One email to students: Can a light-touch intervention make a difference?*†

Darren Page‡and Travis Williams§

January 10, 2022

Click here for latest version

Abstract

Poor performance in introductory courses and lack of individualized assistance may contribute to college non-completion. This research aims to identify the effects of increased, personalized instructor feedback on performance in introductory college courses. We conduct an experiment in which poorly-performing students in large lectures are randomized to receive individualized communication through email about their course performance along with a reminder of their instructors' out-of-class availability. Half of the treated students receive an email from the professor, while the other half receive an email from their teaching assistant. We compare the efficacy of outreach from professors to outreach from teaching assistants. We find that neither treatment measurably increases course performance or perception of instructor quality. Emails from professors decrease attendance at TA office hours, suggesting that students view professor and TA office hours as substitutes. Both types of emails increase the frequency at which students seek help from the email's sender, but do not increase the frequency at which students seek help from the other, non-sending instructor, even though all emails contain office hour information for both the professor and the TA. Thus, changing the identity of the sender changes the effects of the nudge, which has implications for the scale-up of nudging programs. If messages come from a source that is impersonal or unknown to the recipient, the nudges may be less effective than those in a small-scale study where sender and recipient have a closer relationship.

*We would like to thank David Frisvold, Eean Crawford, Stephanie Owen, Richard Patterson, seminar participants at the University of Iowa, and session attendees at the 96th Annual WEAI Conference for their helpful comments. We would like to offer special thanks to the instructors and teaching assistants of the courses at the University of Iowa that participated in the experiment.

†The experiment is preregistered at https://www.socialscienceregistry.org/trials/6396

‡Breech School of Business Administration, Drury University; dpage002@drury.edu

§Department of Economics, University of Iowa; travis-williams@uiowa.edu

1 Introduction

Student engagement predicts success in higher education (Kuh et al., 2006). In particular, academic support and feedback after assessments are associated with increased retention, in part because they help students have better experiences and receive higher grades in their classes (Tinto, 2010). Half of students who drop out of college do so during their first year (National Student Clearinghouse, 2019), so academic support and feedback in introductory classes are essential. Universities frequently offer these introductory courses as large lectures, in which academic support, feedback, or any kind of personalized engagement are more difficult to provide due to the larger class sizes. Course instructors in particular would be hard-pressed to provide frequent, individualized feedback to students.

We conduct an experiment to see whether a simple personalized message can increase student engagement. Specifically, in four introductory, hybrid lecture courses at the University of Iowa we test whether a low-cost, one-time email from instructors affects study habits and help-seeking behavior or improves student experiences more generally, and if these changes improve student performance in the course. The intervention consists of a single, personalized email sent to struggling students that acknowledges their current course performance, offers subtle encouragement, and provides reminders about office hour availability.

Our paper builds on the work of Carrell and Kurlaender (2020), who conduct a similar experiment in which some students receive personalized emails from professors. They find that the emails generally improved students' perceptions of their instructors but improved students' grades in only some settings. Our study differs in two main ways. First, we include two treatment arms. Some students are randomly assigned to receive an email from the professor and others to receive an email from their teaching assistant (TA). The content of the emails is the same. Second, we use an end-of-semester survey to collect data on intermediate outcomes that could influence student grades, including number of email correspondences, number of office hour visits, and hours spent studying.

Varying the sender of the message allows us to assess whether students respond more to the content of these messages or to their relational components. This information contributes to the literature on the effectiveness of nudges in higher education. Nudges are appealing because they

typically carry low costs relative to the potential benefits of improved student performance. However, these interventions are often less effective when implemented on a large scale (DellaVigna and Linos, 2020; Bird et al., 2021). Nudges implemented at a local level may be more effective because the recipients of locally-based nudges are more likely to know and have a relationship with the sender (for example, Avery et al., 2020). We test whether the identity of the sender matters by including schedules and links for both the professor's and the TA's office hours as part of the nudging emails. If the content of the message is all that matters, we would expect the effects of the emails would be the same in the two treatment arms.

We find that the email interventions have no measurable effects on course performance. Students receiving an email from the professor or their TA do not score significantly higher in the course or score significantly higher on subsets of graded assessments. While estimates are somewhat imprecise, we can comfortably rule out large effects. This finding differs from that of Gordanier et al. (2019), who find that struggling students benefit from referrals to academic support resources. In contrast to Carrell and Kurlaender (2020), we find no measurable increases in student perception of the professor or their TA.

We find that the email intervention affects student help-seeking behavior differently based on the sender. Students who receive an email from the professor increase email correspondence with the professor and may increase office hour visits with the professor, while measurably decreasing office hour visits with the TA. Students who receive an email from the TA were slightly more likely to attend both the professor and TA office hours.

Because we conduct the experiment in hybrid courses, our work is also relevant to the literature on the effectiveness of online instruction. Students believe that they don't learn material as well in online sections as they do in face-to-face sections, and this is due in part to reduced interaction with their instructors (Jaggars, 2014). Our results suggest that students perceive greater barriers to help-seeking outside of class in online learning. Students are twice as likely to ever visit their TA's office hours than the professor's office hour. This difference may be driven by the fact that most TAs taught in-person, while professors taught fully online. Both types of emails close this gap in office hour attendance. We also provide suggestive evidence that professor-sent emails have a stronger effect for students enrolled in in-person discussion sections. Given the benefits provided by the flexibility of online and blended learning (Jaggars 2014; Goodman et al. 2019) and the growing

popularity of these courses, exploring methods to aid in learning and retention is worthwhile. Our findings indicate that communication matters for the ways that students seek help in online or hybrid courses.

Our work also contributes to a small literature on the role of TAs in higher education. TAs are a common fixture in university teaching, but few researchers have investigated their impact on students. Feld et al. (2019) compare student and faculty instructors of discussion sections and find that the student instructors are only slightly less effective than the non-students. Lusher et al. (2018) find that minority students benefit from having a same-race TA. In this paper we evaluate the effectiveness of feedback and communication from student instructors. We find that students increase help-seeking from the instructor that sent the email, and that students treat out-of-class help from TAs as a substitute for out-of-class help from the professor.

2 Conceptual Framework

Student engagement predicts success in college (Kuh et al., 2006). For students to be engaged in learning requires both student action and institutional resources for them to interact with (ibid). An important component of student engagement is student-faculty contact (Chickering and Gamson, 1987). In addition to having contact with their instructors in the classroom, students are also able to engage with their instructors over email and during the instructors' office hours. Many institutions mandate a weekly minimum number of hours that instructors must be available for office hours. This outside-of-class availability can provide opportunities for individualized mentoring, feedback, and instruction. For students enrolled in large lectures, contact with instructors outside of class may be particularly valuable due to fewer individualized interactions during class.

Many students, however, do not attend office hours (Griffin et al., 2014) or email their instructors with questions about course material. This may happen for a variety of reasons: (1) students may not perceive out-of-class interactions as valuable, (2) they may not know what office hours are or that they can email their instructors with questions, (3) instructors may schedule office hours at a time or place that makes it costly for students to attend, or (4) students may perceive their instructors as unapproachable or unwilling to help them. Smith et al. (2017) survey students at a public research university and find that all of these factors play a role. Many students are unsure of

the purpose of office hours, or don't know what they would talk about if they attend. Office hours are often scheduled at inconvenient times or places, and students expressed that their instructors were intimidating or even actively discouraged them from using office hours.

If students experience uncertainty in the costs and benefits of engaging with their instructors outside of class, a nudge from instructors could increase engagement. Suppose, for example, that students know some instructors are willing to engage outside of class and others are not. In the absence of signals from instructors, students would not be able to distinguish between types. Because of institutional support for office hours and other out-of-class interaction, an equilibrium in which all unwilling-to-help instructors send a signal of their unwillingness is unlikely; such a signal would be costly due to institutional pressures and possible discipline. If students perceive a high-enough likelihood of an instructor being unwilling to help or if the cost of seeking interaction with an unwilling instructor is high enough, then students may not engage with willing instructors when they receive no signal from them. However, if students receive a signal that the instructor is willing to help, they would likely seek help with greater frequency. These signals would also be credible, because unwilling instructors have no incentive to provide any signals outside of what is prescribed by their institution. It is likely that in the absence of positive nudges from instructors, outside-of-class interaction is inefficiently low.[1]

Nudges could be particularly valuable in large lecture courses. At large state universities, introductory courses often center around large lectures. Instructors are unable to maintain frequent communication with students in the way they might in smaller classes. Instructors in these courses typically communicate with students through broad means either in online portals or through mass emails. They also provide support through office hours, where students can drop in to ask questions or discuss their course performance. Students maintain the responsibility to seek help however, even if they are struggling in the course. Students enrolled in online or hybrid courses may be even more reluctant to seek help because their in-person interactions with instructors are less frequent or nonexistent. Our intervention consists of a single, personalized email from the course instructor or teaching assistant to a struggling student. The email summarizes the student's performance in the course to date, provides subtle encouragement, and reminds them of office hour

[1] This is consistent with Kuryan and Voronina (2021), who find that instructor and student interest in out-of-class communication is higher than actual participation in it.

availability. Because of the nature of large lecture courses, these students would likely receive no other unsolicited, individualized feedback or communication from their instructors.

The email does not fill an informational void. The summary of the student's performance, including a reminder of the first exam grade and a projection of their current course grade, are readily available and easy to calculate through other means. Smith et al. (2018) show that contextualizing student performance after assessments may help their course performance, however. Office hour information is also provided in course syllabi and online portals. Instead, the email signals to the student that the email sender is aware of their current course performance and cares enough to communicate with them. This kind of oversight is uncommon in large lecture courses, which strengthens the signal that the instructor truly cares about the student. Research suggests that students respond positively to perceptions that their instructors care, even in the absence of actual caring (Teven and McCroskey, 1997). In addition to signaling instructor caring and willingness to help, the email also increases the salience of possible out-of-class interactions with the instructor (similar to Betinger et al 2020).

The email nudge directly targets student engagement by reminding students of instructor availability and by signaling the instructor's willingness to engage. Students that receive an email from their instructor may increase email communication or office hour attendance. Whether or not the intervention will increase grades depends on whether out-of-class interaction with instructors is a substitute or complement of other study efforts, and whether or not the intervention is effective in increasing that interaction. For example, if each office hour visit increases student grades by 0.8 percentage points on average (the association found in Guerrero and Rod 2013[2]), then to detect an effect on grades in our sample, the intervention would need to cause students to attend office hours 3.5 more times, on average, than they otherwise would have. Such an increase, though large, is plausible if without the intervention students would not attend office hours or email instructors because they don't know that they can or because they incorrectly believe that the instructor is unwilling to help them. Our study is powered to detect a 2.3 percentage point effect of the interven-

[2]Guerrero and Rod (2013) is a useful reference although their setting (political science courses, and at a different university) is different from ours because in spite of a widespread belief that office hours are helpful to students there are very few direct measures of the correlation between number of office hour visits and course grades. However due to the different settings and the lack of causal identification in Guerrero and Rod (2013) we view their finding of a 0.8 percentage point increase in grades as a useful comparison to our findings rather than as a prediction of what we expect to find.

tion on grades, and an effect of 17 percentage points on the likelihood of a student ever attending the professor's office hours.

Receiving a single email in a single course may improve students' experience in that course, but such a light-touch intervention on its own is less likely to impact their overall college experience or the likelihood that they graduate. However because nudges like this one are easy for individual instructors to implement, it would be possible for them to be applied at scale, in several courses at a university. Increased professor outreach in several of their courses could lead to students engaging more with their instructors overall, improving grades, which could lead to higher completion rates (Denning et al. 2021). However, large-scale nudging campaigns have proven ineffective in many settings (Oreopoulos and Petronijevic 2019; Oreopoulos et al. 2018; Bird et al. 2021).

The success of scaling up the intervention will depend on the underlying mechanisms through which the email affects student behavior. If the email signals to students that the sending professor's office hours will be useful, an institution-level scale up would be ineffective, since requiring all instructors to send these emails would remove any signaling value. However an institutional practice of encouraging instructors to initiate contact with their students–without requiring it–could scale up, as the signal would continue to distinguish between instructor types. Also significant to the value of the signal could be the identity of the sender. Many large-scale nudges rely on impersonal distribution methods, like mass text or email. But if the value of the intervention comes not from the content of the message, but rather from knowing who sent the message (Teven and Gorham 1998), then any scale-up would need to preserve that relational aspect of the emails in order to be effective.

To understand whether the identity of the sender matters for the success of the intervention, we split the treatment group into two arms. One group of students receives an email from the professor leading the lecture course, and the other group receives an email from the TA that leads the discussion section. Control group students receive no personalized email. The content of the emails varies very little across the two treatment arms, only in the order of presentation. If the value of the email is purely in its content, we would expect the effect of the intervention to be the same for both treatment arms. If the relational component does matter, the effect of receiving an email will vary based on the sender's identity.

3 Experimental Design and Data

3.1 Intervention and Data Collection

We conduct the field experiment in 1000 and 2000 level courses that have a large lecture component and a discussion component with a teaching assistant at the University of Iowa. We include four different courses, three of which occur within the Tippie College of Business and are repeat participants from Fall 2020 to Spring 2021. The fourth course is outside the College of Business and only participates in Spring 2021. A total of 3,364 students were enrolled in these courses for an average of 481 students per course. Due to the COVID-19 pandemic, each course was delivered via hybrid or fully-online instruction. Professors in all seven courses delivered lectures either synchronously or asynchronously online. Discussion sections led by teaching assistants took place once a week with in-person instruction in about two-thirds of the sections and online instruction in one-third.

Figure 1 lays out the timeline of the experiment. Within the first month of each semester, professors informed their students of the possibility to earn extra credit on their final exam by participating in a research study or by writing a short essay.[3] Students who chose to participate in the research study followed a link to an initial course survey, where they consented to have their grade data shared. The consent document also informed students that the study aimed to understand the effects of communication with course instructors and that they "may or may not receive different types of communication" from their instructors. Students were informed that they would not be provided with information not available in other formats. While students were aware of their participation in a research study, they were not made aware of the details of the experiment. The initial survey also gathered data by inquiring about student's perceptions of professors and teaching assistants in general, expectations of course difficulty, and demographic information. The initial survey is provided in Appendix C. Of the 3,364 students eligible for the study, 59 percent elected to participate.[4]

The experiment and subsequent analyses include only students scoring below 70 percent on the

[3]For one of the four courses, students were given the opportunity to participate in the research study to earn course credit. The professors provided multiple alternatives to earning these credits, including participation in other research studies throughout the semester.

[4]Differences between average first exam scores for study participants and class averages were minimal, suggesting that participation is not strongly selected on course performance.

first midterm exam. As a fraction of participating students, 31 percent scored below the threshold.[5] We randomize these students within discussion sections to (1) receive an email from the professor, (2) receive an email from their TA, or (3) receive no such email. The emails were sent to the students' university email accounts within two days of the instructor posting or sharing first exam grades. While we do not observe whether students saw, opened, or read the emails, we do know from professors and TAs that several students responded to the emails. These courses were either hybrid or online and (non-personalized) email communication from instructors was common. It is reasonable to believe that the majority of treated students saw the email.

The intervention email is a personalized message from the professor or teaching assistant that acknowledges the student's first exam score, projects their final grade, provides encouragement, and reminds the student of office hour availability. Emails sent by professors and teaching assistants vary only in the ordering of the office hour information. Figure 7 in Appendix A provides a hypothetical example of a professor-sent email and a TA-sent email. Emails are personalized with the student's first name, exam score, and projected grade. Each email also provided direct links to virtual office hours held via Zoom. We instructed professors and TAs to respond to any replies to these personalized emails as they normally would. There were no other components of the intervention.

In the second-to-last week of each semester we administered a second survey to measure whether receiving the personalized message affected student attitudes or behaviors. The second survey asks students about their perceptions of their professor, their perceptions of their TA, the frequency of email communication and office hour visits with each instructor, and their interest in future courses within the discipline. Failure to complete the second survey meant that students received half credit for their participation in the research study. Of students who completed the first survey, about 66 percent successfully completed the second survey. Figure 1 quantifies student participation, experiment eligibility, and survey completion. Our analysis sample includes the 393 students who completed both surveys and scored below 70 percent on the first exam. We also remove three students who dropped the course, two students given incompletes, and five students with a large number of missing survey responses for a final total of 383 students.

[5] We limited the experiment to this subset of students to reduce the burden on instructors and TAs and because students with lower scores on the first exam are more at risk of doing poorly in the course.

3.2 Summary Statistics and Balance

Table 1 summarizes the measures of interest gathered before the intervention. The sample includes all participating students regardless of whether or not they completed the second survey or dropped the course. The first column provides sample means for all students. The second column provides sample means for the experimental group: those who scored below 70 percent on the first exam. The next three columns summarize student measures by treatment arm.

The students who score below 70 percent on the first exam differ from the sample as a whole in more ways than just first exam score. They are less likely to have a high school GPA above 3.5, more likely to be first generation, and end up dropping the course with greater frequency. They are otherwise demographically similar.

We compare means across treatment statuses to examine balance and ensure the validity of the randomization. The final two columns provide p-values from t-tests as a result of regressions of student characteristics on treatment statuses and strata fixed effects. The two treatment arms are not statistically different from the control group for objective measures like first exam score, high school GPA, or demographics. However, students in the control group do subjectively rate professors and TAs at the University of Iowa higher than students assigned to each of the two treatment arms. This tendency to rank instructors lower on average could affect our later estimates of the intervention's effect on perception of instructors at the end of the semester. There are no significant differences in course completion or survey two completion by treatment status, which ensures that dropping non-completes and students lost to follow-up has no effect on our analysis.

Table 2 summarizes the same measures for the students that complete the course and second survey, such that experimental sample comprises our analysis sample. Differences between survey respondents as a whole and the experimental sample are similar to the unrestricted sample. We include an additional table in the Appendix that tests balance across all categorical responses in the first survey, including indicators for missing data.

Similar to the balance for all participating students, the analysis sample is balanced on key objective measures like first exam score and high school GPA. The analysis sample also contains imbalances on impressions of professors and teaching assistants.[6] Students randomized to receive

[6]The impressions variables are indices from 0-4 that count responses of "Strongly agree" to high-quality impressions of professors and teaching assistants. Alternative measures, such as indexing on "Strongly agree" and

no email are also more likely to be first generation students. To account for these imbalances, we carefully examine the sensitivity of our results to the inclusion of control variables.

4 Empirical Strategy

We test the effectiveness of the email intervention on measures of course performance, ratings of course professors and teaching assistants, and frequency of communication with professors and teaching assistants. Given the randomized nature of the email intervention, we estimate the following models by ordinary least squares:

$$y_{ij} = \alpha_0 + \alpha_1 Professor_i + \alpha_2 TA_i + \alpha_3 S_j + \varepsilon_{ij} \tag{1}$$

$$y_{ij} = \alpha_0 + \alpha_1 Professor_i + \alpha_2 TA_i + \alpha_3 S_j + \alpha_4 X_{ij} + \varepsilon_{ij} \tag{2}$$

where $Professor_i$ and TA_i indicate receipt of an intervening email from the professor and TA respectively. The estimates of α_1 and α_2 compare measures for students treated by an email to students who received no intervening email.

As shown previously, there exist imbalances in pre-intervention measures such as general impressions of instructors, expected course difficulty, and first generation across treatment statuses. We account for this by adding in strata fixed effects, S_j, before also including a substantial number of individual-level covariates, X_{ij}. Because randomization is stratified within individual discussion sections, strata fixed effects capture differences that may be driven by teaching assistant quality or driven by selection into different discussion sections such as Honors sections. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty, first exam score, and general impressions of professors and teaching assistants at the University of Iowa.

Our primary analysis includes nine dependent variables. The first three correspond to course success. Overall course percent reflects the final class grade on a 0-100 scale. We also estimate effects on an indicator for scoring at least a B- in the course and on second exam score.[7] Additional

"Somewhat agree" or averaging across the scales, are more balanced across treatment statuses.

[7]Each course included in the study uses different thresholds translating course percents into letter grades. A score of B- corresponds to the 2.75 GPA threshold necessary for admission to the Tippie College of Business.

outcome variables include ratings of course professors and TAs collected in the second survey. Finally, we estimate effects for communication with professors and TAs through email and visits to office hours. Our primary analysis is conducted on the sample of students who successfully completed both surveys and completed the course. We also include analysis of effects on gradebook data solely for the sample of students that successfully completed the course, regardless of their completion of the second survey. Finally, we conduct heterogeneous analysis across different groups of students to assess the potential for effects to vary by type of sender or type of recipient. Standard errors are clustered at the strata level in all analyses.

5 Results

5.1 Main results

Table 3 shows the main results from our preregistered analysis plan. We find small, positive, and not statistically significant effects of the intervention on the final percent score among students that responded to the second survey and completed the course (survey 2 sample). Students receiving emails from professors earned higher grades (a 1.5 percentage point increase in the overall score in the class relative to the control group) than students that received an email from their TA (a 0.4 percentage point increase), though effect sizes do not statistically differ. We can rule out effects larger than 3.5 percentage points for the professor email and effects larger than 2.5 percentage points for the TA email. These bounds are smaller than the 6.4 percentage point increase that Carrell and Kurlaender (2020) find in their pilot study[8], and are larger than but not statistically different from the results of their scaled-up study, in which the professor-sent email intervention had no measurable effect on overall grades.

When we also include in our sample those students that responded to only the first of the two surveys (experimental sample), as shown in the bottom half of the table, the estimates are closer to zero, with 95% confidence intervals within -2 and 1.5 in all specifications for both the professor and the TA interventions. We similarly find no evidence of strong positive effects of the interventions on the other grade outcomes that we observe: percent score on the second exam, percent score on

[8]Carrell and Kurlaender also conduct a follow-up experiment to replicate the findings from their pilot. In that experiment, they find that on average, students receiving an email from their professors earned 4.4 percentage points more in their overall grade in the class than students receiving no email.

the final exam, earning at least a B- in the course, and failing or not completing the course. If these light-touch email interventions improve student grades on average, it is by only a small amount.

These results add to the evidence that college students are mostly unresponsive to behavioral interventions in terms of course achievement. Oreopoulos and Petronijevic (2019), for example, find no significant improvement in achievement across interventions of varying intensities aimed at aiding students in large introductory courses. In their study, interventions ranging from goal-setting exercises to repeated, intensive coaching have no measurable effects on course performance.

In addition to being a one-time intervention, our email design also lacks actionable steps for students to take other than an indirect suggestion that office hours may be helpful. This differs from other interventions in higher education that have proven helpful by nudging students to take certain actions. Examples include encouragements to finish college applications (Oreopoulos and Ford, 2019), renew financial aid (Castleman and Page, 2016), and enroll in selective colleges (Dynarski et al., 2018).

We also do not find statistically significant positive effects among survey 2 respondents on their perceptions of the instructors. In fact, students that received an email from either a professor or a TA were slightly less likely to rate their instructors highly than their peers in the control group. This contrasts with the findings of Carrell and Kurlaender (2020), who consistently estimate positive effects of additional email communication on student perceptions of their instructors, even in settings where there is no effect on grades. Our negative estimate of the effect on student perceptions of their instructors may be due to differences in the baseline characteristics of the treatment and control groups. In the first survey, when asked four questions about their overall perceptions of professors and four questions about their overall perceptions of TAs at the university, students in both treatment arms gave their instructors the highest rating ("Strongly agree") on 0.3 - 0.5 fewer questions on average that the students in the control group gave theirs. If the students in the treatment arms are predisposed to give lower ratings to instructors or if they incorporated in their answers to the survey one questions a negative view of the instructor teaching the course in which we conducted the experiment, our results would be biased downward.

Although we find no evidence that the intervention improved students' perceptions of their instructors, we do find evidence that the email messages affected the likelihood that students would email their instructors or attend office hours, and that the effect varies depending on who

sends the message. This is consistent with students updating their beliefs based on the identity of the sender, and not on the content of the message. The Survey 2 sample estimates in Table 3 show the effect that the professor and TA emails had on further email communication[9] and office hour attendance.

5.2 Spillovers

Interventions like ours can have spillover effects if treated students share information they receive with untreated students. If this occurs, positive spillovers would bias our estimates toward zero. We test for spillovers by comparing untreated students across discussion sections, leveraging variation in the number of same-section peers who received intervening emails.[10] There are two types of untreated students: those who scored below 70 percent on the first exam and were randomized to receive no email (control group) and those that who scored above 70 percent on the first exam (non-experimental students). We distinguish between these groups due to the correlation between one's own treatment status and the statuses of peers. A control group student is likely to have more treated classmates as a purely statistical result rather than as a result of random variation in the first exam score of peers. We also generate separate estimates to gauge the effects of spillovers on our estimates, as control students provide the counterfactual in our main analysis.

First, we examine the possibility of spillovers with simple averages. Figure 5 includes eight panels. Each panel has a different combination of sample (control vs. non-experimental) and outcome variable. Within each panel, we show the average of the outcome variable as the number of in-section peers receiving an email from the professor changes from a minimum of 0 to a maximum of 5. Because we hypothesize that professor emails and TA emails have different effects, we show these averages separately by the number of in-section peers receiving an email from the TA.[11] The size of each point is scaled by the number of students represented.

The movement in averages are suggestive of potential spillovers. Consider the outcome of Exam 2 percent, the second row from the bottom. We see a downward trend, indicating that more in-section peers receiving an email from the professor correlates with worse performance on subsequent

[9]

[10] We omit sections with no treated students; these are honors sections and have unusually high average scores.

[11] There exists variation in N_{prof} for given values of N_{TA} because the number of students eligible for the experiment in a given discussion section was not always divisible by three. Therefore, treatment is not perfectly balanced within discussion sections.

exams. Estimating the specification

$$y_{ij} = \alpha_0 + \alpha_1 N_{prof} + \alpha_2 N_{TA} + \alpha_3 X_i + \varepsilon_{ij} \tag{3}$$

y includes the main outcome variables of interest, N_{prof} is the number of peers in the discussion section that received an email from the professor, N_{TA} is the number of peers in the discussion section that received an email from the TA, and X is the full set of covariates less strata fixed effects,[12] would yield a negative coefficient for α_1 for the exam 2 outcome. However, if the number of a student's peers that receive treatment is associated with the outcome in a way that does not involve spillovers from the treatment, this would bias the spillover estimates.

Figure 6 summarizes the pathways through which a student's outcomes can be influenced by their own and their peers' characteristics. The first important observation is that a students' own characteristics that influence their outcomes also have a causal effect on the number of students in each treatment arm. This creates a backdoor path between the number of a student's treated peers and the student's outcomes. A student's characteristics affect the size of the treatment groups in their own section because the number of treated students is weakly increasing in the number of students scoring below 70 percent on the first exam. Therefore, students that would tend to get worse grades anyway will also tend to have more peers receiving treatment. In our tests for spillovers, we can eliminate this particular backdoor path by limiting the sample to control-group students only.

A backdoor path that is more difficult to close could exist if there are peer effects. Spillovers themselves are a type of peer effect, so if we are interested in measuring spillovers it is natural to be concerned about other types of peer effects. A student's peers affect the student's exposure to treated students, which could affect the student's outcomes. If the other students also influence student outcomes directly, as shown in the lower-right corner of Figure 6, then estimates of spillovers from our treatment will be biased unless we can adequately control for those peer effects.

To address this concern, we estimate spillover effects with the following empirical model:

$$y_{ij} = \alpha_0 + \alpha_1 N_{prof} + \alpha_2 N_{TA} + \alpha_3 X_i + \alpha_4 \overline{\text{Exam 1}}_{-i,j} + \varepsilon_{ij} \tag{4}$$

[12] We omit the strata fixed effects out of necessity, since there is no within-stratum variation in the number of treated students.

where y, N_{prof}, N_{TA}, and X are as described above, and $\overline{\text{Exam 1}}_{-i,j}$ is the average exam 1 score of the other students in the discussion section. Controlling for $\overline{\text{Exam 1}}_{-i,j}$ certainly does not for all the ways that a students' peers could affect their outcomes, however, within the constraints of our experiment, it provides one way to test whether other peer effects are biasing the spillover estimates. We estimate both equations 3 and 4 to be able to compare specifications with and without controls for peer effects. Estimates of α_1 measure spillover effects of an additional peer receiving a professor email and estimates of α_2 measure spillover effects of an additional peer receiving an email from the TA.

Table 4 provides the estimates for control group students. The first two columns show no statistically significant effects of having additional peers receive emails on the 209 control group students with full grade data. When we control for the average first exam score within the section, estimates indicate that students perform worse in the course when additional peers receive an email. By looking at the results with survey data, we see that students are more likely to engage in email communication with the professors and attend TA office hours when an additional peer receives an email from the professor.

In total, these estimates provide no clear, consistent pattern of spillover effects. We lack clear explanations for negative effects on course performance. Those estimates are also inconsistent with the effects showing increased communication with the professor and TA office hours visits. The estimated effects that reach statistical significance are also implausible effects of having one additional peer receive an email. There is no clear evidence that control group students were uniformly affected by having treated peers. Therefore, our main estimates are unlikely to be consistently biased toward zero.

Table 5 provides estimates of spillovers on students who scored above 70 percent on the first exam and were not experiment participants. Estimated spillover effects are smaller, consistent with the idea that students who did not receive an email were not heavily affected by having peers who did.

5.3 Exploratory Analysis

We conduct the experiment with student scoring low on the first exam in order to assess the effectiveness of the intervention on students at risk of failing the course. We can also look directly

at the effects of the intervention on disadvantaged groups of students or students most at risk of failing introductory courses independent of their first exam score.

Table 6 provides the results by first-generation status. Table 7 provides the results by whether or not students expected the course to be more or less difficult than their other courses. Table 8 provides the results for students with a high school GPA above and below 3.5. The estimates for these subgroup analyses are all subject to statistical imprecision, and the point estimates vary quite a bit depending on the inclusion of control variables. We can safely rule out that heterogeneity is masking large effects within the main subgroups of interest however, as point estimates on course achievement are consistently small.

Because we identify different responses to different sender in the main analysis, we take a further interest in the characteristics of the email senders. First, we divide the sample based on TA quality, which we measure by taking average scores for each TA across untreated and control group students. We restrict the sample to include only students enrolled in a discussion section led by a TA with average scores in the top 13 out of the 27 TAs who sent emails.[13] Table 9 presents the effects of both email interventions for these students. The most notable finding is that substitution away from TA office hours persists despite the measured quality of the TA. The statistical precision of this result is sensitive to the set of controls, however.

Out of the 104 discussion sections with at least one student who received an email, 37 were fully online. We conduct separate analysis on students enrolled in the 67 sections with in-person instruction with the TA and present those results in Table 10. Students receiving an email from the TA respond differently than students receiving an email from the instructor in terms of help-seeking behavior. Most notably, professor emails led to more students scoring at least a B- in the course. A few caveats should be noted for these results. First, two of the four courses included in the study used online discussion sections exclusively. Therefore, we are restricting the sample to two of the four courses. Second, students may select into online vs. in-person discussions sections in ways that make these comparisons difficult. We provide equivalent tables that show heterogeneous effects of the email interventions on experimental samples in the Appendix.

Table 11 provides the results for additional exploratory outcomes collected in the second survey.

[13]Students enroll in courses and discussion sections before listings include the name of the TA, greatly reducing the chances of student selection into sections based on TA characteristics.

We ask students to report which days their professors and TAs hold office hours. Correct answers consist of survey responses that are a subset or an exact match for the actual days office hours were held. We also ask students if they are registered or plan to register for more courses in the discipline. Finally, they report the number of hours they spent on the course in categorical responses which we group into five hours and over and eight hours and over. The results are statistically insignificant for all outcomes, including knowledge of the days of office hours which have negative point estimates.

6 Discussion and Conclusion

A majority of students who drop out of college do so within their first year of enrollment. Many of these students will have only taken introductory courses with large enrollments in which outreach and personalized communication with instructors is limited.

We test the effectiveness of an intervention that consists of a single email to struggling students early in the semester. The email acknowledges and quantifies the students course performance, provides brief encouragement, and reminds students of office hour availability. We vary the sender to assess the effectiveness of emails sent by a professor and emails sent by a TA, comparing students randomized to receive these emails to students randomized to receive no such email.

We find no significant effects of either email type on final course score, scoring at least a B- in the course, or scores on subsequent exams. We also find no evidence that student perceptions of the quality of the professor or the TA shift in response to receiving an email. Receiving an email from the professor or TA does change the way students seek help in the course. We show that professor emails increase communication with the professor through email or office hours and lead students to substitute away from attending TA office hours. TA emails generally increase help-seeking from both the TA and the professor.

Our results imply that light-touch interventions are unlikely to drastically improve course performance or increase retention for students in introductory courses. Successful interventions that rely on information provision or communication from instructors likely require repeated follow-up with students.

7 References

[1] C. Avery, Benjamin L. Castleman, Michael Hurwitz, B. Long, and Lindsay C. Page. Digital messaging to improve college enrollment and success. *National Bureau of Economic Research*, 2020.

[2] Eric Bettinger, Nina Menezes Cunha, Guilherme Finkelfarb Lichand, and Ricardo A. Madeira. Are the effects of informational interventions driven by salience? *Randomized Social Experiments eJournal*, 2020.

[3] Kelli A Bird, Benjamin L Castleman, Jeffrey T Denning, Joshua Goodman, Cait Lamberton, and Kelly Ochs Rosinger. Nudging at scale: Experimental evidence from fafsa completion campaigns. *Journal of Economic Behavior & Organization*, 183:105–128, 2021.

[4] Scott E Carrell and Michal Kurlaender. My professor cares: Experimental evidence on the role of faculty engagement. Working Paper 27312, National Bureau of Economic Research, June 2020.

[5] Benjamin L Castleman and Lindsay C Page. Freshman year financial aid nudges: An experiment to increase fafsa renewal and college persistence. *Journal of Human Resources*, 51(2):389–415, 2016.

[6] National Student Clearinghouse. Yearly success and progress rates (fall 2012 entering cohort). Technical report, National Student Clearinghouse Research Center, 2019.

[7] Stefano DellaVigna and Elizabeth Linos. RCTs to scale: Comprehensive evidence from two nudge units. Working Paper 27594, National Bureau of Economic Research, July 2020.

[8] Jeffrey T Denning, Eric R Eide, Kevin Mumford, Richard W Patterson, and Merrill Warnick. Why have college completion rates increased? an analysis of rising grades. Working Paper 28710, National Bureau of Economic Research, April 2021.

[9] Susan Dynarski, CJ Libassi, Katherine Michelmore, and Stephanie Owen. Closing the gap: The effect of a targeted, tuition-free promise on college choices of high-achieving, low-income students. Technical report, National Bureau of Economic Research, 2018.

[10] Jan Feld, Nicolás Salamanca, and Ulf Zölitz. Students are almost as effective as professors in university teaching. *Economics of Education Review*, 73:101912, 2019.

[11] Joshua Goodman, Julia Melkers, and Amanda Pallais. Can online delivery increase access to education? *Journal of Labor Economics*, 37(1):1–34, 2019.

[12] John Gordanier, William Hauk, and Chandini Sankaran. Early intervention in college classes and improved student outcomes. *Economics of Education Review*, 72:23–29, 2019.

[13] Shanna Smith Jaggars. Choosing between online and face-to-face courses: Community college student voices. *American Journal of Distance Education*, 28(1):27–38, 2014.

[14] Lester Lusher, Doug Campbell, and Scott Carrell. Tas like me: Racial interactions between graduate teaching assistants and undergraduates. *Journal of Public Economics*, 159:203–224, 2018.

[15] J. Monks and Richárd Schmidt. The impact of class size on outcomes in higher education. *The B.E. Journal of Economic Analysis Policy*, 11, 2011.

[16] Philip Oreopoulos and Reuben Ford. Keeping college options open: A field experiment to help all high school seniors through the college application process. *Journal of Policy Analysis and Management*, 38(2):426–454, 2019.

[17] Philip Oreopoulos, Richard W Patterson, Uros Petronijevic, and Nolan G Pope. When studying and nudging don't go as planned: Unsuccessful attempts to help traditional and online college students. Working Paper 25036, National Bureau of Economic Research, September 2018.

[18] Philip Oreopoulos and Uros Petronijevic. The remarkable unresponsiveness of college students to nudging and what we can learn from it. Technical report, National Bureau of Economic Research, 2019.

[19] Ben O Smith, Dustin R White, Patricia C Kuzyk, and James E Tierney. Improved grade outcomes with an e-mailed "grade nudge". *The Journal of Economic Education*, 49(1):1–7, 2018.

[20] Jason J. Teven and Joan Gorham. A qualitative analysis of low-inference student perceptions of teacher caring and non-caring behaviors within the college classroom. *Communication Research Reports*, 15(3):288–298, 1998.

[21] Vincent Tinto. Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45:125 – 89, 1975.

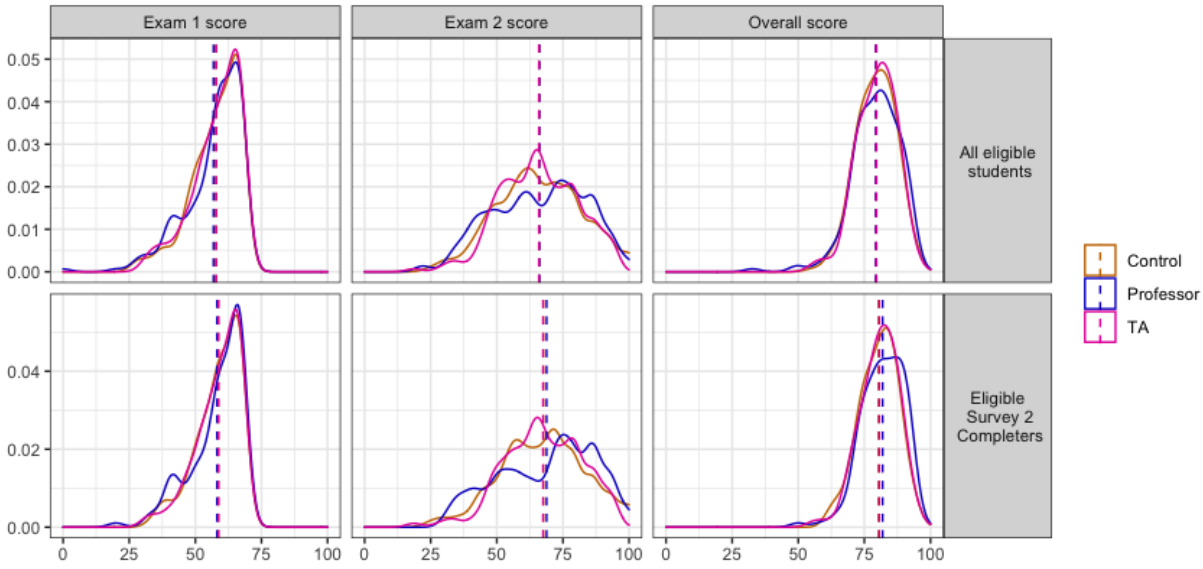
[22] Vincent Tinto. From theory to action: Exploring the institutional conditions for student retention. 2010.

8 Figures

Figure 1: Experiment Timeline

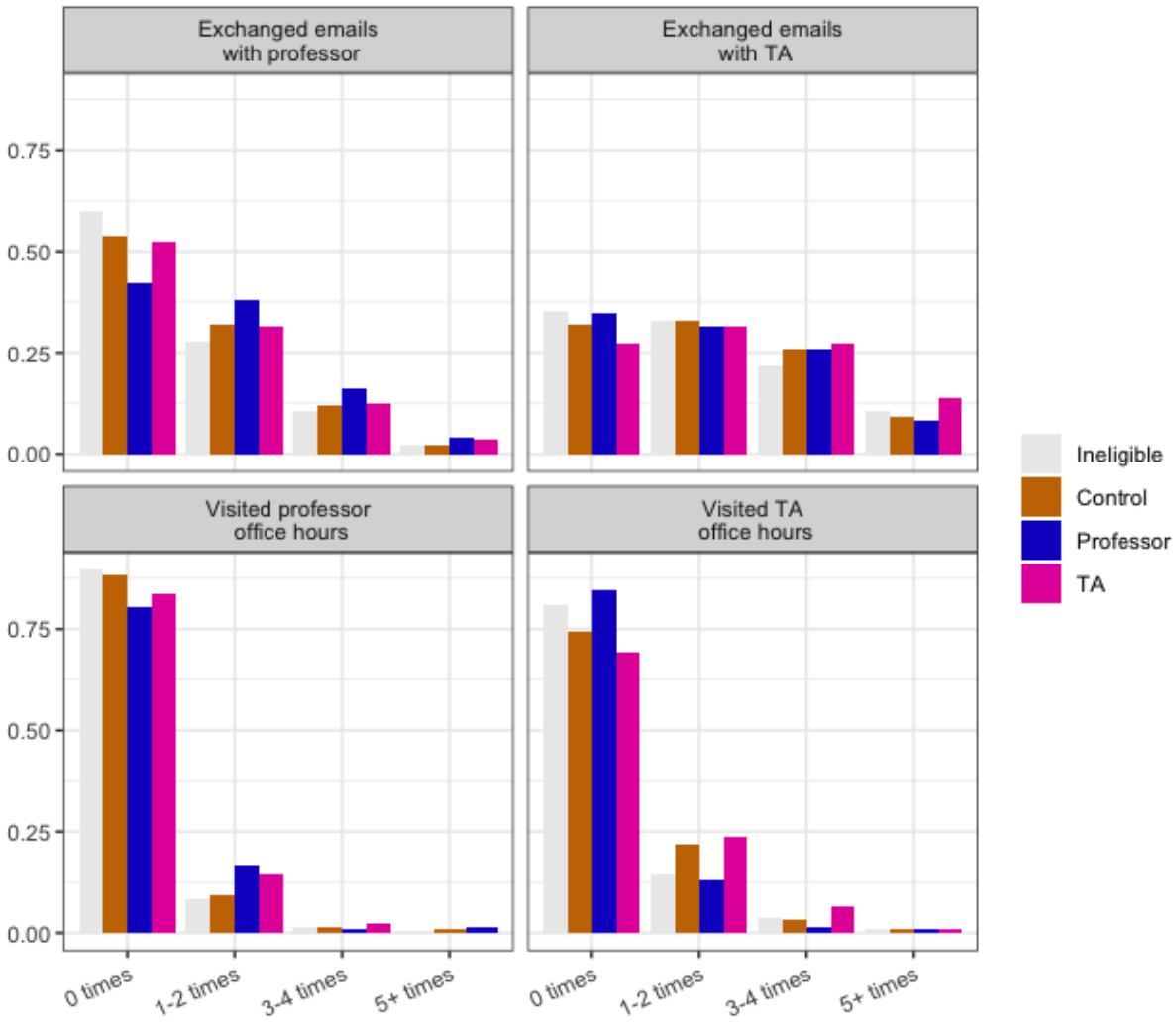
Notes: This figure shows the timing and structure of the intervention and data collection. Students were invited to take part in the study by the professors in the participating courses. Students that consented to participate in the study and scored less than 70% on the first midterm were eligible for the email intervention. Randomization to treatment status took place within discussion sections.

Figure 2: Grade distributions by treatment status



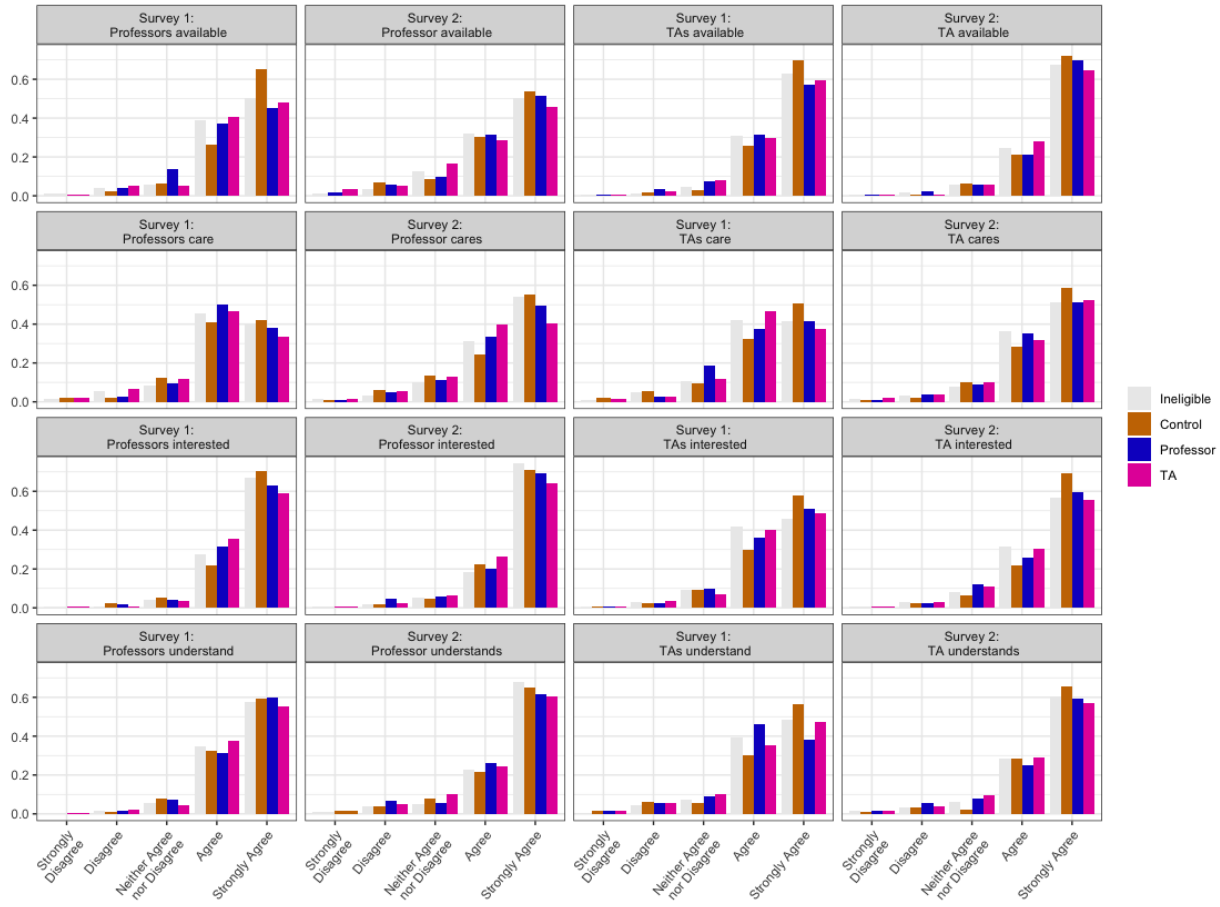
Notes: This figure shows the distribution of grades on the first and second exams, and the distribution of final scores for the class overall. Dashed lines represent the mean value for each treatment status. The densities shown are from kernel density estimation using a Gaussian kernel with a standard deviation of 3.

Figure 3: Email frequency and office hour attendance by treatment status



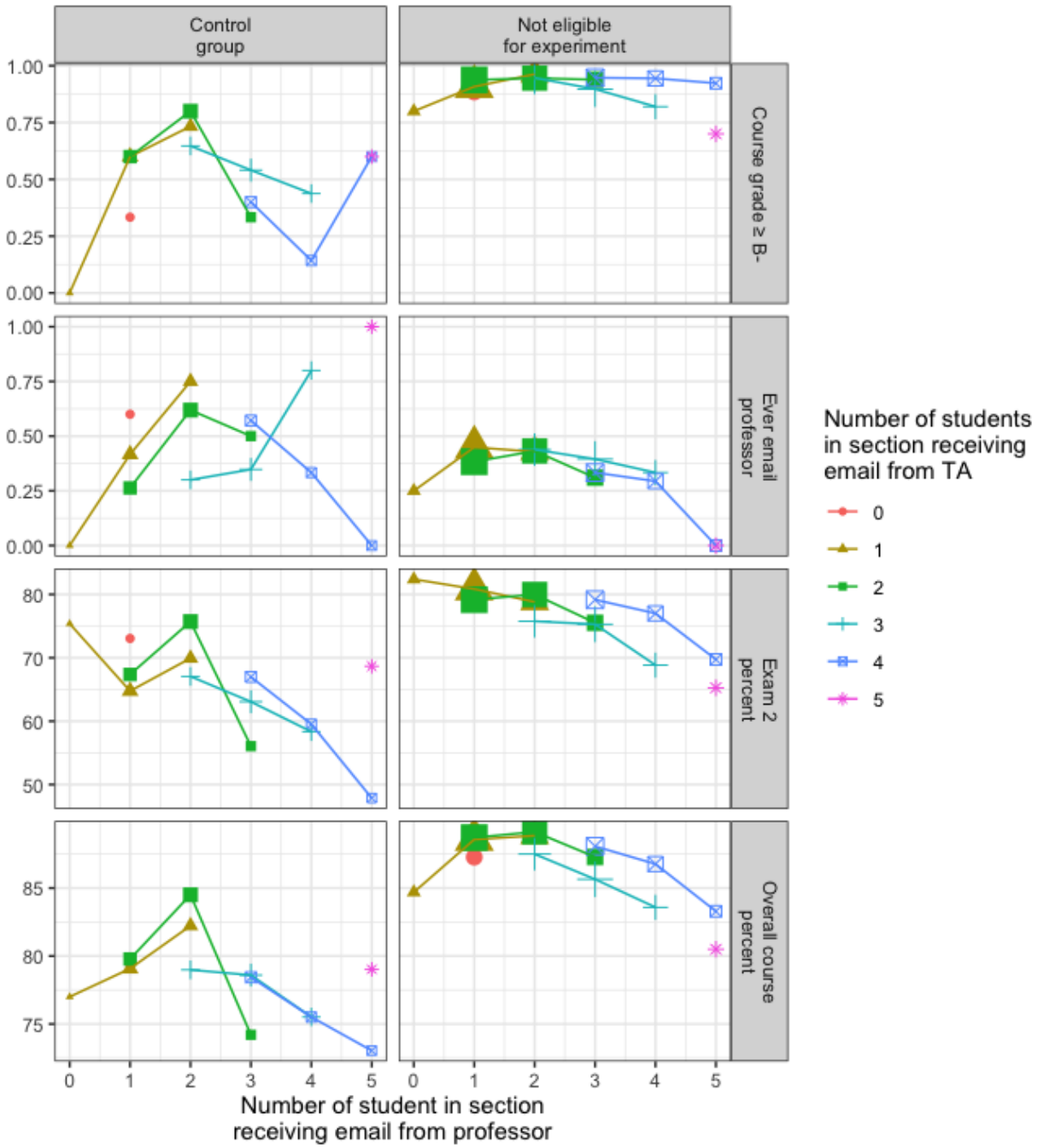
Notes: This figure shows self-reported data from Survey 2 on the frequency of email exchanges with instructors and office hour attendance. The vertical axis measures the share of students that reported emailing or attending office hours at the frequency given on the horizontal axis. The survey first asked students whether they had exchanged emails with (attended the office hours of) the instructor. Students that answered yes were then asked how many times they had exchanged emails (attended office hours).

Figure 4: Perceptions of instructors by treatment status



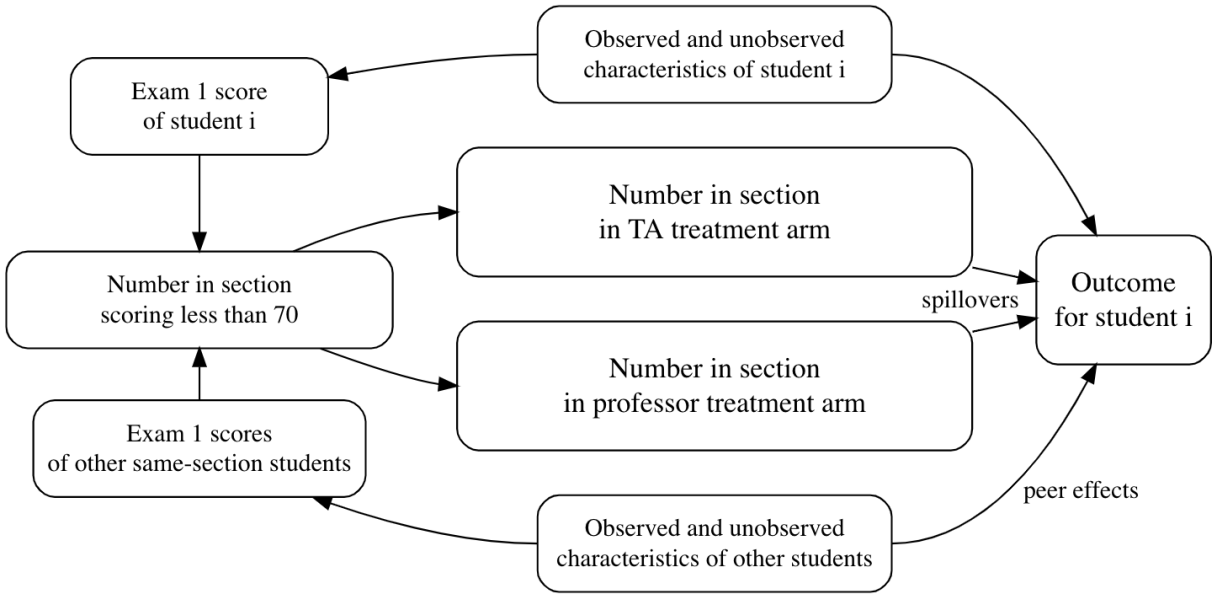
Notes: This figure shows self-reported data from Survey 2 on student perceptions of their instructors. The vertical axis measures the share of students choosing the response given on the horizontal axis. In Survey 1, students were asked to about their perceptions of current and former professors and TAs overall at the University of Iowa. In Survey 2 students were asked for their perceptions of their professor and TA for the lecture class participating in the experiment.

Figure 5: Student outcomes as a function of the number of classmates in each treatment arm



Notes: This figure shows average outcomes for untreated students by number of treated peers. In each panel, the horizontal axis measures the number of students in the same discussion section that received an email from the professor. Each line shows how average outcomes for the untreated students change as the number of same-discussion students in the professor treatment arm increases, while holding constant the number of same-discussion students in the TA treatment arm.

Figure 6: Causal paths of spillovers and other peer effects for untreated students



Notes: This figure shows causal pathways from the number of students in a discussion section that are in each treatment arm to the outcomes of a untreated student in the same discussion section. Nodes connected by arrows have a causal relationship, with the arrow showing the direction of the causality.

9 Tables

Table 1: Sample Means and Balance Tests - Experimental Sample

	Means					P-values	
	All Students	Experiment Sample	No Email	Professor Email	TA Email	Professor Email	TA Email
Exam 1 Score	75.2	57.5	57.6	56.9	57.9	0.439	0.819
	(0.439)	(0.432)	(0.626)	(0.744)	(0.647)		
HS GPA $\geq$ 3.5	0.756	0.610	0.632	0.604	0.592	0.743	0.482
	(0.013)	(0.023)	(0.037)	(0.035)	(0.036)		
Female	0.431	0.426	0.411	0.427	0.441	0.808	0.566
	(0.013)	(0.023)	(0.040)	(0.035)	(0.034)		
URM	0.083	0.095	0.070	0.104	0.109	0.306	0.367
	(0.007)	(0.013)	(0.020)	(0.020)	(0.022)		
First Generation	0.184	0.230	0.256	0.249	0.185	0.891	0.147
	(0.010)	(0.019)	(0.032)	(0.030)	(0.031)		
First Year	0.366	0.331	0.294	0.348	0.351	0.327	0.147
	(0.032)	(0.036)	(0.041)	(0.042)	(0.043)		
Expect Difficulty	0.425	0.460	0.478	0.433	0.469	0.384	0.786
	(0.013)	(0.021)	(0.033)	(0.036)	(0.036)		
Professors Impression	2.133	2.131	2.277	2.085	2.029	0.155	0.078
	(0.036)	(0.064)	(0.100)	(0.109)	(0.096)		
TAs Impression	2.032	2.086	2.327	1.960	1.971	0.023	0.022
	(0.042)	(0.073)	(0.111)	(0.115)	(0.096)		
Dropped Course	0.025	0.064	0.067	0.064	0.063	0.920	0.984
	(0.004)	(0.013)	(0.018)	(0.019)	(0.019)		
Survey 2 Complete	0.675	0.633	0.614	0.613	0.671	0.781	0.251
	(0.021)	(0.028)	(0.038)	(0.042)	(0.033)		
N	1977	621	210	204	207		

Notes: Each cell provides the mean for the given measure and group. Standard errors of the mean are in parentheses. All students refers to research participants. Experiment sample refers to participants who score below 70% on the first exam. No email, professor email, and TA email indicate treatment status groups. Each sample includes all students regardless of survey 2 completion or course completion. Exam 1 score is percent score on first exam. Professors impression and TAs impression are index 0-4 index variables measuring impressions of instructors generally at the University of Iowa. All other variables are indicators for the given description. URM = underrepresented minority. Expect difficulty = expect the course to be as difficult or more difficult than other courses. Dropped course = course noncompletion. P-values are from t-tests of the treatment status coefficients in a regression of the characteristic on treatment status dummies and strata fixed effects. Strata = discussion sections.

Table 2: Sample Means and Balance Tests - Survey 2 Sample

	Means					P-values	
	All Students	Experiment Sample	No Email	Professor Email	TA Email	Professor Email	TA Email
Exam 1 score	76.3	58.6	58.6	58.2	58.8	0.704	0.825
	(0.403)	(0.451)	(0.737)	(0.897)	(0.717)		
High school GPA $\geq$ 3.5	0.796	0.657	0.683	0.686	0.606	0.959	0.331
	(0.011)	(0.025)	(.042)	(0.042)	(0.043)		
Female	0.482	0.489	0.520	0.463	0.485	0.322	0.664
	(0.014)	(0.026)	(0.045)	(0.046)	(0.044)		
URM	0.073	0.080	0.056	0.066	0.115	0.785	0.210
	(0.007)	(0.014)	(0.021)	(0.023)	(0.028)		
First generation	0.172	0.189	0.246	0.175	0.147	0.110	0.088
	(0.010)	(0.020)	(0.039)	(0.035)	(0.031)		
First year	0.456	0.413	0.344	0.471	0.424	0.227	0.136
	(0.014)	(0.026)	(0.043)	(0.046)	(0.043)		
Expected difficulty	0.433	0.508	0.571	0.438	0.511	0.075	0.774
	(0.014)	(0.026)	(0.044)	(0.045)	(0.043)		
Professors impression	2.14	2.12	2.36	2.03	1.96	0.290	0.043
	(0.040)	(0.073)	(0.121)	(0.130)	(0.125)		
TAs impression	2.01	2.03	2.34	1.86	1.90	0.178	0.110
	(0.043)	(0.081)	(0.139)	(0.146)	(0.134)		
N	1313	383	126	122	135		

Notes: Each cell provides the mean for the given measure and group. Standard errors of the mean are in parentheses. All students refers to research participants. Experiment sample refers to participants who score below 70% on the first exam. No email, professor email, and TA email indicate treatment status groups. Each sample is conditional on survey 2 completion and course completion. Experiment sample also excludes 10 respondents with a large number of missing measures. Exam 1 score is percent score on first exam. Professors impression and TAs impression are index 0-4 index variables measuring impressions of instructors generally at the University of Iowa. All other variables are indicators for the given description. URM = underrepresented minority. Expect difficulty = expect the course to be more difficult than other courses. P-values are from t-tests of the treatment status coefficients in a regression of the characteristic on treatment status dummies and strata fixed effects. Strata = discussion sections.

Table 3: Effects on Primary Outcomes

	Professor Email	TA Email	Professor Email	TA Email	Control Mean
Survey 2 sample:					
Grades:					
Overall Course Percent	1.032	0.336	0.520	0.137	80.33
	(1.011)	(1.038)	(0.890)	(0.994)	
Course Grade $\geq$ B-	0.061	0.044	0.021	0.024	0.66
	(0.059)	(0.071)	(0.058)	(0.070)	
Exam 2 Percent	0.808	0.098	-0.469	-0.508	67.32
	(2.367)	(2.029)	(2.464)	(2.488)	
Perceptions:					
Professor Quality Index	-0.136	-0.492	0.087	-0.174	2.45
	(0.237)	(0.225)	(0.257)	(0.225)	
TA Quality Index	-0.232	-0.314	-0.078	-0.085	2.66
	(0.240)	(0.268)	(0.272)	(0.279)	
Engagement Behaviors:					
Ever Email Professor	0.101	0.001	0.085	-0.036	0.45
	(0.075)	(0.068)	(0.080)	(0.077)	
Ever Attend Professor Office Hours	0.055	0.064	0.040	0.033	0.12
	(0.064)	(0.058)	(0.065)	(0.054)	
Ever Email TA	-0.090	-0.005	-0.073	-0.019	0.68
	(0.072)	(0.081)	(0.072)	(0.079)	
Ever Attend TA Office Hours	-0.130	0.084	-0.122	0.059	0.25
	(0.059)	(0.068)	(0.064)	(0.076)	
$N = 383$					
Experimental sample:					
Grades:					
Overall Course Percent	-0.141	-0.145	-0.119	-0.280	79.4
	(0.728)	(0.730)	(0.672)	(0.773)	
Course Grade $\geq$ B-	0.015	0.004	-0.010	-0.024	0.58
	(0.041)	(0.050)	(0.042)	(0.052)	
Exam 2 Percent	-0.450	-0.567	-0.444	-0.497	66.06
	(1.590)	(1.435)	(1.678)	(1.646)	
Final Exam Percent	1.660	0.519	0.962	-0.851	69.47
	(1.981)	(1.813)	(1.873)	(1.793)	
Drop/Incomplete	-0.003	0.001	0.008	0.019	0.07
	(0.030)	(0.028)	(0.030)	(0.029)	
Fail/Drop/Incomplete	0.021	0.008	0.033	0.030	0.08
	(0.034)	(0.031)	(0.030)	(0.030)	
$N = 621$					
Strata fixed effects	✓	✓	✓	✓	
Covariates			✓	✓	

Notes: Standard errors in parentheses clustered at randomization strata: discussion sections. Both samples include students scoring below 70 percent on exam 1. Survey 2 sample includes students who completed both surveys. Experimental sample includes students who completed the first survey. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester.

Table 4: Effects on Primary Outcomes: Spillovers

	Experimental Sample				Survey 2 Sample			
	N_{prof}	N_{TA}	N_{prof}	N_{TA}	N_{prof}	N_{TA}	N_{prof}	N_{TA}
Overall Course Percent	-1.229	0.968	-1.864	0.812	-0.746	0.458	-1.700	0.337
	(0.829)	(0.927)	(0.840)	(0.871)	(0.980)	(0.991)	(1.092)	(0.890)
Course Grade $\geq$ B-	-0.055	0.050	-0.109	0.037	-0.000	0.033	-0.082	0.022
	(0.045)	(0.048)	(0.045)	(0.046)	(0.071)	(0.062)	(0.072)	(0.054)
Exam 2 Percent	-3.034	2.443	-4.201	2.156	-2.720	1.806	-4.177	1.621
	(1.662)	(1.757)	(1.768)	(1.660)	(2.366)	(2.009)	(2.706)	(1.993)
Professor Quality Index					0.251	-0.135	0.258	-0.134
					(0.170)	(0.168)	(0.182)	(0.168)
TA Quality Index					0.064	-0.047	0.072	-0.046
					(0.197)	(0.164)	(0.211)	(0.165)
Ever Email Professor					0.200	-0.111	0.187	-0.113
					(0.064)	(0.057)	(0.067)	(0.056)
Ever Attend Professor Office Hours					0.006	0.040	0.011	0.040
					(0.046)	(0.058)	(0.044)	(0.059)
Ever Email TA					0.007	0.067	0.048	0.072
					(0.066)	(0.058)	(0.073)	(0.055)
Ever Attend TA Office Hours					0.147	-0.080	0.155	-0.079
					(0.056)	(0.054)	(0.065)	(0.054)
Covariates	✓		✓		✓		✓	
Section average exam 1 score			✓				✓	
N	209		209		129		129	

Notes: Point estimates are estimated effects of an additional peer who received an email from the given sender. Samples restricted to students scoring below 70 percent on the first exam who receive no email. Standard errors in parentheses clustered at randomization strata: discussion sections. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester. Section average exam 1 score indicates additional control for the average exam 1 score within the discussion section.

Table 5: Effects on Primary Outcomes: Spillovers - Non-Experimental Sample

	Non-Experimental Sample				Non-Experimental / Survey 2 Sample			
	N_{prof}	N_{TA}	N_{prof}	N_{TA}	N_{prof}	N_{TA}	N_{prof}	N_{TA}
Overall Course Percent	-0.699	-0.036	-0.664	-0.028	-0.503	-0.362	-0.471	-0.351
	(0.283)	(0.285)	(0.329)	(0.279)	(0.310)	(0.296)	(0.348)	(0.287)
Course Grade $\geq$ B-	-0.006	0.008	-0.007	0.008	0.008	-0.006	0.004	-0.007
	(0.010)	(0.009)	(0.011)	(0.009)	(0.012)	(0.011)	(0.014)	(0.011)
Exam 2 Percent	-1.611	-0.449	-1.598	-0.446	-1.619	-0.605	-1.548	-0.581
	(0.640)	(0.579)	(0.741)	(0.569)	(0.678)	(0.603)	(0.801)	(0.579)
Professor Quality Index					0.058	0.078	0.288	0.156
					(0.097)	(0.093)	(0.105)	(0.092)
TA Quality Index					0.092	-0.050	0.196	-0.015
					(0.100)	(0.088)	(0.108)	(0.086)
Ever Email Professor					-0.019	-0.024	-0.022	-0.025
					(0.026)	(0.023)	(0.029)	(0.022)
Ever Attend Professor Office Hours					0.010	-0.013	-0.010	-0.020
					(0.017)	(0.016)	(0.019)	(0.015)
Ever Email TA					0.020	-0.015	0.033	-0.011
					(0.031)	(0.027)	(0.036)	(0.026)
Ever Attend TA Office Hours					0.000	0.013	0.015	0.018
					(0.026)	(0.028)	(0.029)	(0.029)
Covariates	✓		✓		✓		✓	
Section average exam 1 score			✓				✓	
N	1301		1301		904		904	

Notes: Point estimates are estimated effects of an additional peer who received an email from the given sender. Sample: students scoring above 70 percent on the first exam and ineligible for email receipt. Standard errors in parentheses clustered at randomization strata: discussion sections. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester. Section average exam 1 score indicates additional control for the average exam 1 score within the discussion section.

Table 6: Heterogeneous Effects: First Generation Status

	First Generation			Non First Generation		
	Professor Email	TA Email	Control Mean	Professor Email	TA Email	Control Mean
Grades:						
Overall Course Percent	0.411	1.408	79.5	0.862	0.739	80.6
	(3.234)	(3.269)		(1.103)	(1.120)	
Course Grade $\geq$ B-	0.086	0.195	0.61	0.031	0.032	0.67
	(0.174)	(0.158)		(0.078)	(0.082)	
Exam 2 Percent	-4.865	-2.322	67.0	1.945	0.064	67.4
	(7.820)	(7.401)		(3.030)	(2.804)	
Perceptions:						
Professor Quality Index	-0.035	-0.839	2.87	0.000	-0.092	2.32
	(0.605)	(0.645)		(0.325)	(0.308)	
TA Quality Index	-0.795	-0.405	2.90	0.046	0.093	2.58
	(0.579)	(0.715)		(0.391)	(0.374)	
Engagement Behaviors:						
Ever Email Professor	0.049	-0.230	0.42	0.087	-0.042	0.46
	(0.210)	(0.248)		(0.090)	(0.083)	
Ever Attend Professor Office Hours	0.121	0.135	0.10	0.006	-0.029	0.13
	(0.137)	(0.149)		(0.084)	(0.061)	
Ever Email TA	-0.038	0.076	0.52	-0.139	-0.045	0.74
	(0.211)	(0.230)		(0.083)	(0.090)	
Ever Attend TA Office Hours	-0.170	-0.012	0.29	-0.076	0.090	0.24
	(0.169)	(0.194)		(0.087)	(0.091)	
Strata fixed effects				✓	✓	
Covariates	✓	✓		✓	✓	
	$N = 71$			$N = 304$		

Notes: Standard errors in parentheses clustered at randomization strata: discussion sections. Both subsamples include students scoring below 70 percent on exam 1 and completing both surveys. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester. Strata fixed effects omitted from analysis with first generation students due to small sample.

Table 7: Heterogeneous Effects: Expected Difficulty Compared to Other Courses

	Expect: More Difficult			Expect: As or Less Difficult		
	Professor Email	TA Email	Control Mean	Professor Email	TA Email	Control Mean
Grades:						
Overall Course Percent	3.236	0.719	79.6	-0.772	0.017	81.3
	(2.237)	(1.805)		(1.801)	(2.192)	
Course Grade $\geq$ B-	0.171	0.067	0.62	-0.110	-0.030	0.70
	(0.159)	(0.153)		(0.149)	(0.177)	
Exam 2 Percent	-1.553	1.025	66.8	0.261	-0.702	68.1
	(5.904)	(4.572)		(4.872)	(5.305)	
Perceptions:						
Professor Quality Index	0.145	0.040	2.49	-0.170	-0.074	2.41
	(0.652)	(0.531)		(0.482)	(0.539)	
TA Quality Index	0.136	-0.067	2.68	-0.554	0.095	2.63
	(0.702)	(0.563)		(0.510)	(0.587)	
Engagement Behaviors:						
Ever Email Professor	-0.086	-0.247	0.47	0.124	0.039	0.43
	(0.162)	(0.144)		(0.190)	(0.191)	
Ever Attend Professor Office Hours	-0.043	-0.075	0.15	0.077	0.103	0.07
	(0.135)	(0.087)		(0.101)	(0.117)	
Ever Email TA	-0.011	0.086	0.71	-0.211	-0.106	0.65
	(0.168)	(0.178)		(0.189)	(0.165)	
Ever Attend TA Office Hours	-0.189	-0.054	0.35	-0.060	0.051	0.13
	(0.147)	(0.145)		(0.137)	(0.175)	
Strata fixed effects	✓	✓		✓	✓	
Covariates	✓	✓		✓	✓	
	$N = 194$			$N = 188$		

Notes: Standard errors in parentheses clustered at randomization strata: discussion sections. Both subsamples include students scoring below 70 percent on exam 1 and completing both surveys. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester.

Table 8: Heterogeneous Effects: High School GPA

	High School GPA $\geq$ 3.5			High School GPA $<$ 3.5		
	Professor Email	TA Email	Control Mean	Professor Email	TA Email	Control Mean
Grades:						
Overall Course Percent	0.779	0.060	81.0	-4.419	-3.093	79.1
	(1.301)	(1.129)		(1.856)	(3.004)	
Course Grade $\geq$ B-	0.076	0.087	0.68	-0.367	-0.326	0.64
	(0.075)	(0.096)		(0.190)	(0.263)	
Exam 2 Percent	-1.954	-0.893	69.0	-9.280	-7.028	64.2
	(3.333)	(2.988)		(6.009)	(8.694)	
Perceptions:						
Professor Quality Index	0.071	0.178	2.21	0.074	-0.893	3.03
	(0.407)	(0.373)		(0.823)	(0.653)	
TA Quality Index	0.016	0.155	2.52	-0.914	-1.024	2.97
	(0.383)	(0.465)		(0.857)	(0.927)	
Engagement Behaviors:						
Ever Email Professor	0.066	-0.153	0.45	-0.008	-0.427	0.49
	(0.104)	(0.117)		(0.302)	(0.283)	
Ever Attend Professor Office Hours	0.110	-0.018	0.13	-0.077	0.028	0.10
	(0.083)	(0.068)		(0.179)	(0.243)	
Ever Email TA	0.014	0.096	0.62	-0.308	-0.142	0.82
	(0.102)	(0.131)		(0.221)	(0.273)	
Ever Attend TA Office Hours	-0.122	-0.027	0.26	-0.115	0.055	0.23
	(0.099)	(0.115)		(0.214)	(0.239)	
Strata fixed effects	✓	✓		✓	✓	
Covariates	✓	✓		✓	✓	
	$N = 247$			$N = 129$		

Notes: Standard errors in parentheses clustered at randomization strata: discussion sections. Both subsamples include students scoring below 70 percent on exam 1 and completing both surveys. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester.

Table 9: Heterogeneous Effects: High-Rated vs. Low-Rated TA

	High-Rated TA			Low-Rated TA		
	Professor Email	TA Email	Control Mean	Professor Email	TA Email	Control Mean
Grades:						
Overall Course Percent	1.694	1.303	81.2	-0.132	-0.250	79.6
	(1.630)	(1.434)		(1.059)	(1.363)	
Course Grade $\geq$ B-	0.141	0.031	0.69	-0.047	0.037	0.63
	(0.089)	(0.089)		(0.084)	(0.098)	
Exam 2 Percent	0.491	1.381	71.2	-1.674	-1.656	63.7
	(4.209)	(3.060)		(3.279)	(3.684)	
Perceptions:						
Professor Quality Index	-0.038	-0.205	2.38	-0.186	-0.423	2.52
	(0.418)	(0.392)		(0.375)	(0.274)	
TA Quality Index	0.108	0.088	2.90	-0.503	-0.308	2.43
	(0.392)	(0.519)		(0.439)	(0.336)	
Engagement Behaviors:						
Ever Email Professor	0.017	-0.087	0.48	0.158	-0.003	0.43
	(0.118)	(0.135)		(0.110)	(0.120)	
Ever Attend Professor Office Hours	0.061	0.043	0.11	0.017	0.047	0.12
	(0.121)	(0.073)		(0.085)	(0.082)	
Ever Email TA	-0.027	0.065	0.67	-0.141	-0.126	0.69
	(0.105)	(0.104)		(0.108)	(0.128)	
Ever Attend TA Office Hours	-0.107	0.099	0.30	-0.133	0.094	0.22
	(0.116)	(0.125)		(0.095)	(0.110)	
Strata fixed effects	✓	✓		✓	✓	
Covariates	✓	✓		✓	✓	
	$N = 180$			$N = 202$		

Notes: Standard errors in parentheses clustered at randomization strata: discussion sections. Both subsamples include students scoring below 70 percent on exam 1 and completing both surveys. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester.

Table 10: Heterogeneous Effects: Online vs. In-Person Discussions

	Online Discussion			In-Person Discussion		
	Professor Email	TA Email	Control Mean	Professor Email	TA Email	Control Mean
Grades:						
Overall Course Percent	0.211	0.287	78.1	2.085	0.951	81.9
	(1.505)	(1.838)		(1.133)	(1.206)	
Course Grade $\geq$ B-	-0.057	0.035	0.53	0.153	0.079	0.75
	(0.139)	(0.149)		(0.056)	(0.072)	
Exam 2 Percent	0.638	0.247	62.9	1.489	0.352	70.4
	(4.512)	(5.196)		(3.072)	(2.782)	
Perceptions:						
Professor Quality Index	1.055	0.111	2.63	-0.124	-0.202	2.33
	(0.356)	(0.399)		(0.326)	(0.312)	
TA Quality Index	-0.063	-0.136	2.86	0.101	0.105	2.52
	(0.531)	(0.546)		(0.351)	(0.390)	
Engagement Behaviors:						
Ever Email Professor	0.193	-0.165	0.47	0.141	0.046	0.44
	(0.133)	(0.124)		(0.088)	(0.093)	
Ever Attend Professor Office Hours	-0.018	0.003	0.12	0.088	0.028	0.12
	(0.095)	(0.099)		(0.088)	(0.063)	
Ever Email TA	-0.202	-0.117	0.76	-0.019	0.013	0.63
	(0.135)	(0.148)		(0.097)	(0.104)	
Ever Attend TA Office Hours	-0.110	0.062	0.27	-0.110	0.050	0.24
	(0.141)	(0.159)		(0.091)	(0.089)	
Strata fixed effects	✓	✓		✓	✓	
Covariates	✓	✓		✓	✓	
	$N = 138$			$N = 244$		

Notes: Standard errors in parentheses clustered at randomization strata: discussion sections. Both subsamples include students scoring below 70 percent on exam 1 and completing both surveys. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester.

Table 11: Effects: Exploratory Analysis

	Professor Email	TA Email	Professor Email	TA Email	Control Mean
Correct: Professor Office Days	-0.063	0.015	-0.048	0.041	0.33
	(0.077)	(0.069)	(0.075)	(0.071)	
Correct: TA Office Days	-0.104	-0.019	-0.094	-0.000	0.31
	(0.066)	(0.076)	(0.068)	(0.081)	
More Courses	-0.021	-0.001	-0.033	-0.024	0.5
	(0.067)	(0.072)	(0.071)	(0.079)	
Study Hours ≥ 5	-0.050	0.045	-0.048	0.032	0.67
	(0.080)	(0.077)	(0.089)	(0.082)	
Study Hours ≥ 8	0.001	0.041	0.034	0.025	0.16
	(0.054)	(0.052)	(0.063)	(0.061)	
$N = 383$					
Strata fixed effects	✓	✓	✓	✓	
Covariates			✓	✓	

Notes: Standard errors in parentheses clustered at randomization strata: discussion sections. Sample = Survey 2 sample: students who completed both surveys. Correct variables indicate knowledge of office hour schedules. More courses indicate desire to enroll for more courses in the discipline in the future. Study hours variables indicate self reports of study hours per week above the threshold. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester.

Table 12: Heterogeneous Effects: Same-Sex vs. Opposite-Sex TA

	Same-Sex TA			Opposite-Sex TA		
	Professor Email	TA Email	Control Mean	Professor Email	TA Email	Control Mean
Grades:						
Overall Course Percent	1.010	2.169	81.55	1.331	-1.478	79.38
	(1.947)	(1.968)		(2.144)	(2.220)	
Course Grade $\geq$ B-	-0.005	0.089	0.77	0.045	-0.091	0.58
	(0.137)	(0.132)		(0.135)	(0.152)	
Exam 2 Percent	3.919	3.452	69.86	-0.349	1.034	65.16
	(6.560)	(5.324)		(5.041)	(5.283)	
Perceptions:						
Professor Quality Index	0.082	0.383	2.29	0.181	-0.652	2.55
	(0.711)	(0.664)		(0.633)	(0.502)	
TA Quality Index	0.329	0.750	2.67	0.094	-0.836	2.63
	(0.769)	(0.767)		(0.537)	(0.490)	
Engagement Behaviors:						
Ever Email Professor	0.111	-0.025	0.44	0.067	-0.125	0.47
	(0.209)	(0.197)		(0.180)	(0.166)	
Ever Attend Professor Office Hours	-0.066	0.059	0.1	-0.026	-0.100	0.14
	(0.135)	(0.106)		(0.123)	(0.112)	
Ever Email TA	-0.177	-0.002	0.73	-0.124	-0.093	0.66
	(0.195)	(0.182)		(0.155)	(0.150)	
Ever Attend TA Office Hours	-0.175	0.051	0.27	-0.060	0.027	0.23
	(0.167)	(0.213)		(0.176)	(0.162)	
Strata fixed effects	✓	✓		✓	✓	
Covariates	✓	✓		✓	✓	
	$N = 178$			$N = 200$		

Notes: Standard errors in parentheses clustered at randomization strata: discussion sections. Both subsamples include students scoring below 70 percent on exam 1 and completing both surveys. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester.

A Email example

Figure 7: Example of Professor and TA Email

Professor Email

Hi Tim,

I am emailing you after looking over the results of the first exam. You scored 52.5%, which puts you on track for a C- in the class. If you are aiming for a higher grade, there is still time to improve your score.

I want to remind you of my availability if you ever have questions. I hold virtual office hours on Mondays, Wednesdays, and Fridays from 10:00am-11:00am via Zoom. You must make a reservation on ICON to attend my virtual office hours. Instructions are included in the syllabus. At or before the time of your reservation, go to https://uiowa.zoom.us/j/97483382053.

Your TA, Jane, is also available through office hours on Tuesdays from 11:00am-12:00pm at https://uiowa.zoom.us/j/95163939950 and on Wednesdays from 12:30pm-1:30pm at https://uiowa.zoom.us/j/99075417977.

TA Email

Hi Tim,

I am emailing you after looking over the results of the first exam. You scored 52.5%, which puts you on track for a C- in the class. If you are aiming for a higher grade, there is still time to improve your score.

I want to remind you of my availability if you ever have questions. I hold virtual office hours on Tuesdays from 11:00am-12:00pm at https://uiowa.zoom.us/j/95163939950 and Wednesdays from 12:30pm-1:30pm at https://uiowa.zoom.us/j/99075417977.

Professor Smith is also available through office hours on Mondays, Wednesdays, and Fridays from 10:00am-11:00am via Zoom. You must make a reservation on ICON to attend her virtual office hours. Instructions are included in the syllabus. At or before the time of your reservation, go to https://uiowa.zoom.us/j/97483382053.

B Balance

Table 13: Sample Means and Balance Tests Categorical - Part 1

	Means					P-values	
	All Students	Experiment Sample	No Email	Professor Email	TA Email	Professor Email	TA Email
Female	0.516	0.504	0.476	0.533	0.504	0.339	0.739
	(0.015)	(0.023)	(0.046)	(0.045)	(0.043)		
Male	0.478	0.483	0.516	0.459	0.474	0.292	0.560
	(0.015)	(0.024)	(0.047)	(0.045)	(0.043)		
Missing sex	0.005	0.013	0.008	0.008	0.022	0.574	0.263
	(0.002)	(0.006)	(0.008)	(0.008)	(0.013)		
Underrepresented minority	0.914	0.903	0.937	0.926	0.852	0.756	0.090
	(0.009)	(0.016)	(0.023)	(0.023)	(0.030)		
Nonminority	0.072	0.078	0.056	0.066	0.111	0.815	0.215
	(0.008)	(0.015)	(0.020)	(0.022)	(0.027)		
Missing race/ethnicity	0.014	0.018	0.008	0.008	0.037	0.771	0.215
	(0.003)	(0.007)	(0.008)	(0.008)	(0.016)		
First generation	0.818	0.794	0.754	0.811	0.815	0.129	0.287
	(0.010)	(0.020)	(0.038)	(0.032)	(0.033)		
Non-first generation	0.167	0.185	0.246	0.172	0.141	0.099	0.063
	(0.010)	(0.019)	(0.038)	(0.032)	(0.031)		
Missing first generation	0.015	0.021	0.000	0.016	0.044	0.475	0.032
	(0.003)	(0.008)	(0.000)	(0.011)	(0.017)		
First year	0.442	0.402	0.333	0.459	0.415	0.274	0.187
	(0.032)	(0.039)	(0.054)	(0.050)	(0.049)		
Second year	0.362	0.389	0.413	0.385	0.370	0.880	0.387
	(0.024)	(0.034)	(0.051)	(0.049)	(0.045)		
Third year	0.121	0.141	0.159	0.115	0.148	0.300	0.744
	(0.012)	(0.021)	(0.038)	(0.031)	(0.030)		
Fourth year	0.034	0.029	0.048	0.008	0.030	0.142	0.567
	(0.006)	(0.008)	(0.019)	(0.008)	(0.015)		
Fifth year	0.009	0.013	0.016	0.008	0.015	0.762	0.876
	(0.002)	(0.006)	(0.011)	(0.008)	(0.010)		
Missing year	0.033	0.026	0.032	0.025	0.022	0.610	0.191
	(0.013)	(0.011)	(0.016)	(0.014)	(0.013)		
Score exam 1	76.32	58.55	58.62	58.19	58.82	0.704	0.825
	(0.477)	(0.461)	(0.728)	(0.849)	(0.675)		
HS GPA 0-2.49	0.001	0.003	0.008	0.000	0.000	0.397	0.397
	(0.001)	(0.003)	(0.008)	(0.000)	(0.000)		
HS GPA 2.5-2.99	0.028	0.060	0.071	0.033	0.074	0.526	0.660
	(0.005)	(0.012)	(0.025)	(0.016)	(0.023)		
HS GPA 3.0-3.49	0.175	0.274	0.230	0.279	0.311	0.605	0.380
	(0.012)	(0.024)	(0.039)	(0.040)	(0.041)		
HS GPA 3.5-3.99	0.466	0.483	0.540	0.492	0.422	0.908	0.301
	(0.013)	(0.024)	(0.046)	(0.045)	(0.042)		
HS GPA 4.0	0.318	0.162	0.127	0.189	0.170	0.711	0.741
	(0.016)	(0.021)	(0.030)	(0.038)	(0.032)		
Missing HS GPA	0.012	0.018	0.024	0.008	0.022	0.477	0.952
	(0.003)	(0.007)	(0.014)	(0.008)	(0.012)		
N	1313	383	126	122	135		

Notes: Each cell provides the mean for the given measure and group. Standard errors of the mean are in parentheses. All students refers to research participants. Experiment sample refers to participants who score below 70% on the first exam. No email, professor email, and TA email indicate treatment status groups. Each sample includes all students regardless of survey 2 completion or course completion. Each variable indicates a given survey response. P-values are from t-tests of the treatment status coefficients in a regression of the characteristic on treatment status dummies and strata fixed effects. Strata = discussion sections.

Table 14: Sample Means and Balance Tests Categorical - Part 2

	Means					P-values	
	All Students	Experiment Sample	No Email	Professor Email	TA Email	Professor Email	TA Email
Expect: much easier	0.022	0.005	0.008	0.008	0.000	0.593	0.294
	(0.005)	(0.004)	(0.008)	(0.008)	(0.000)		
Expect: a little easier	0.118	0.081	0.079	0.090	0.074	0.783	0.939
	(0.010)	(0.014)	(0.024)	(0.028)	(0.022)		
Expect: average difficulty	0.428	0.405	0.341	0.459	0.415	0.083	0.634
	(0.014)	(0.026)	(0.040)	(0.045)	(0.045)		
Expect: a little more difficult	0.362	0.431	0.452	0.369	0.467	0.135	0.768
	(0.014)	(0.027)	(0.043)	(0.045)	(0.046)		
Expect: much more difficult	0.066	0.076	0.119	0.066	0.044	0.416	0.312
	(0.007)	(0.012)	(0.027)	(0.023)	(0.018)		
Missing expected difficulty	0.003	0.003	0.000	0.008	0.000	0.390	0.389
	(0.002)	(0.003)	(0.000)	(0.008)	(0.000)		
Study time: 0-5 hours	0.295	0.324	0.341	0.393	0.244	0.561	0.117
	(0.013)	(0.021)	(0.039)	(0.040)	(0.035)		
Study time: 5-8 hours	0.533	0.501	0.500	0.451	0.548	0.594	0.423
	(0.015)	(0.025)	(0.043)	(0.044)	(0.041)		
Study time: 8-10 hours	0.137	0.154	0.135	0.131	0.193	0.975	0.285
	(0.009)	(0.018)	(0.031)	(0.032)	(0.033)		
Study time: more than 10 hours	0.003	0.003	0.000	0.008	0.000	0.390	0.389
	(0.002)	(0.003)	(0.000)	(0.008)	(0.000)		
Missing expected study time	0.032	0.018	0.024	0.016	0.015	0.883	0.636
	(0.005)	(0.007)	(0.013)	(0.012)	(0.010)		
Professors index = 0	0.173	0.162	0.103	0.180	0.200	0.415	0.052
	(0.011)	(0.020)	(0.027)	(0.033)	(0.036)		
Professors index = 1	0.185	0.209	0.206	0.205	0.215	0.866	0.796
	(0.010)	(0.020)	(0.037)	(0.036)	(0.034)		
Professors index = 2	0.200	0.188	0.167	0.205	0.193	0.423	0.802
	(0.010)	(0.020)	(0.033)	(0.039)	(0.035)		
Professors index = 3	0.184	0.193	0.246	0.172	0.163	0.110	0.176
	(0.012)	(0.022)	(0.038)	(0.034)	(0.031)		
Professors index = 4	0.248	0.232	0.262	0.230	0.207	0.746	0.236
	(0.014)	(0.025)	(0.039)	(0.038)	(0.033)		
Missing professors index	0.009	0.016	0.016	0.008	0.022	0.930	0.549
	(0.003)	(0.006)	(0.011)	(0.008)	(0.013)		
TAs index = 0	0.256	0.258	0.183	0.303	0.289	0.090	0.140
	(0.014)	(0.023)	(0.038)	(0.042)	(0.038)		
TAs index = 1	0.165	0.159	0.167	0.172	0.141	0.681	0.547
	(0.011)	(0.019)	(0.034)	(0.034)	(0.032)		
TAs index = 2	0.148	0.138	0.127	0.115	0.170	0.738	0.494
	(0.010)	(0.020)	(0.029)	(0.029)	(0.034)		
TAs index = 3	0.148	0.159	0.159	0.148	0.170	0.663	0.807
	(0.009)	(0.019)	(0.030)	(0.036)	(0.034)		
TAs index = 4	0.273	0.274	0.357	0.246	0.222	0.393	0.081
	(0.014)	(0.029)	(0.047)	(0.040)	(0.037)		
Missing TAs index	0.009	0.010	0.008	0.016	0.007	0.517	0.826
	(0.003)	(0.005)	(0.008)	(0.011)	(0.007)		
N	1313	383	126	122	135		

Notes: Each cell provides the mean for the given measure and group. Standard errors of the mean are in parentheses. All students refers to research participants. Experiment sample refers to participants who score below 70% on the first exam. No email, professor email, and TA email indicate treatment status groups. Each sample includes all students regardless of survey 2 completion or course completion. Index variables count number of high scores given to instructors. Each other variable indicates a given survey response. P-values are from t-tests of the treatment status coefficients in a regression of the characteristic on treatment status dummies and strata fixed effects. Strata = discussion sections.

C Surveys

The beginning of semester survey was distributed during the first month of each semester.

Figure 8: Caption

Beginning of Semester Survey - Micro

Start of Block: Consent

Q4 Are you age 18 or older?

- Yes (1)
- No (2)

Skip To: End of Survey If Are you age 18 or older? = No

Q23
We invite you to participate in a research study. The purpose of the study is to understand the effects that communication with course instructors can have on students' experience in a large, online or hybrid class.
We are inviting you to be in this study because you are an undergraduate student at the University of Iowa enrolled in ECON 1100: Principles of Microeconomics. Approximately 1700 people will take part in this study at the University of Iowa.
If you agree to participate, we would like you to complete two surveys. The surveys ask for your demographic information, perceptions of the course, and other details related to courses at the University of Iowa. Each survey will take roughly 5 to 10 minutes. The link to the second survey will be provided near the end of the semester. If you do not wish to participate, do not complete either survey. You are free to not answer any of the questions. If you agree to participate, you may or may not receive different types of communication from your instructor. Your participation in the experiment will not provideyou with any information that is not already available to you in other formats, such as the course website. If you choose to participate, we will access educational records related to your performance in Principles of Microeconomics.
We will keep the information you provide confidential, however federal regulatory agencies and the University of Iowa Institutional Review Board (a committee that reviews and approves research studies) may inspect and copy records pertaining to this research. We will collect your student ID number and retain it for identification purposes for the duration of the semester. At the conclusion of the semester, your student ID number and other personally identifiable information will be permanently deleted from the data. These data will only be used for the research purpose described herein. These data will not be used for any further research. If we write a report about this study we will do so in such a way that you cannot be identified.
You may experience minor and temporary emotional or psychological discomfort from your participation in the study. There are no known long-term risks to participating in this study. You will not benefit personally from your participation. However, we hope that others may benefit in the future from what we learn as a result of this study.

Page 1 of 9

You will not have any costs for being in this research study.
You will not be paid for being in this research study.
You will be rewarded with extra credit in the course for the completion of both surveys. The extra credit is an additional two percentage points on your final exam grade. If you complete only the first survey, you will receive one additional percentage point on your final exam grade. If you do not wish to take part in the study, an alternative method of earning the equivalent amount of extra credit is described in the email that included the link to this survey.
Taking part in this research study is completely voluntary. If you decide not to be in this study, or if you stop participating at any time, you won't be penalized or lose any benefits for which you otherwise qualify. If you ever wish to discontinue your participation in the study, please contact Darren Page at darren-page@uiowa.edu.
If you have any questions about the research study itself, please contact Darren Page at darren-page@uiowa.edu. If you experience a research-related injury, please contact: Darren Page at darren-page@uiowa.edu. If you have questions about the rights of research subjects, please contact the Human Subjects Office, 105 Hardin Library for the Health Sciences, 600 Newton Rd, The University of Iowa, Iowa City, IA 52242-1098, (319) 335-6564, or e-mail irb@uiowa.edu. To offer input about your experiences as a research subject or to speak to someone other than the research staff, call the Human Subjects Office at the number above.
This document is not a contract. It is a written explanation of what will happen during the study if you decide to participate. Your signature indicates that this research study has been explained to you, that your questions have been answered, and that you agree to take part in this study. Your signature also indicates that you authorize the research team to access educational records related to your performance in Principles of Microeconomics. Please print or save a copy of this form for your records.
Sincerely,Darren PagePhD Candidate, University of Iowa

Q24 Signature

End of Block: Consent

Start of Block: ID

Q3 Please enter your student ID number. (This is necessary to make sure you get credit for completing the survey and to link your responses to your educational records. We will not ask for identifiable information again.)

_____

Q26 Please enter the last three digits of your phone number. (We will ask you the same question on the second survey. This helps keep your survey responses anonymous.)

_____

Q27 Please enter the first three letters of your mother's maiden name. (We will ask you the same question on the second survey. This helps keep your survey responses anonymous.)

_____

End of Block: ID

Start of Block: Block 1

Q1 In comparison to other classes you are taking this semester, you expect this course to be

- ○ much more difficult (1)
- ○ a little more difficult (2)
- ○ average difficulty (3)
- ○ a little easier (4)
- ○ much easier (5)

Q2 On average, how many hours each week do you expect to spend attending (in-person or virtually), studying and working on assignments for this class?

- ○ 0-5 hours (1)
- ○ 5-8 hours (2)
- ○ 8-10 hours (3)
- ○ More than 10 hours (4)

End of Block: Block 1

Start of Block: Block 2

Q5 The following questions are about your perceptions of current and former professors overall at the University of Iowa.

Q6 My professors care whether or not I learn the material.

- Strongly agree (1)
- Somewhat agree (2)
- Neither agree nor disagree (3)
- Somewhat disagree (4)
- Strongly disagree (5)

Q8 My professors are available to talk to me individually and answer my questions.

- Strongly agree (1)
- Somewhat agree (2)
- Neither agree nor disagree (3)
- Somewhat disagree (4)
- Strongly disagree (5)

Q9 My professors are interested in teaching their courses.

- Strongly agree (1)
- Somewhat agree (2)
- Neither agree nor disagree (3)
- Somewhat disagree (4)
- Strongly disagree (5)

Q10 My professors understand the material and are able to communicate it clearly.

- Strongly agree (1)
- Somewhat agree (2)
- Neither agree nor disagree (3)
- Somewhat disagree (4)
- Strongly disagree (5)

Q11 The following questions are about your perceptions of current and former TAs overall at the University of Iowa.

Q12 My TAs care whether or not I learn the material.

- ○ Strongly agree (1)
- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)
- ○ Somewhat disagree (4)
- ○ Strongly disagree (5)

Q13 My TAs are available to talk to me individually and answer my questions.

- ○ Strongly agree (1)
- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)
- ○ Somewhat disagree (4)
- ○ Strongly disagree (5)

Q14 My TAs are interested in teaching their courses.

- ○ Strongly agree (1)
- ○ Somewhat agree (2)
- ○ Neither agree nor disagree (3)
- ○ Somewhat disagree (4)
- ○ Strongly disagree (5)

Q15 My TAs understand the material and are able to communicate it clearly.

- Strongly agree (1)
- Somewhat agree (2)
- Neither agree nor disagree (3)
- Somewhat disagree (4)
- Strongly disagree (5)

End of Block: Block 2

Start of Block: Block 3

Q16 These final questions ask for your demographic information. We ask these questions so that we can understand how online and hybrid classes might disproportionately affect different groups of students. For as many questions as you want, you can select "prefer not to answer".

Q17 What year in school are you?

- First year (1)
- Second year (2)
- Third year (3)
- Fourth year (4)
- Fifth year or more (5)
- Prefer not to answer (6)

Q18 What is your sex?

- ○ Male (1)
- ○ Female (2)
- ○ Prefer not to answer (3)

Q19 What is your race? Select all that apply.

- ☐ Asian (1)
- ☐ Black or African-American (2)
- ☐ Middle Eastern or North African (3)
- ☐ White (4)
- ☐ Other (5)
- ☐ Prefer not to answer (6)

Q20 Are you of Hispanic, Latino, or Spanish origin?

- ○ Yes (1)
- ○ No (2)
- ○ Prefer not to answer (3)

Q21 Are you a first-generation student?

- Yes (1)
- No (2)
- Prefer not to answer (3)

Q28 What was your high school GPA?

- 4.0 or higher (1)
- 3.5-3.99 (2)
- 3.0-3.49 (3)
- 2.5-2.99 (4)
- 0-2.49 (5)
- Prefer not to answer (6)

Q22 Are you an international student?

- Yes (1)
- No (2)
- Prefer not to answer (3)

End of Block: Block 3

Page 9 of 9

D Additional tables

Table 15: Heterogeneous Effects: Online vs. In-Person Discussions

	Online Discussion			In-Person Discussion		
	Professor Email	TA Email	Control Mean	Professor Email	TA Email	Control Mean
Grades:						
Overall Course Percent	-0.420	-0.671	76.8	0.890	0.769	81.8
	(0.724)	(1.080)		(1.059)	(1.097)	
Course Grade $\geq$ B-	-0.047	-0.054	0.43	0.065	0.030	0.72
	(0.062)	(0.085)		(0.054)	(0.067)	
Exam 2 Percent	-0.079	-0.677	60.9	1.013	0.552	70.9
	(2.126)	(2.414)		(2.431)	(2.351)	
Final Exam Percent	2.262	-1.073	65.2	1.016	0.342	73.4
	(2.777)	(2.698)		(2.700)	(2.697)	
Drop/Incomplete	0.022	0.084	0.09	-0.018	-0.033	0.05
	(0.052)	(0.048)		(0.024)	(0.021)	
Fail/Drop/Incomplete	0.043	0.085	0.10	0.010	-0.011	0.07
	(0.052)	(0.048)		(0.030)	(0.031)	
Strata fixed effects	✓	✓		✓	✓	
Covariates	✓	✓		✓	✓	
	$N = 293$			$N = 326$		

Notes: Standard errors in parentheses clustered at randomization strata: discussion sections. Both subsamples include students scoring below 70 percent on exam 1. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester.

Table 16: Heterogeneous Effects: High-Rated vs. Low-Rated TA

	High-Rated TA			Low-Rated TA		
	Professor Email	TA Email	Control Mean	Professor Email	TA Email	Control Mean
Grades:						
Overall Course Percent	0.564	0.531	81.6	-0.322	-0.239	77.8
	(1.278)	(1.267)		(0.766)	(1.024)	
Course Grade $\geq$ B-	0.054	-0.033	0.69	-0.053	-0.021	0.50
	(0.078)	(0.075)		(0.051)	(0.073)	
Exam 2 Percent	0.016	1.798	71.9	-0.395	-0.605	61.9
	(3.144)	(2.794)		(1.955)	(1.964)	
Final Exam Percent	1.348	1.064	72.3	1.196	-0.619	67.4
	(3.141)	(3.558)		(1.956)	(2.105)	
Drop/Incomplete	-0.042	-0.028	0.05	0.026	0.052	0.08
	(0.030)	(0.029)		(0.044)	(0.041)	
Fail/Drop/Incomplete	0.005	0.007	0.07	0.046	0.050	0.09
	(0.034)	(0.037)		(0.046)	(0.042)	
Strata fixed effects	✓	✓		✓	✓	
Covariates	✓	✓		✓	✓	
	$N = 247$			$N = 372$		

Notes: Standard errors in parentheses clustered at randomization strata: discussion sections. Both subsamples include students scoring below 70 percent on exam 1. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester.

Table 17: Heterogeneous Effects: Same-Sex vs. Opposite-Sex TA

	Same-Sex TA			Opposite-Sex TA		
	Professor Email	TA Email	Control Mean	Professor Email	TA Email	Control Mean
Grades:						
Overall Course Percent	-0.875	0.285	79.9	0.879	0.070	78.8
	(1.265)	(1.272)		(1.217)	(1.392)	
Course Grade $\geq$ B-	-0.091	-0.009	0.63	0.043	-0.036	0.52
	(0.072)	(0.078)		(0.077)	(0.089)	
Exam 2 Percent	-0.650	-0.970	67.3	-0.151	1.429	64.4
	(3.642)	(2.671)		(2.554)	(2.947)	
Final Exam Percent	-1.549	-0.663	70.4	4.436	-1.864	68.8
	(3.226)	(2.678)		(3.390)	(3.200)	
Drop/Incomplete	-0.002	0.061	0.07	0.010	-0.007	0.06
	(0.043)	(0.060)		(0.054)	(0.048)	
Fail/Drop/Incomplete	-0.019	0.078	0.09	0.061	0.015	0.07
	(0.046)	(0.061)		(0.061)	(0.053)	
Strata fixed effects	✓	✓		✓	✓	
Covariates	✓	✓		✓	✓	
	$N = 299$			$N = 311$		

Notes: Standard errors in parentheses clustered at randomization strata: discussion sections. Both subsamples include students scoring below 70 percent on exam 1. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester.

Table 18: Heterogeneous Effects: First Generation Status

	First Generation			Non First Generation		
	Professor Email	TA Email	Control Mean	Professor Email	TA Email	Control Mean
Grades:						
Overall Course Percent	-0.344	2.158	77.5	0.271	-0.523	80.1
	(1.780)	(1.581)		(0.741)	(0.827)	
Course Grade $\geq$ B-	0.098	0.140	0.42	-0.046	-0.079	0.64
	(0.094)	(0.094)		(0.050)	(0.058)	
Exam 2 Percent	-1.427	2.730	63.3	1.117	-1.064	67.0
	(4.571)	(3.723)		(1.643)	(1.767)	
Final Exam Percent	2.849	6.350	63.5	1.041	-0.758	71.8
	(5.017)	(5.150)		(2.058)	(2.090)	
Drop/Incomplete	0.058	0.049	0.12	0.009	0.026	0.05
	(0.072)	(0.078)		(0.021)	(0.024)	
Fail/Drop/Incomplete	0.078	0.032	0.13	0.024	0.022	0.06
	(0.071)	(0.080)		(0.025)	(0.028)	
Strata fixed effects				✓	✓	
Covariates	✓	✓		✓	✓	
	$N = 138$			$N = 462$		

Notes: Standard errors in parentheses clustered at randomization strata: discussion sections. Both subsamples include students scoring below 70 percent on exam 1. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester. Strata fixed effects omitted from analysis with first generation students due to small sample.

Table 19: Heterogeneous Effects: Expected Difficulty Compared to Other Courses

	Expect: More Difficult			Expect: As or Less Difficult		
	Professor Email	TA Email	Control Mean	Professor Email	TA Email	Control Mean
Grades:						
Overall Course Percent	0.684	-0.390	79.7	0.155	0.361	79.1
	(1.370)	(1.512)		(1.100)	(1.389)	
Course Grade $\geq$ B-	-0.015	-0.037	0.59	-0.040	-0.004	0.56
	(0.092)	(0.098)		(0.077)	(0.103)	
Exam 2 Percent	-2.416	-1.356	67.4	1.522	1.810	64.6
	(3.713)	(3.272)		(2.719)	(2.928)	
Final Exam Percent	3.852	-0.453	69.8	-0.338	-2.134	69.2
	(3.513)	(3.558)		(2.829)	(2.901)	
Drop/Incomplete	0.033	0.044	0.05	0.000	-0.015	0.08
	(0.041)	(0.044)		(0.052)	(0.057)	
Fail/Drop/Incomplete	0.011	0.033	0.07	0.015	-0.024	0.09
	(0.043)	(0.049)		(0.055)	(0.060)	
Strata fixed effects	✓	✓		✓	✓	
Covariates	✓	✓		✓	✓	
	$N = 285$			$N = 334$		

Notes: Standard errors in parentheses clustered at randomization strata: discussion sections. Both subsamples include students scoring below 70 percent on exam 1. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester.

Table 20: Heterogeneous Effects: High School GPA

	High School GPA $\geq$ 3.5			High School GPA $<$ 3.5		
	Professor Email	TA Email	Control Mean	Professor Email	TA Email	Control Mean
Grades:						
Overall Course Percent	0.616	0.103	80.6	-2.000	-0.396	77.3
	(1.004)	(0.937)		(1.598)	(1.769)	
Course Grade $\geq$ B-	0.058	0.033	0.62	-0.196	-0.192	0.51
	(0.062)	(0.073)		(0.082)	(0.115)	
Exam 2 Percent	-0.601	-0.299	67.7	-0.219	0.983	63.1
	(2.365)	(2.217)		(4.085)	(4.183)	
Final Exam Percent	4.169	1.509	69.2	-4.483	-5.569	69.5
	(2.547)	(2.473)		(3.717)	(3.635)	
Drop/Incomplete	-0.005	0.001	0.07	0.071	0.047	0.07
	(0.042)	(0.039)		(0.066)	(0.086)	
Fail/Drop/Incomplete	0.008	0.005	0.08	0.115	0.102	0.09
	(0.046)	(0.039)		(0.069)	(0.084)	
Strata fixed effects	✓	✓		✓	✓	
Covariates	✓	✓		✓	✓	
	$N = 370$			$N = 237$		

Notes: Standard errors in parentheses clustered at randomization strata: discussion sections. Both subsamples include students scoring below 70 percent on exam 1. Quality indices have ranges from 0 to 4 measuring the number of "Strongly Agree" responses to the 4 quality perception questions. Covariates include controls for sex, race, international, first generation, year in school, high school GPA, expected course difficulty at the beginning of the semester, and general impressions of professors and teaching assistants at the University of Iowa at the beginning of the semester.