: Tastes, castes and culture: The influence of society on preferences. *The Economic Journal*, 121(556):F396–F412.

Fisher, L. A. and Bauman, K. E. (1988). Influence and selection in the friend-adolescent relationship: Findings from studies of adolescent smoking and drinking1. *Journal of Applied Social Psychology*, 18(4):289–314.

Foster, A. D. and Rosenzweig, M. R. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of political Economy*, 103(6):1176–1209.

Foster, G. (2006). It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism. *Journal of Public Economics*, 90(8):14551475.

Frank, K. A., Muller, C., and Mueller, A. S. (2013). The embeddedness of adolescent friendship nominations: The formation of social capital in emergent network structures 1. *American Journal of Sociology*, 119(1):216–253.

Glaeser, E. L. and Sacerdote, B. (1999). Why is there more crime in cities? *Journal of political economy*, 107(S6):S225–S258.

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 91(3):481–510.

Greenwald, A. G., Carnot, C. G., Beach, R., and Young, B. (1987). Increasing voting behavior by asking people if they expect to vote. *Journal of Applied Psychology*, 72(2):315.

Guryan, J., Kroft, K., and Notowidigdo, M. J. (2009). Peer Effects in the Workplace" Evidence from Random Groupings in Professional Tournaments. *American Economic Journal: Applied Economics*, 1(4):34–68.

Hahn, J., Hausman, J., and Kuersteiner, G. (2004). Estimation with weak instruments: Accuracy of higher-order bias and mse approximations. *The Econometrics Journal*, 7(1):272–306.

Han, L. and Li, T. (2009). The gender difference of peer influence in higher education. *Economics of Education review*, 28(1):129–134.

Hanushek, E. A., Kain, J. F., Markman, J. M., and Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18(5):527–544.

Harrison, G. W. and List, J. A. (2004). Field experiments. *Journal of Economic Literature*, 42(4):1009–1055.

Hill, A. J. (2015). The girl next door: The effect of opposite gender friends on high school achievement. *American Economic Journal: Applied Economics*, 7(3):147–177.

Hoxby, C. and Weingarth, G. (2005). Taking race out of the equation: School reassignment and the structure of peer effects. *Manuscript*.

Hoxby, C. M. (2000). Peer effects in the classroom: Learning from gender and race variation. *NBER Working Paper*, No. 7867.

Kandel, D. B. (1978). Homophily, selection, and socialization in adolescent friendships. *American journal of Sociology*, 84(2):427–436.

Karlan, D., Mobius, M., Rosenblat, T., and Szeidl, A. (2009). Trust and social collateral. *The Quarterly Journal of Economics*, 124(3):1307–1361.

Knack, S. and Keefer, P. (1997). Does social capital have an economic payoff? a cross-country investigation. *The Quarterly journal of economics*, 112(4):1251–1288.

Kossinets, G. and Watts, D. J. (2006). Empirical analysis of an evolving social network. *science*, 311(5757):88–90.

Leider, S., Möbius, M. M., Rosenblat, T., and Do, Q.-A. (2009). Directed altruism and enforced reciprocity in social networks. *The Quarterly Journal of Economics*, 124(4):1815–1851.

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

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

Marlow, C., Naaman, M., Boyd, D., and Davis, M. (2006). Ht06, tagging paper, taxonomy, flickr, academic article, to read. In *Proceedings of the seventeenth conference on Hypertext and hypermedia*, pages 31–40. ACM.

Marmaros, D. and Sacerdote, B. (2002). Peer and social networks in job search. *European Economic Review*, 46(4):870–879.

Maroulis, S. and Gomez, L. M. (2008). Does "connectedness" matter? Evidence from a social network analysis within a small-school reform. *Teachers College Record*, 110(9):1901–1929.

Mayer, A. and Puller, S. L. (2008). The old boy (and girl) network: Social network formation on university campuses. *Journal of public economics*, 92(1):329–347.

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1):415–444.

Meyer, C. and Waller, G. (2001). Social convergence of disturbed eating attitudes in young adult women. *The Journal of nervous and mental disease*, 189(2):114–119.

Miguel, E. and Kremer, M. (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1):159–217.

Milkman, K. L., Beshears, J., Choi, J. J., Laibson, D., and Madrian, B. C. (2011). Using implementation intentions prompts to enhance influenza vaccination rates. *Proceedings of the National Academy of Sciences*, 108(26):10415–10420.

Mobius, M. M., Niehaus, P., and Rosenblat, T. S. (2005). Social learning and consumer demand. *Working Paper*.

Morwitz, V. G., Johnson, E., and Schmittlein, D. (1993). Does measuring intent change behavior? *Journal of Consumer Research*, 20(1):46–61.

Nickerson, D. W. and Rogers, T. (2010). Do you have a voting plan? Implementation intentions, voter turnout, and organic plan making. *Psychological Science*, 21(2):194–199.

Ors, E., Palomino, F., and Peyrache, E. (2013). Performance gender gap: does competition matter? *Journal of Labor Economics*, 31(3):443–499.

Oster, E. and Thornton, R. (2012). Determinants of technology adoption: Peer effects in menstrual cup take-up. *Journal of the European Economic Association*, 10(6):1263–1293.

Pintrich, P. R. and De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1):33.

Puma, M. J. et al. (1993). Prospects: The congressionally mandated study of educational growth and opportunity. the interim report.

Rogers, T., Milkman, K. L., John, L. K., and Norton, M. I. (2015). Beyond good intentions: Prompting people to make plans improves follow-through on important tasks. *Behavioral Science & Policy*, 1(2):33–41.

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

Sacerdote, B. (2011). Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? *Handbook of the Economics of Education*, 3:249277.

Stevens, E. A. and Prinstein, M. J. (2005). Peer contagion of depressogenic attributional styles among adolescents: A longitudinal study. *Journal of abnormal child psychology*, 33(1):25–37.

Stinebrickner, R. and Stinebrickner, T. R. (2006). What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics*, 90(8):1435–1454.

Tajfel, H. and Turner, J. C. (1979). An integrative theory of intergroup conflict. *The Social Psychology of Intergroup Relations*, 33(47):74.

Topa, G. (2001). Social interactions, local spillovers and unemployment. *The Review of Economic Studies*, 68(2):261–295.

Wentzel, K. R. (1998). Social relationships and motivation in middle school: The role of parents, teachers, and peers. *Journal of Educational Psychology*, 90(2):202.

Xu, C. W. (2016). Peer effects in randomly assigned study groups. *Working Paper*.

Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics*, 85(1):9–23.
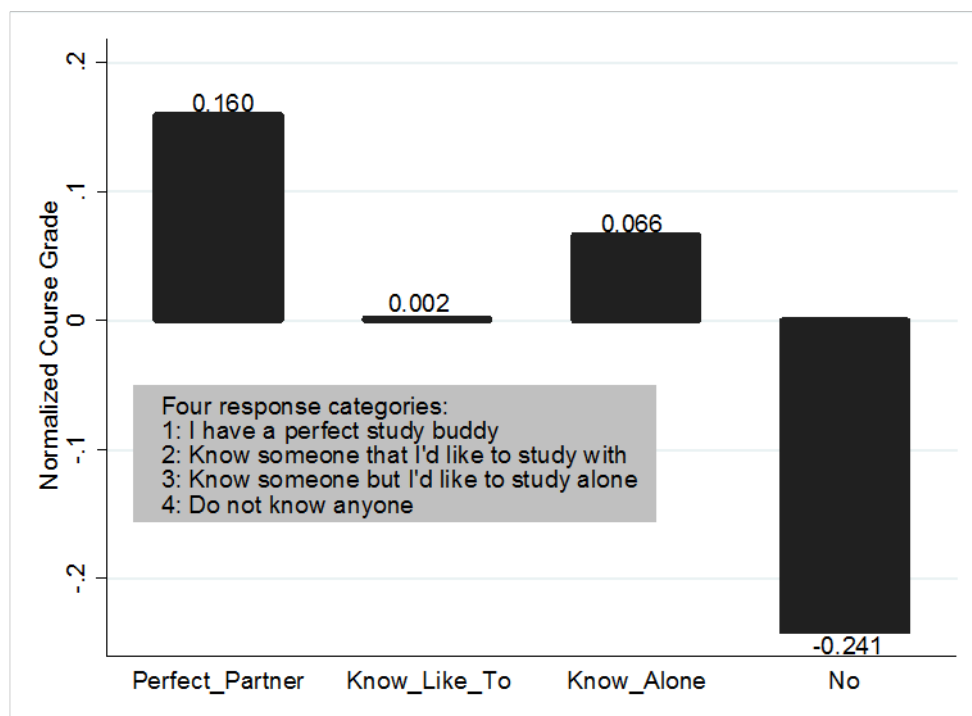
APPENDIX A

Appendix for Chapter One

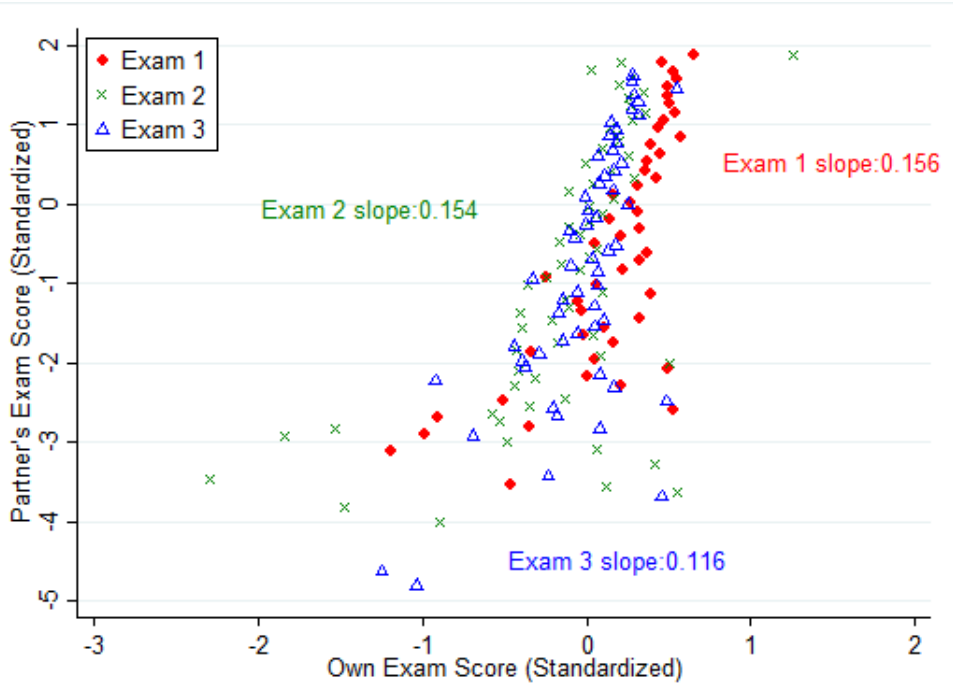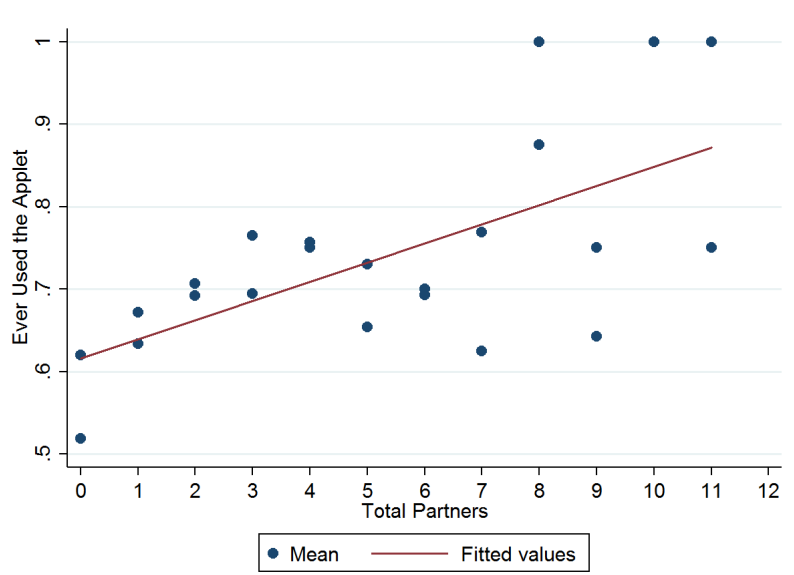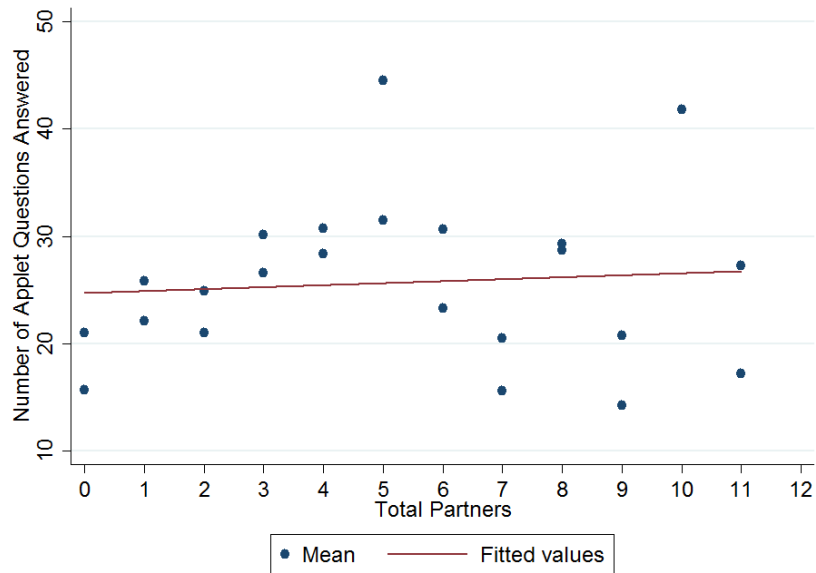Figure A.1: Positive Correlation btw. Study Buddy and Grades based on Past Feedback

Figure A.2: Score Correlation

(a) Likelihood of Using

(b) Intensity of Use

Figure A.3: Increasing Usage with More Total Partners

Table A.1: Comparing Respondents and Nonrespondents

	Respondents			Nonrespondents			T Test	
	Mean	S.D.	N	Mean	S.D.	N	Diff.	t.stat.
Female	0.583	0.493	1564	0.417	0.493	592	0.166***	6.95
White	0.613	0.487	1564	0.606	0.489	592	0.007	0.29
Black	0.025	0.156	1564	0.052	0.223	592	-0.027**	-3.21
Asian	0.221	0.415	1564	0.193	0.395	592	0.029	1.45
Hispanic	0.041	0.198	1564	0.051	0.220	592	-0.010	-0.99
In State	0.590	0.492	1564	0.600	0.490	592	-0.010	-0.43
Freshman	0.108	0.310	1564	0.102	0.303	592	0.006	0.39
Sophomore	0.474	0.499	1564	0.380	0.486	592	0.094***	3.93
Junior	0.304	0.460	1564	0.371	0.484	592	-0.067**	-2.97
Senior	0.113	0.317	1564	0.146	0.353	592	-0.032*	-2.05
HS GPA	3.833	0.167	1441	3.761	0.230	536	0.072***	7.65
College GPA	3.413	0.414	1538	3.148	0.529	577	0.266***	12.15
Withdraw	0.001	0.025	1564	0.046	0.209	592	-0.045***	-8.36
Final Grade	3.275	0.706	1562	2.621	1.055	564	0.655***	16.39

Notes: Nonrespondents refer to those students who did not respond to the network elicitation question after Exam 2.

Table A.2: Network Dynamics

Logit Source → Target	(1) Drop	(2) Is Dropped	(3) Add	(4) Is Added
Source is Treated	-0.086 (0.122)		-0.111 (0.290)	
Target is Treated		0.088 (0.104)		0.001 (0.301)
Obs.	2248	2248	47709	47709
Cluster	Source	Target	Source	Target

Notes: This table compare the two networks, one elicited after Exam 2 and the other elicited after Exam 3, in order to see what predicts the network dynamics.

Table A.3: Study Buddy "Or" Networks Summary Statistics

Post Exam Surveys	Exam 2	Exam 3
# Respondents (%)	1564 (73%)	1896 (88%)
% Female	0.58	0.55
% Listing a study or talking buddy	0.71	0.62
# Nodes	1604	1668
# Edges	2422	2558
avg. # buddies listed (out degree)	2.2	2.2
Unique pairs	1915	2005
% edges between female students	0.47	0.44
% edges between male students	0.25	0.28
% edges between male and female students	0.27	0.28
% edges reciprocated if both in survey	0.52	0.47

Table A.4: Node Position and Academic Performance

Centrality	OLS	(1) Exam 1 Score	(2) Exam 2 Score	(3) Exam 3 Score
	In-degree	0.149	0.602**	0.503**
		(0.181)	(0.260)	(0.212)
	Out-degree	0.200	0.854***	0.526***
		(0.145)	(0.210)	(0.186)
	Page Rank	601.8	2154***	1634***
		(384.5)	(500.8)	(484.7)
	Eigenvector	-1.388	2.947	3.088
		(2.748)	(2.933)	(2.287)
	Closeness	0.913	3.181***	2.962***
		(0.567)	(0.884)	(0.734)
	Betweenness	-828.8	-379.5	-443.8
		(839.6)	(1200.0)	(1009.0)

Notes: In each column, each row is a separate OLS regression between i's grade and a network position measure, controlling for the same set of control variables as used in all the other regressions. All the exam scores are standardized. In-degree is the number of students that list i as a study buddy. Out-degree is the number of students that i lists. Page rank is an algorithm to quantify the importance of a node. Eigenvector is another global measure that is calculated based on the network adjacency matrix. Closeness centrality is defined as the inverse of farness, which in turn, is the sum of distances to all other nodes. Finally betweenness centrality is equal to the number of shortest paths from all nodes to all others that pass through i.

Table A.5: Treatment and Spillover Effects on Applet Use (Using Fractions Treated)

Panel A: Ever Used	(1) LPM	(2) Logit	(3) Tobit	
T_i (β)	0.108**	0.101**	0.103***	
	(0.046)	(0.043)	(0.035)	
FriTreated%$_i$ (γ)	0.072	0.064	0.085***	
	(0.052)	(0.050)	(0.042)	
T_i*FriTreated%$_i$ (ρ)	-0.138**	-0.130**	-0.128***	
	(0.057)	(0.055)	(0.048)	
Control Mean	0.555	0.555	0.555	
Control S.D.	(0.498)	(0.498)	(0.498)	
p-value: $\gamma + \rho = 0$	0.043	0.048	0.133	
Obs.	3054	3054	3054	
Panel B: #Questions Answered	(1) OLS	(2) Tobit	(3) Poisson	(4) Negative Binomial
T_i (β)	8.050***	8.477***	9.688***	9.163***
	(2.592)	(2.842)	(3.243)	(3.404)
FriTreated%$_i$ (γ)	7.378*	6.948**	8.910**	7.236*
	(3.844)	(3.481)	(4.167)	(3.934)
T_i*FriTreated%$_i$ (ρ)	-10.13**	-10.53***	-11.90***	-10.33**
	(4.163)	(3.877)	(4.565)	(4.471)
Control Mean	16.83	16.83	16.83	16.83
Control S.D.	(23.17)	(23.17)	(23.17)	(23.17)
p-value: $\gamma + \rho = 0$	0.341	0.133	0.284	0.246
Obs.	3054	3054	3054	3054

Notes: All the columns stack results from Exam 2 and 3 together. "Ever Used" in Panel A equals to 1 if a student has used the applet for at least once for exam studying, and 0 otherwise. "#Questions" in Panel B counts the number of questions a student answers through the applet. The estimates are the mean marginal effects. Standard errors are clustered at the lab section level (68 clusters). The network is constructed based on the network elicited after Exam 2, using the "or" definition. The individual controls included but not shown here are gender, college cumulative and high school GPA, cohort and lecture section dummies, and Exam 1 test score. Control mean is the mean outcome for an untreated student with no treated study buddies. The p-value shown is under the null hypothesis that there are no peer effects on the treated student.

Detailed Model Derivations

To derive the expected utility:

$$
\begin{aligned}
E\big(U_i(e_i)\big) &= E(G_i - mG_i^2 - ke_i^2) \\
&= E(G_{0i} + e_i v_i) - mE(G_{0i}^2 + 2G_{0i}e_i + e_i^2 v_i^2) - kE(e_i^2) \\
&= G_0 + e_i E(v_i) - mG_{0i}^2 - 2mG_{0i}e_i E(v_i) - me_i^2 E(v_i^2) - ke_i^2 \\
&= G_0 + e_i \mu_i - mG_i^2 - 2mG_{0i}e_i \mu_i - me_i^2(\mu_i + \sigma_i^2) - ke_i^2
\end{aligned}
$$

The first order condition:

$$
\frac{\partial E\big(U_i(e_i)\big)}{e_i} = \mu_i - 2mG_{0i}\mu_i - 2me_i(\mu_i + \sigma_i^2) - 2ke_i
$$

$$
e_i^* = \frac{\mu_i - 2mG_{0i}\mu_i}{2\big(m(\mu_i + \sigma_i^2) + k\big)}
$$

Comparative statics:

$$
\begin{aligned}
\frac{\partial e_i^*}{\mu_i} &= \frac{2(1 - 2G_{0i}m)(m\sigma_i + m\mu_i + k) - 2(1 - 2G_{0i}m)m\mu_i}{\big(2\big(m(\sigma_i + \mu_i) + k\big)\big)^2} \\
&= \frac{2(1 - 2G_{0i}m)(m\sigma_i + k)}{\big(2\big(m(\sigma_i + \mu_i) + k\big)\big)^2}
\end{aligned}
$$

when $m < \frac{1}{2G_{0i}}$, $\frac{\partial e_i^*}{\mu_i}$ is positive.

An Alternative Target Input Model

Under a target input model, a student i's utility from spending a_i amount of time on the applet is modeled by equation (3). This model assumes that the utility increases in the target level (A_i) and decreases in the square of the deviation from the target:

$$U_i(a_i) = A_i - (a_i - A_i)^2 \tag{A.1}$$

The target input level, A_i, is not known when students choose how much to study. A_i is determined by:

$$A_i = A + \mu_i$$

where A is the mean optimal amount of time spent on the applet. The actual optimal input level fluctuates due to student specific characteristics. The fluctuations are reflected by an idiosyncratic shock, μ_i. Assume $\sigma_\mu^2(T_i = 1) < \sigma_\mu^2(T_i = 0)$ and a prior belief about A that is $N(A_{0i}, \sigma_{0i}^2)$.

Students choose a_i to maximize the utility specified (1). It is straightforward to see that i should choose the input that equals to her expectation of the target level: $a_i = E(A_i)$. Therefore the expected utility is:

$$E(U_{0i}) = E(A_i) - E\left(E(A_i) - A_i\right)^2 = A_{0i} - \sigma_{0i}^2 - \sigma_\mu^2(T_i) \tag{A.2}$$

With social learning, students get information about the optimal applet usage from partners by observing their usage choice a_j. a_j is distributed as $N(A_{0j}, \sigma_{0j}^2 + \sigma_\mu^2(T_j)) \equiv N(A_{0j}, \sigma_{\mu j}^2(T_j))$. Then i Bayesian updates his/her expected utility after social learning to be:

$$E(U_{1i}) = E(A_{1i}) - \sigma_{1i}^2 - \sigma_\mu^2(T_i) \tag{A.3}$$

where A_{1i} is the updated belief and is a random normal variable with the following mean and variance:

$$E(A_{1i}) = \frac{\frac{A_{0i}}{\sigma_{0i}^2} + \sum_j g_{ij} \frac{A_{0j}}{\sigma_{\mu j}^2(T_j)}}{\frac{1}{\sigma_{0i}^2} + \sum_j g_{ij} \frac{1}{\sigma_{\mu j}^2(T_j)}}$$

$$Var(A_{1i}) = \sigma_{1i}^2 = \frac{1}{\frac{1}{\sigma_{0i}^2} + \sum_j g_{ij} \frac{1}{\sigma_{\mu j}^2(T_j)}}$$

The same comparative statics from the main model thus follow as well.

Estimating Endogenous Peer Effects

Consider the following general peer effects model to write a student's outcome (y_i) as a function of own characteristics (x_i), n_i friends' average characteristics ($\frac{\sum_{a_{ij}=1} x_j}{n_i}$) and their average outcome ($\frac{\sum_{a_{ij}=1} y_j}{n_i}$):

$$y_i = \alpha + \beta \frac{\sum_{a_{ij}=1} y_j}{n_i} + \gamma x_i + \delta \frac{\sum_{a_{ij}=1} x_j}{n_i} + \epsilon_i \tag{A.4}$$

$$E(\epsilon_i | x) = 0$$

$a_{ij} = 1$ if two are study partners and 0 otherwise. β captures the endogenous effect and δ captures the exogenous effect. I treat students as being in one big network since most of the students are in a densely connected component.[1]

Let G be the row normalized adjacency matrix, so that $G_{ij} = \frac{1}{n_i}$ if $a_{ij} = 1$ and 0 otherwise. The above equation can be written as:

$$\boldsymbol{y} = \boldsymbol{\alpha} + \beta \boldsymbol{G} \boldsymbol{y} + \gamma \boldsymbol{x} + \delta \boldsymbol{G} \boldsymbol{x} + \boldsymbol{\epsilon}$$

It is well known that simply regressing $\boldsymbol{y}$ on the endogenous variable $\boldsymbol{G}\boldsymbol{y}$ is problematic. Following Bramoullé et al. (2009) that show β can be identified as long as I, G and G^2 satisfy the independent condition. Satisfying the independence condition means that there are at least two students who are not friends but have a common friend. The identification strategy is to use $\boldsymbol{G^2 x}$ as instruments for $\boldsymbol{G y}$. In other words, one can use friends' friends' characteristics to instrument for friends' outcomes. Since the identifying condition holds in my elicited network, I use the instrumental variable approach with the GMM continuously updated estimator since it is robust to potential weak instruments (Hahn et al., 2004).

Table A6 column (1) shows the results from the naive OLS regression with $\boldsymbol{G y}$ as an endogenous regressor. The endogenous peer effects estimate β is positive and statistically significant. The IV estimate is 0.006. This means that when partners on average answer one more question through the applet, a student answers 0.006 question more. The OLS estimate is larger than the IV estimate. This can be caused by positive selection where students using the applet

---

[1]This is different from previous studies that use the Add Health friendship data and are able to partition the whole sample into subnetworks naturally by schools.

more choose to study with those who also tend to use the applet more. The instrument has an F-statistic of 2.94 which signals that the estimate is subject to weak instrument biases.

Table A6: Estimating Endogenous Peer Effects

Outcome: Applet Q's Answered	OLS	IV
Endogenous Peer Effects (β)	0.179***	0.006
	(0.041)	(0.352)
Individual Controls (γ)	Y	Y
Contextual Effects (δ)	Y	Y
Cragg-Donald F Stat.	-	2.94
Overidentification Test p-value	-	0.145
Obs.	2372	2372

Notes: All the standard errors clustered at the lab section level and are shown in the brackets (68 clusters). The network is constructed based on the network elicited after Exam 2, using the "or" definition. The regressors included but not shown here include individual background control variables such as gender, college cumulative GPA before the semester starts, cohort dummies and high school GPA, partners' average of those background variables, and lecture section fixed effects.

APPENDIX B

Appendix for Chapter Two
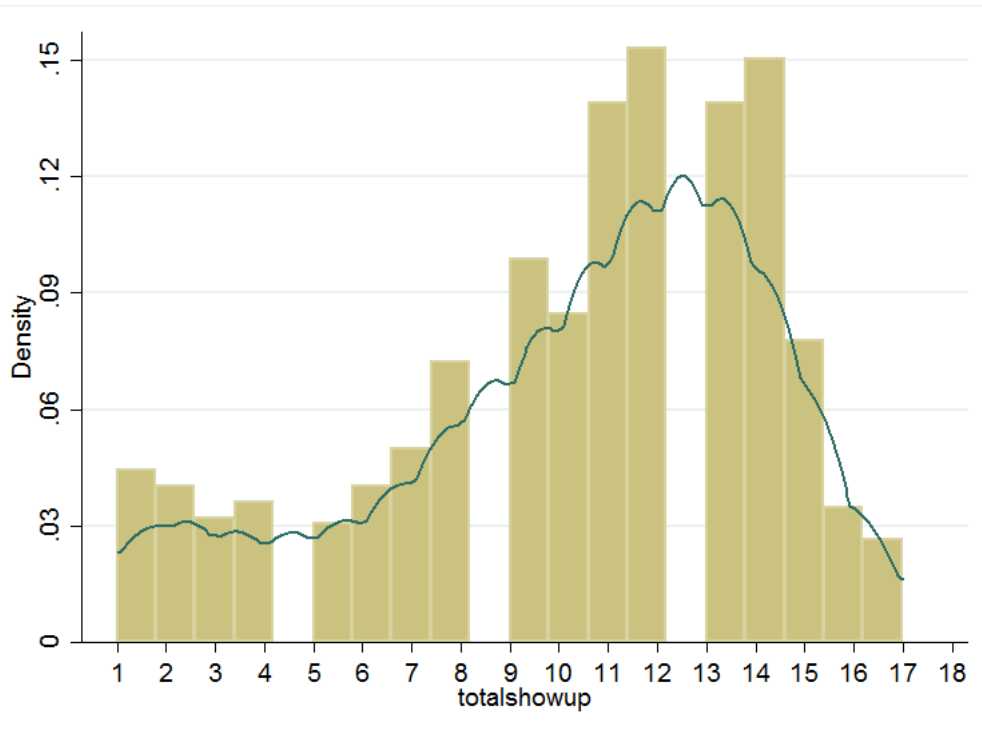
Figure B.1: Distribution of Attendance
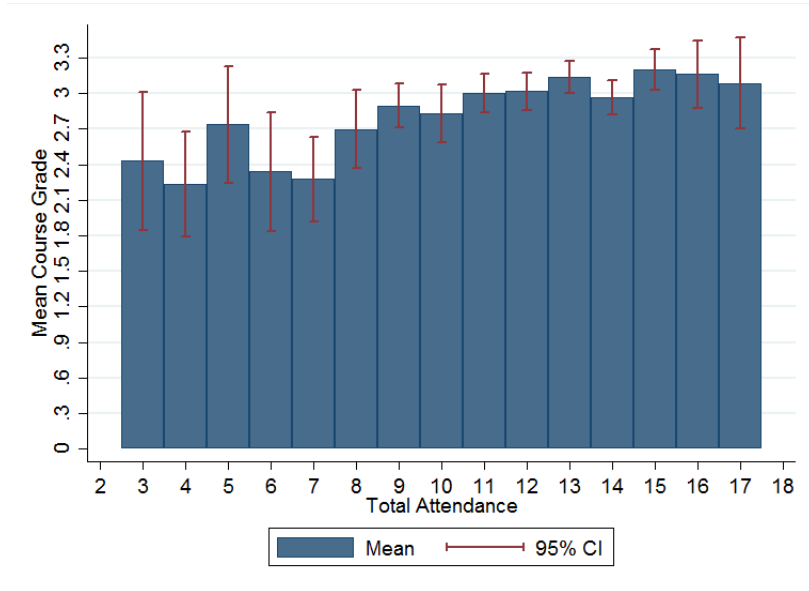
(a) Full
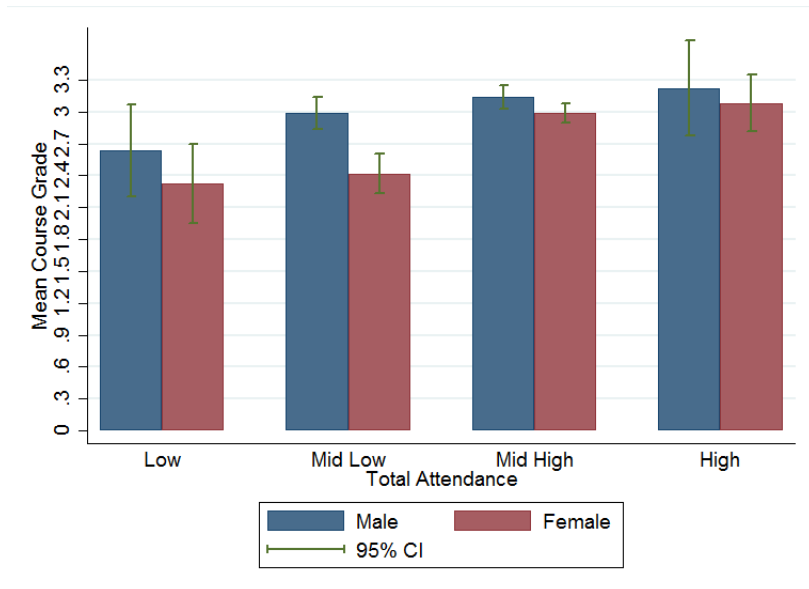
(b) By Gender and Attendance Brackets

Figure B.2: Attendance and Grade

Note: The attendance categories are created in this way. "Low": fewer than 6 attendances. Mid Low: 6-10 attendances. Mid High: 11-15 attendances, High: more than 15 attendances.

Table B.1: Course Grade and Study Group Participation

Course Grade	(1)	(2)	(3)	(4)
Study Group	0.279***	0.293***	0.297***	0.317***
	(0.049)	(0.050)	(0.066)	(0.066)
Female	-0.226***	-0.252***	-0.202***	-0.221***
	(0.049)	(0.049)	(0.077)	(0.076)
Female X Study Group			-0.040	-0.052
			(0.098)	(0.105)
High School GPA	1.523***	1.541***	1.522***	1.540***
	(0.164)	(0.167)	(0.164)	(0.167)
Math Performance	0.027***	0.026***	0.027***	0.026***
	(0.003)	(0.003)	(0.003)	(0.003)
Eng. Performance	0.009***	0.007**	0.009***	0.007***
	(0.003)	(0.003)	(0.003)	(0.003)
Constant	-6.404***	-6.167***	-6.413***	-6.178***
	(0.621)	(0.670)	(0.622)	(0.067)
Control	N	Y	N	Y
R^2	0.21	0.23	0.21	0.23
N	1355	1355	1355	1355

Notes: This is an OLS specification. Study Group is a dummy variables which equals to 1 if one is in a study group participant and 0 otherwise. The other variables are coded the same as in Table 1. Standard errors are shown in brackets. Control variables include dummies for cohort, female gender, hispanic, asian and black ethnicity, lecture sections, and special education programs. In all regressions, dummies for missing background performance measures are included. Then those missing background performances are coded as the median scores of the population.

APPENDIX C

Appendix for Chapter Three

Since you are allowed to name up to ten of your best freshman friends, you have ten independent chances to win the prize. Please make sure that you do list as many as ten different friends. Moreover, please make sure that you list think carefully about the friends you list and the time spent with each friend in order to maximize your winnings.

You can go back to review your responses by clicking the backward button.

Please select the name of a 1st freshman friend.

Last Initial   C ▼
Last Name   Cadagin   ▼
First Name
UMich Email   Cadagin
            Cadel
            Cadoux
            Cahalan
Is this frie Cahen   mate?
            Cahill
            Cai
  ○ Yes   Cain
            Cakar
  ○ No   Calabrese
            Calcaterra Jr
            Calderon
            Callahan
How muc Calvaneso   rage do you spend with this friend per week? Only count
the time s Camaj   one or in small social gatherings (do not include classes).
            Camhi
            Campbell
  ○ 0-30 mir Campos
            Camras
  ○ 30 min to an nour

  ○ 1 to 2 hours

  ○ 2 to 4 hours

  ○ 4 to 8 hours

  ○ More than 8 hours

Figure C.1: Friends Elicitation

115

For each row please indicate whether you prefer the sure payment or the lottery.

$0 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0
$10 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0
$20 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0
$30 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0
$40 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0
$50 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0
$60 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0
$70 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0
$80 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0
$90 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0
$100 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0
$110 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0
$120 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0
$130 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0
$140 for sure	○ ○	50% chance of winning $200 and 50% chance of getting $0

Figure C.2: Risk Preference Elicitation

For each row please indicate whether you prefer the two week column or the future payment.

$75 in two weeks	○ ○	$80 in 3 months
$70 in two weeks	○ ○	$80 in 3 months
$65 in two weeks	○ ○	$80 in 3 months
$60 in two weeks	○ ○	$80 in 3 months
$55 in two weeks	○ ○	$80 in 3 months
$50 in two weeks	○ ○	$80 in 3 months
$45 in two weeks	○ ○	$80 in 3 months
$40 in two weeks	○ ○	$80 in 3 months
$35 in two weeks	○ ○	$80 in 3 months
$30 in two weeks	○ ○	$80 in 3 months
$25 in two weeks	○ ○	$80 in 3 months
$20 in two weeks	○ ○	$80 in 3 months
$15 in two weeks	○ ○	$80 in 3 months
$10 in two weeks	○ ○	$80 in 3 months
$5 in two weeks	○ ○	$80 in 3 months

Figure C.3: Time Preference Elicitation

In this game, you will be randomly and anonymously matched with another student at the University of Michigan who is participating in this survey. You must decide how to divide 100 tokens between yourself and the other person. Each token is worth $0.75 to you, and worth $1.50 to the other person.

Please indicate how many tokens you want to keep to yourself and pass to the other person. The two numbers must add up to 100.

I choose [0] tokens to keep to myself.

I choose [0] tokens to pass to the other person.

Figure C.4: Social Preference Elicitation

Question 1/5

How do you see yourself?

Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?

Please select a value on the scale, where the value **0** means: **"risk averse"** and the value **10** means: **"fully prepared to take risks"**. You can use the values in between to make your estimate.

	risk averse 0	1	2	3	4	5	6	7	8	9	fully prepared to take risks 10
How do you see yourself?	○	○	○	○	○	○	○	○	○	○	○

Question 2/5

How do you see yourself?

Are you generally an impatient person, or someone who always shows great patience?

Please select a value on the scale, where the value **0** means: **"very impatient"** and the value **10** means: **"very patient"**.

	very impatient 0	1	2	3	4	5	6	7	8	9	very patient 10
How do you see yourself?	○	○	○	○	○	○	○	○	○	○	○

Figure C.5: Self-Reported Risk Tolerance

Question 1/5

How do you see yourself?

Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?

Please select a value on the scale, where the value **0** means: **"risk averse"** and the value **10** means: **"fully prepared to take risks"**. You can use the values in between to make your estimate.

	risk averse 0	1	2	3	4	5	6	7	8	9	fully prepared to take risks 10
How do you see yourself?	○	○	○	○	○	○	○	○	○	○	○

Question 2/5

How do you see yourself?

Are you generally an impatient person, or someone who always shows great patience?

Please select a value on the scale, where the value **0** means: **"very impatient"** and the value **10** means: **"very patient"**.

	very impatient 0	1	2	3	4	5	6	7	8	9	very patient 10
How do you see yourself?	○	○	○	○	○	○	○	○	○	○	○

Figure C.6: Self-Reported Patience

Table C.1: Summary Statistics of Choices/Reports from the Follow-up Sample

	Behavioral Measures			Self-Reported Measures	
Phase 1	8.53	3.21	59.5	5.76	5.88
Phase 2	8.58	3.21	59.86	5.79	6.16
Phase 3	8.17	3.51	60	5.83	6.04
Follow-up: Network	8.82	3.52	63.12	5.68	6.91
Follow-up: Lab Pool	8.45	3.36	61.04	5.77	6.7

Notes: Each column reports the average dependent variable for the three Phases of the main study, as well as for the subjects in the follow-up study recruited using the Network Elicitation protocols and using the standard laboratory subject pool.

Table C.2: Similar Choice/Reports between Subjects and Follow-up Sample

DV:	Network Protocol Indicator	
	Coefficient	Std. Err.
Incentivized Risk Measure	0.364	-0.684
Incentivized Patience Measure	-0.114	-0.657
Incentivized Generosity Measure	2.091	-5.223
Self-Reported Risk Measure	-0.0793	-0.377
Self-Reported Patience Measure	0.138	-0.374

Notes: Each row reports the results of a logistic regression of the corresponding dependent variable on an indicator variable for subjects recruited via the network elicitation survey (as compared to subjects from the standard laboratory subject pool).