

NBER WORKING PAPER SERIES

MY PROFESSOR CARES:
EXPERIMENTAL EVIDENCE ON THE ROLE OF FACULTY ENGAGEMENT

Scott E. Carrell
Michal Kurlaender

Working Paper 27312
<http://www.nber.org/papers/w27312>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2020

This project was funded through a grant from the College Futures Foundation: <https://collegefutures.org>. We thank Joy Salvetti for her partnership and assistance, Monica Bhatt for research support and collaboration on the initial pilot, and Sherrie Reed for technical support. We also thank Ryan Bandy, Claudia Escobar, Brian Fernandez-Oquendo, Betsey Friedman, and Derek Rury for providing excellent research assistance. We are grateful for comments from participants at the numerous conferences and seminars where we have presented this project. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Scott E. Carrell and Michal Kurlaender. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

My Professor Cares: Experimental Evidence on the Role of Faculty Engagement
Scott E. Carrell and Michal Kurlaender
NBER Working Paper No. 27312
June 2020
JEL No. I20

ABSTRACT

Despite a growing body of literature that instructors “matter” in higher education, there is virtually no evidence about how their actions influence student outcomes. We provide experimental evidence on the impact of specific faculty behaviors aimed at increasing student success. We test the effect of professor feedback on student success in higher education classrooms through a "light-touch" randomized intervention. We present results from a small pilot in an introductory-level microeconomics course at a comprehensive research university, and the scale-up conducted in over 43 classrooms and nearly 4,000 students at a large broad-access university. The intervention consisted of several strategically-timed E-mails to students from the professor indicating keys to success in the class, the students' current standing in the course, and a reminder of when the professor is available. Results from the pilot show that students in the treatment group scored higher on exams, homework assignments, and final course grade. Results from the scaled-up experiment are more mixed—we find significant positive effects on student perceptions of the professor and course for all students. However, we only find positive achievement effects for our target population, first year students from underrepresented minority groups. Finally, we replicated the pilot to test the robustness of these results and again find positive effects on student achievement. We conclude that in certain settings and with some students, targeted feedback from professors can lead to meaningful gains in achievement.

Scott E. Carrell
Department of Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616
and NBER
secarrell@ucdavis.edu

Michal Kurlaender
University of California, Davis
One Shields Avenue
School of Education
Davis, CA 95616
mkurlaender@ucdavis.edu

A randomized controlled trials registry entry is available at <http://www.socialscienceregistry.org/trials/5875>

1. Introduction

The rising value of a college degree has been well documented among social scientists (Pew Research Center, 2014; Baum, Ma, & Payea, 2013), and more broadly in the popular press (Leonhardt, New York Times, 2014) and in policy efforts (Turner, 2018). However, despite increases in college attendance, college completion has not kept up (Holzer & Baum, 2017; Pew Research Center, 2014; Snyder & Dillow, 2013). Moreover, many disparities by social origin and race/ethnicity exist in college access and college completion (Holzer & Baum, 2017; Hoxby & Avery, 2013; Bailey & Dynarski, 2011). Despite a growing number of randomized control trials on improving college access, particularly for low income and other underrepresented groups (Phillips & Reber, 2019; Castelman, Page, & Schooley, 2014; Carrell & Sacerdote, 2017; Hoxby & Turner, 2013; Bettinger, Long, Oreopoulos, & Sanbonmatsu, 2012; Avery & Kane, 2004; Barr & Castleman, 2017), the research base is decidedly thin on how to keep students in college, and on improving college success and degree completion.

Interventions that have focused directly on increasing student supports for college retention and completion efforts have been met with mixed results. For example, financial incentives and need-based aid programs reveal inconsistent results on college persistence and degree receipt (Carlson et al., 2019; Anderson et al., 2018; Angrist et al. 2014; Angrist et al. 2009), as does coaching and advising (Oreopoulos & Petronijevic, 2019; Bettinger & Baker, 2014; Angrist et al.; 2009). Several psychological interventions aimed at improving students' academic mindsets and sense of belonging found positive impacts on persistence, performance, and reductions in achievement and persistence gaps by race and gender (Walton & Cohen, 2011; Yaeger et al., 2016); yet similar interventions fail to replicate in other settings (Dobronyi et al., 2019; Oreopoulos & Petronijevic, 2019). Notably, nearly all of the college persistence

interventions have been targeted at students, neglecting a potential key input in the education production function—faculty.

In fact, a growing body of literature suggests that college instructors matter (Braga et al. 2014; Carrell & West, 2010). In particular, prior work has demonstrated that demographic characteristics of professors such as gender (Carrell et al. 2010; Price, 2010; Hoffmann & Oreopoulos, 2009) and race (Fairlie et al., 2014; Price, 2010) can influence student performance and attainment in particular courses. And, instructor status (i.e. adjunct employment or academic rank) (Ran & Xu, 2018; Figlio, Schapiro, & Soter, 2015; Bettinger and Long, 2006; Carrell & West, 2010; Ehrenberg & Zhang, 2005), as well as the effectiveness of student evaluations (Braga et al., 2014; Beleche et al., 2012; Carrell & West, 2010) as predictors of both contemporaneous and longer run outcomes of students. However, prior studies on college instructors have focused almost exclusively on their innate traits, job features or unobservable characteristics. The one exception is Brownback and Sadoff (2019) who conducted a field experiment testing the impact of performance-based financial incentives for community college instructors. They find that instructor incentives significantly improved students' performance and completion in a course, and had broader spillovers for credit accumulation and transfer. This study provides evidence that instructor effectiveness in the postsecondary environment is malleable and that financial incentives—at least in the community college context—may improve instructor effectiveness. However, the literature, as a whole, leaves the question about *how* faculty could improve their effectiveness unanswered.

In this paper, we provide the first experimental evidence on the impact of specific faculty behaviors aimed at increasing student success. We test a theoretically grounded treatment designed to address a fundamental aspect of the college experience: faculty-student engagement.

Moreover, unlike the unique settings of many of the prior studies on the role of college instructors (e.g., elite universities, military academies, or community colleges), the setting for our study is a large representative broad-access four-year university campus. Thus, the study represents, to our knowledge, the first experiment in higher education aimed at inducing a pedagogical change in faculty behavior.

Specifically, we test the effect of increased and individualized professor feedback on student success. The paper presents the full development of the intervention: Exploratory qualitative work with our target population—underrepresented minority students attending a large broad-access university, the pilot phase of the intervention, two waves of full-scale implementation, and finally a replication of the pilot. The “light-touch” intervention consisted of strategically-timed personalized e-mails to students from the professor indicating the professor’s knowledge of the students’ current standing in the course, keys to success in the class, and a reminder of when the professor is available. Results from the pilot were promising-- students in the randomly selected treatment group exhibited increased effort on homework as well as a significant increase in academic achievement, motivating a scale-up of the intervention at the same large, broad-access, four-year institution where we conducted the focus groups. We implemented the intervention with 34 faculty members across 43 classrooms in 27 different course subjects and nearly 4,000 students during the spring of 2016 and fall of 2017. Results showed positive effects on students’ perceptions of instructor support for all students, and positive effects on course grades for the target population—students from underrepresented minority backgrounds, but not for all students. We examine other treatment heterogeneity, fidelity of implementation, and adjust our inference for multiple comparisons. Finally, to test the robustness of the intervention, we replicated the pilot in the spring of 2019. Results again show

significant positive effects on student achievement. We conclude that context likely matters, and that in certain settings and with some students, targeted feedback from professors can lead to meaningful gains in student achievement, particularly for those students in their first year of college.

2. Faculty Feedback and Student Engagement

2.1 Focus Groups

To explore ways to improve college success, particularly for underrepresented students in their first-year of college, we conducted a series of open-ended qualitative focus groups with African American and Latino male students at a large, broad access, four-year university in Northern California during the winter of 2014. We chose this population given the documented low six-year completion rate they experienced at that institution—less than 30 percent. The interviews focused on student experiences and struggles while in college. More specifically, we asked students to reflect on their experiences in the classroom. Two key themes emerged. First, students expressed a general lack of interaction with faculty, they found it hard to engage with instructors both in and outside of formal class times (i.e. office hours). Second, students felt unsure of what they needed to do to be more successful in their courses, something they described as very much a departure from the success they felt as high school students in their courses. In short, students did not believe most college instructors were accessible, clear about their expectations of students, or supportive of their learning.

2.2 Developing an Intervention

These focus groups inspired a pilot intervention aimed at providing personalized information and encouragement from a professor to her/his students. The intervention is “light touch” in that it requires a modest amount of extra time on the part of the faculty member to implement. We piloted the intervention to students enrolled in a large introductory course at a comprehensive university. The intervention itself consists of personalized emails from the professor at the beginning and middle of the term, providing students with specific information about the necessary steps to succeed in the course and encouragement about how to be successful in college.

The specific treatment is built upon theories from behavioral economics about information, from education on the role of feedback and student outcomes, and from social psychology on self-efficacy and affirmation. Moreover, the intervention rests upon one key premise: Faculty are an important and (potentially) under-utilized resource to increase student success more generally and retention and completion more specifically. Our hypothesis is that receiving additional information about course performance and positive directions and encouragement regarding college success can improve students’ sense of self-efficacy and influence their decision to persist towards, and ultimately complete, the degree. Moreover, we speculate that such information may be particularly valuable to students early in their college careers and to students who have been historically underrepresented at the University (e.g. Black and Latinx students).

Specifically, students in the treatment condition received two e-mails with the explicit purpose of providing information about (1) how they are progressing in the class; (2) how to be successful in the class moving forward; and (3) the availability of the professor and other

supports. The goal was to test whether these personalized messages from faculty influence short-term outcomes such as homework and midterm exam performance, and medium-run outcomes such as course completion and grade. We also tested potential mechanisms for the interpretation of results by surveying students on their perception of the professor and the course after the submission of the final exam.

At the heart of our treatment is the notion that increased and individualized information provided by faculty to students regarding their performance and direction for future work will improve their sense of self-efficacy and their course and college outcomes. We know from human capital theory that the individual decision to invest in education (i.e. persist in college) should be based on an interaction of students' resources (financial or otherwise) to enroll, tastes for the college experience, and ability to do the work. Students rely on many sources of information to make these decisions. That is, students will use information about the cost of college, their experience in college (grades, friends, etc.), and, arguably, some knowledge about the long-term benefit of having a college degree to make the optimal decision about whether to stay in school (Avery & Kane, 2004).¹ However, recent work in behavioral economics is more critical of rational choice, and posits that human behavior is more psychologically driven, suggesting that decisions are heavily influenced by factors such as how the information is conveyed, by whom, and in what context (Thaler & Sunstein, 2008). Here, we hypothesize that a small increase in information that is personalized and provided directly from students' course instructors can influence performance in that course, and ultimately, their persistence in college overall. We also conceive of the information being provided to students as a form of personalized feedback, given that it happens after faculty have some indication of student

¹ Students may display hyperbolic discounting (Laibson, 1997) in evaluating the costs and benefits of staying in college. That is, short sightedness causes them to highly discount the benefits of increased earnings, which are likely years away.

performance in the course, and that the information is specifically tailored toward students in light of their performance.

Feedback in the teaching and learning literature refers to the information provided in response to one's performance or understanding. As such, feedback is considered a "consequence of performance" (Hattie & Timperley, 2007). Empirical evidence from the literature on feedback suggests that it can be a powerful influence on achievement in the K-12 context, but that it is also highly variable. A meta-analysis found that studies showing the highest effect sizes involved students receiving feedback about a task and how to do it more effectively, while lower effect sizes were associated with feedback in the form of praise, rewards, and punishment (Hattie & Timperley, 2007; Kluger & DeNisi, 1998). Our intervention is not focused on changing how faculty grade assignments or provide feedback on specific course tasks, rather on how they can provide feedback about the processes underlying the tasks expected of students in their courses and the strategies students can incorporate to improve performance. Feedback at this "process level" has been found to be particularly effective (Balzer, Doherty, & O'Connor, 1989) and is the basis for the information that faculty in our intervention provide. Specifically, the goal is to provide feedback on how to seek help (a learned process) and how to overcome potential self-doubt or embarrassment about such help-seeking behavior (Karabenick & Knapp, 1991). In fact, a critical mediator to feedback is the perception of self-efficacy (Hattie and Timperley, 2007; Kluger & DeNisi, 1998). That is, feedback is particularly valuable if it also encourages and promotes students' sense of self-efficacy.

Although largely framed as an information and feedback intervention, our underlying theory of change suggests that this information can have important consequences for students' sense of support, self-efficacy and help-seeking behavior. Self-efficacy is a key component to

how students may handle challenging or unpredictable situations and, importantly, how much effort they may decide to expend or how long they persist in light of challenging or unpredictable situations (Bandura, 1993). Individuals' perceived sense of efficacy can influence actions indirectly, for example, by its impact on goals and aspirations, their effort and commitments to different pursuits, and how they cope with stressful situations (Steele, 1988; Bandura and Schunk, 1981). Experiments from social psychology demonstrate that accentuating positive growth rather than shortfalls enhance self-efficacy, aspirations, and performance (Bandura, 1993; Yeager and Walton, 2011). Thus, the nature of the feedback and information provided by faculty may play an important role in perceived self-efficacy and ultimately in course success.

3. Piloting the Intervention

3.1 Pilot Design

The pilot setting was a large, introductory-level microeconomics course with an initial enrollment of 420 students at a large selective comprehensive university. In this course, students are required to complete 5 of 7 homework assignments throughout the term. Data from prior years of this course indicate that failure to complete the first homework is a good early indication of struggling students.²

During the spring quarter of 2014, the research team randomized students who did not submit or failed the first homework assignment into a treatment group and control group. Students in the treatment group received a two-tiered intervention in the form of e-mails from the professor reminding them of the behaviors that lead to success in the course (attend class, complete practice problems, attend section and utilize office hours as needed), as well as a reminder of when the professor is available.

² Students who fail to complete the first homework score about 10 percentage points lower in the course, on average.

The first e-mail to the treatment group was sent as a result of failing the first homework assignment. The second e-mail to the treatment group was sent after the first midterm exam and feedback to students was based on their exam performance:

- **Group A:** Students that received a B+ or higher: E-mail text commends the student on a job well done and reminds the student of the professor's office hours availability.
- **Group B:** Students that received between a C- and B: E-mail text tells the student what their grade in the course is likely to be based on this midterm performance and highlights that it is not too late to improve their grade and the set of behaviors that will help the student be successful in the course, as well as reminds the student of the professor's office hours availability.
- **Group C:** Students that received lower than a C- on the midterm: E-mail text warns the student that based on his/her trajectory, the student may be at risk of failing the course, but reminds them there is time to recover and details the behaviors that would allow them to pass the course successfully, and reminds students to seek additional supports and the professor's office hours availability.³

During the course of the term, we tracked students' course dropout status, homework completion, time spent on homework, midterm and final exam scores, final course grades, and office hour attendance. We also asked students at the end of the class about their personal motivation to do well in the course and their perception of how much the professor cared about their performance.

³ A fourth group of five students, who had dropped out of the course from the treatment group at the time the second e-mail was sent, received no e-mails.

Data were collected via the online homework portal through which students submitted assignments, office-hour sign-in sheets, course gradebooks, and two survey questions placed on the final exam. In addition, we merged student-level data from the University registrar on student sex, underrepresented minority status, whether or not a student was a first-generation college student, high school GPA, residency status, and the year in which they entered college.

3.2 Pilot Data and Methods

Of the 69 students who did not submit the first homework assignment, 35 were assigned to the treatment group and 34 to the control group, and 16 students dropped out of the course. The sample of students overall is 68% male, 89% California residents, 26% of students are first-generation college students, 23% of students are underrepresented minorities (Table 1). The treatment and control groups are comparable with respect to student characteristics with some minor differences. For example, the treatment group has slightly more first-generation college students than the control group (28% as compared to 25%), but the control group has a greater proportion of underrepresented minority students (29%) than the control group (17%). We conduct randomization checks on the comparability of treatment and control group by regressing student characteristics on an indicator variable for treatment status (Table A1). The results indicate that student characteristics are not significantly predictive of treatment status and provides evidence that randomization created groups that were equal in expectation for receipt of the treatment.

[Insert Table 1]

The study design, random assignment of study subjects to treatment or control status, allowed for a simple analytic strategy. Specifically, we use ordinary least squares (OLS)

regression analysis to calculate the average causal treatment effect for our “light-touch” feedback intervention with the following regression:

$$Y_{is} = \alpha + \beta * treat_{is} + \gamma * X_i + \lambda_s + \varepsilon_{is}$$

Where Y represents our respective outcomes of interest for student i in section s , " $treat$ " is a dummy variable for treatment versus control status and X is a vector of individual student characteristics. In this analysis, β represents the average causal effect of the (intervention) on student outcomes. We investigate several outcomes: exam grades, total course score and grade, homework score, time spent on homework, office hour attendance, attitudinal measures towards the course and professor, and course completion.⁴

We calculate a treatment effect for each outcome variable of interest using three specifications. The first specification includes only a dummy indicator for treatment status. The second specification includes TA fixed effects (λ_s) to account for variation in teaching and learning across each of the four TAs in the course. Each student in the course was assigned to one TA and attended his/her small-group section once a week.⁵ Attendance at section was not mandatory, nor was seeking out TA assistance in office hours. The TA fixed effects are represented by a dummy indicator for each TA and allows comparisons between individuals with the same TA while eliminating between-TA differences. The third and final specification includes both TA fixed effects and student-level controls. Individual control variables include whether the student is male, first-generation college student status, under-represented minority status, California residency status, entering cohort year, and high school GPA. In addition, we conducted a randomization check by regressing individual student characteristics on treatment

⁴ For analyzing treatment effects on survey questions “The professor cares about my performance” and “I am motivated to do well in the course” we use a probit model that accounts for a binomial outcome.

⁵ Importantly, students do not choose their TA as the TA’s are assigned to sections after the student’s primary registration period ends.

status both with and without TA fixed effects. Results showed that there are no statistically significant relationships between individual characteristics and treatment status.

3.3 Pilot Results

Results are displayed in Table 2 for each outcome variable of interest over three specifications: (1) no controls, (2) TA fixed effects, and (3) TA fixed effects and student demographic controls. Results are presented for students in the sample who did not dropout of the course. Results presented in Panel A of Table 2 indicate a strong positive treatment effect of 14 percentage points on students' second midterm scores, which followed after the second e-mail of the intervention. Perhaps driven by this treatment effect on the second midterm, students in the treatment group also performed 8 percentage points (or approximately half a letter grade) higher compared to their control group peers on their final course grade.

[Insert Table 2]

Students in the treatment group also scored approximately 15 percentage points higher than students in the control group on their overall homework assignments (Table 2, Panel B). Results in Panel B of Table 2 also indicate that there is some evidence that students in the treatment group spent as much as two hours or more on their homework assignments, as measured by time spent in the homework portal; however, these results are not statistically significant. Additional results on plausible mechanisms (Panel C) suggest that there are small, positive treatment effects on the number of office hour visits and negative effects on the likelihood of dropping out of the course, though, these results are also not statistically significant. Finally, there is some evidence that students in the treatment group are more likely to report that their professor cares about their

performance but less likely to report that they are motivated to do well in the course. Again, these results are not statistically significantly different from zero.

Overall, the pilot results suggest that a light-touch intervention of increased professor feedback (i.e. engagement) can significantly affect students' course performance. Potential mechanisms for this treatment effect may be that students spend more time on assignments and devote more time to course material. Alternatively, students may feel more comfortable seeking help from the professor or TA and therefore understand the material better. A third reason may be that students feel the professor cares about their experiences, causing them to be more motivated and engaged in the course. Additional qualitative feedback provided through students' replies to professor emails indicate that the third explanation may be at play; specifically, that the professor's engagement and concern for their well-being, was an important feature of the course for students in the treatment group. Student email replies expressing their gratitude towards this individual attention are instructive, examples of this feedback include:

- *I'd...like to thank you for offering your help in such a kind manner, I've rarely seen teachers at this school respond to missed assignments the way you have. I'll be sure to complete future assignments in a timely manner, the first practice homework was indeed pretty helpful.*
- *Thanks for talking to me about my homework and test scores. Even though you have a couple hundred students, I really appreciate the effort you put into making it personal for your students. I would have gone to office hours the first time you emailed me, but I simply forgot by the end of the week.*
- *This class is fulfilling a GE for me and my other classes do rank higher on my list. While I do enjoy microeconomics and usually aim to only take GE's that interest me,*

I'm just too busy to spend too much time on the homework. I definitely plan to study and practice extensively for the tests, but not make it a priority, which I know is not ideal but a reality for this quarter. But once again, thanks for the email, I really appreciate what you're doing and I do wish more professors do what you are doing.

It is worth noting that each of these comments suggest that students are appreciative primarily of the contact between them and the professor, rather than the information provided itself.

Moreover, these e-mails indicate that students are not accustomed to receiving individualized attention from their professors in large, introductory courses and that they are appreciative of such gestures. Given the promising results from the pilot, we scaled-up the intervention at the same large, broad-access, four-year institution where we conducted the initial focus groups, and where completion rates—particularly for Latinx and Black student—are low.

4. Scale-up

The study was implemented in two separate waves during the spring of 2016 and fall of 2017 at a large representative, broad-access four-year institution where the original focus groups took place.⁶ We randomly chose 30 large undergraduate courses (serving over 120 students) in each respective term and identified the instructor of record. Collaborating with the Campus Center for College and Career Readiness we recruited these professors by sending personalized letters signed by both the Provost and Dean of Undergraduate Studies.⁷ In total, 34 faculty members across 27 different course subjects participated in the study, with nearly 4,000 total students in the treatment and control groups. All participating faculty were given templates of E-

⁶ This study was registered at the American Economic Association's registry for randomized controlled trials: <https://www.socialscienceregistry.org/trials/5875>

⁷ All recruitment materials available from the authors upon request.

mails that they were encouraged to personalize to their own courses (available in Appendix B). Given the autonomy faculty have in the college classroom, our goal in the scale-up was to allow faculty to individualize to their own teaching style (i.e. what each respectively believed was the feedback students needed about how to be successful in their course). However, all emails had to meet three basic criteria: 1) they had to be personalized to the student; 2) they had to acknowledge a student's performance in the course thus far; and 3) they had to provide feedback about what students could do to improve grade performance and/or seek additional help. Participating professors received a \$500 payment to be a part of the experiment, with an additional \$100 gift card for completing a survey at the end of the semester.⁸

4.1 Scale-up Design and Sample

There are several key differences between the scale-up intervention and pilot that are worth mentioning. First, because of both logistical concerns and the overall low graduation rate at the institution we study, rather than condition on the first assignment as was done in the pilot, we chose to randomly select students from the entire course. In the spring of 2016 the treatment group comprised of a randomly selected one-half of the students in 14 large undergraduate classes. To assess whether spillovers may be biasing our estimates, in the Fall of 2017, we drew our treatment group in two different ways. In eight large classes (>120 students) we randomly selected one-third of students into treatment (in contrast to the half randomized in the first wave). Additionally, we randomly selected the entire class to receive treatment in the ten cases where the professor taught two sections of the identical course. Second, rather than providing two targeted emails as was done in the pilot (at a 10-week quarter system course), we chose to have three targeted emails in the scale-up (at 16-week semester system courses). The timing of these

⁸ Post implementation surveys for participating faculty were only included in the first wave (spring 2016).

emails differed slightly in the two waves. During spring of 2016, the first email entailed an initial “welcome to my class” message containing strategies to succeed in the course. The second and third emails were targeted performance feedbacks at the midway point in the course and just before the final exam. In the fall of 2017, similar to the pilot, we asked professors to give students in the treatment targeted feedback based on the first “meaningful” assignment as well as midway through the course and just before the final exam.

Since professors volunteered to participate in the study, it is important to know whether there are differences in the types of professors who chose to participate in the study versus those who chose not to participate. Though, this level of selection will not bias the internal validity of our estimated effects (i.e., our estimates are unbiased for the sample of professors in the study), professor selection may bias the external validity (i.e., the effects could differ for the average professor at the university). Comparing professor characteristics in Table A2 shows there are no significant differences in Rate My Professor ratings of participating versus non-participating professors. However, participating professors are significantly *more* likely to be female, Black or Latino. It is unclear, however, how this selection may affect external validity.

4.2 Scale-up Data and Methods

Our sample consists of a broad set of academic subjects from Art to Engineering (Table A3 in the Appendix lists the set of course subjects taught by professors in the study). The sample for the scale-up intervention is very diverse across race/ethnicity, with Latino students who represent the largest group on campus at 25% of the sample. Fifty-percent of students are female and 37% are first-time freshman. The average student is midway through their sophomore year, having completed just over 45 units, with an average combined SAT score of 964. To ensure

balance across randomization we regress treatment status on the set of observable characteristics. Results indicate that student characteristics are largely uncorrelated with treatment status, with only 2 of 30 coefficients significant at the 10-percent level (Table A4 in the Appendix).

To assess plausible mechanisms of treatment effects, we collected data on student perceptions of the professor and course, as well as academic achievement measures. Student perceptions were obtained via survey at the end of the course. The survey was administered via email, where students were incentivized to complete the survey by being entered into a lottery to win an iPad or Amazon gift card. In addition, at the end of the Fall 2017 semester the survey was administered in-person by our research assistants in select classes. Table 3 includes the list of questions on the survey as well as summary statistics for responses measured on a 1-5 Likert-scale. The overall survey response rate was 25%, which is quite similar to other college surveys (Carrell & Sacerdote, 2017). Table A5 in the Appendix provides results from models regressing the probability of response on student background characteristics. Unsurprisingly, response is positively correlated with college GPA and SAT scores. Importantly, response is uncorrelated with treatment status.

Our primary measure of academic achievement is course grade, which was obtained from the university registrar. Given the coarseness of this measure, we also collected gradebooks from willing professors, which allows us to examine the effects of the intervention on the total percentage of points earned in the course. Summary statistics for the scale-up in Table 3 show that the average course grade is 2.53, with 82% of students earning a passing grade. Among classes where we were able to obtain gradebooks, the average percentage of points earned in the course (after the first feedback email) is 72%.

[Insert Table 3]

Similar to the pilot, random assignment of study subjects to treatment status again allows for a simple OLS regression analysis to calculate the average treatment:

$$Y_{ict} = \alpha + \beta * treat_{ict} + \gamma * X_i + \lambda_{ct} + \varepsilon_{ict}$$

Where Y represents our respective outcomes of interest for student i in course c , during semester t , " $treat$ " is a dummy variable for treatment versus control status and X is a vector of individual student characteristics. β represents the average causal effect of the (intervention) on student outcomes. We calculate treatment effects for each outcome variable using two specifications. The first specification includes a dummy indicator for treatment status and a professor by semester fixed effect (λ_{ct}). The professor by semester fixed effect is used to account for unobserved differences across professor and semester. The second specification includes both professor by semester fixed effects and student-level controls. Individual control variables include: Female, Cumulative Units Earned, College GPA, remediation status, First Time Freshman, SAT Combined Score, Black, Latino, Asian, White, Missing_SAT. For statistical inference, and to address for multiple hypothesis testing, we follow Athey and Imbens (2017) and List, Shaikh, and Xu (2019) and use a bootstrap-based procedure for testing these null hypotheses simultaneously using experimental data in which random sampling is used to assign treatment status. Hence, instead of reporting traditional clustered standard errors, square brackets contain empirical p-values from randomization-based inference using a counterfactual of randomly assigning treatment status within classrooms 500 times.⁹

⁹ Athey and Imbens (2017) recommend the use of randomization-based inference in lieu of sampling-based inference for experiments. Additionally, as discussed by List, Shaikh, and Xu (2019), "by incorporating information about dependence ignored in classical multiple testing procedures, such as the Bonferroni (1935) and Holm (1979) corrections randomization-based inference has much greater ability to detect truly false null hypotheses."

4.3 Scale-up Results

4.3.1 Average Treatment Effects

Findings of average treatment effects for the scale-up are presented in Table 4. Results are displayed for each outcome variable of interest over two specifications: (1) Professor by semester fixed effects, and (2) Professor by semester fixed effects with student controls. Importantly, across all outcomes, our results are largely unchanged when including controls, providing further evidence that the randomization is valid and that treatment is uncorrelated with survey response. Results in Table 4, columns 1-4 show the treatment had no discernable effect on student performance in the course. Across all measures of achievement, including course grade, percentage of points earned in the course, passing the course, or earning an A or B grade, we find null effects. These null findings are quite precise—for instance, we can rule out effects sizes larger than 0.12 grade points for final course grade and three percentage points for the percentage of points earned in the course.

Table 4, columns 5-9 present how students experienced this intervention, specifically results from the survey, which examines student perceptions of the professor and course. We provide evidence of a strong positive treatment effect on student perceptions of the instructor and course. Students in the treatment group respond more positively on questions asking whether the professor was approachable, available, and cared as well as how well the student felt supported and informed. The largest treatment effects are for the questions that asked students how much they believed the *professor cared* about their success and how well the professor kept them *informed* about their progress in the class. The magnitudes of these effects are relatively large with students in the treatment group, on average, responding nearly a quarter of a standard deviation higher than students in the control condition. Of note is the fact that the question

asking about the *usefulness* of professor feedback is statistically insignificant and of much smaller magnitude (.104 standard deviation). Finally, to test whether our null effects could be driven by attrition, in column 10, we see that the treatment had no effect on whether students dropped the course.

[Insert Table 4]

4.3.2 *Heterogeneous Treatment Effects*

Our focus groups targeted male students of color, the group with the lowest graduation rates, and the target population for this intervention. Moreover, our pre-experiment hypothesis was that targeted information from the professor would most likely help students in their first year of college. Thus, in Tables 5a and 5b we present heterogeneous treatment effects by race/ethnicity, gender, and measures of experience and academic preparedness (results from models that include additional controls are in the Appendix). To adjust inference for multiple hypothesis testing, as recommended by List, Shaikh, and Xu (2019), in square brackets we again report empirical p-values from randomization-based inference using a counterfactual of randomly assigning treatment status within classrooms 500 times. Results in Table 5a show treatment effects for African American and Latinx students. Here we first note positive average treatment effects on course grades, which is driven by first year students, particularly those who enter less prepared (as measured by their high school GPA). For these groups, (freshman and those least academically prepared) the effect sizes are quite large with students in the treatment significantly outperforming control students by upwards of a third of a letter grade (0.31 and 0.43, respectively). Although, the treatment had no overall effect on dropping the course, we do note that treatment students in the most vulnerable group (freshman men of color) were 6-

percentage points less likely to drop the course (p-value = 0.040). Treatment effects on students' perceptions of their instructor were universally positive and by and large statistically significant for Black and Latinx students, irrespective of gender and first year status.

Table 5b presents results for non-Black and Latinx students. We note the positive and statistically significant effect of the treatment for first year female students' course grade, and a correspondingly negative and statistically significant effect on first year male students' course grade. Similar to the underrepresented minority population, treatment effects on students' perceptions of their instructor were largely positive and statistically significant for all non-black and Latinx students, irrespective of gender and first year status.

Together the results suggest that the treatment had a positive effect on students' perception of instructor support. However, this only translated to higher course grades for some students, in particular students of color and those early in their college career, suggesting that students may interpret targeted emails from the professor differently depending on their background and when it comes in their college career.

[Insert Table 5a and 5b]

4.3.3 Treatment Effects on Other (Non-Treated) Courses

An important question, particularly for our target population (first year students from underrepresented minority groups) where we find positive treatment effects, is whether the intervention in one course affects academic performance in treated students' other (non-treated) courses. On the one hand, the positive treatment effects we observe could be driven by a reallocation of student effort from non-treated classes to the treated class, resulting in no overall gain in average academic achievement. On the other hand, faculty engagement from one

professor could result in an overall increase in treated students' self-efficacy or sense of belonging, thereby improving performance in their non-treated courses. To test these possibilities, in Table 6 we present results for student grades in courses other than those in the experimental study. For comparison purposes, we also show in odd numbered columns results for treated courses. Overall, the pattern of results shows compelling evidence of a positive spillover from treated courses to non-treated courses. For students in the target populations (Black/Latinx first year students, Black/Latinx first year males, and Black/Latinx with low high school GPAs), the magnitude of the treatment effect on grades in non-treated courses is roughly one-half the size of the effects in treated courses and statistically significant at the 0.05-level. For example, while the treatment effect for Black/Latinx freshman is 0.31 grade points for treated courses, the effect in other (non-treated) courses is a 0.15 grade points ($p=0.024$).

[Insert Table 6]

5. Discussion

Overall, these scale-up results suggest that a light-touch intervention that demonstrates professor engagement significantly improves all students' perceptions of the professor and course, and the course performance of some students, namely underrepresented minority students in their first year of college. Differences in these course performance effects could be driven by a number of factors, including: (1) fidelity of implementation of the treatment, (2) differential response to the feedback, or (3) contextual differences with the pilot setting. We explore these plausible explanations below.

5.1. Fidelity of Implementation and Course/Instructor Differences

One potential concern with the scale-up is the fact that professors may have not followed through with the experiment, or were just motivated by the monetary incentive, rather than the potential for improving student outcomes. To address potential differences across courses we do several things: First, we assigned a research assistant to all participating faculty to assist with E-mail drafts and ensure no contamination of the control group; second, we examine the timeliness and quality of the emails sent to students by all instructors. Although we find differences in the quality of the text of the emails provided to students (i.e. specificity and/or encouragement), and timeliness of the feedback (i.e. in conjunction with key course assignments or exams), all participating instructors met the three criteria required by the emails—personalized feedback responsive to student performance. Thus, we can confidently rule out fidelity as a potential mechanism. In addition to fidelity of implementation, we also explore whether treatment effects may have differed by other course and instructor types. Specifically, we look at differences by course discipline, faculty characteristics (gender and race/ethnicity), and appointment status (lecturer, tenured faculty, tenured-track faculty). We find no systematic pattern of treatment effects by course or instructor type.¹⁰

5.2 Student response to feedback

All participating professors collected any student replies to their E-mails, which we analyze qualitatively. Similar to the pilot, many students from the treatment group wrote E-mails expressing their appreciation and gratitude towards this individual attention. Examples of this feedback include:

- *Thank you for your email, I will keep that in mind for the future. I appreciate all the help.*

¹⁰ These results are available upon request.

- *Hello Professor, It means a great deal to receive feedback and am appreciative of your time and help. I love what im learning and will reach out if when I need guidance.*
- *Hi Professor, Thank you for all of this information. It's very useful and I'm looking forward to learning a lot from your class. I was struggling in the beginning because I've never taken a one part lecture and one part discussion based class, but I think I'm starting to get the hang of it. If I have any questions, I'll be sure to stop by your office hours. Thanks once again!"*

Several other themes in the qualitative data emerged. Students were apologetic, often expressing regret for their actions, and students also communicated a host of explanations that included both academic and personal challenges:

- *Hello professor. I attend every class, go to the review sessions, and have turned in the extra credit so I am defiantly trying to do well but I am still struggling. I will come to office hours and try to meet up with our TA as well. Let me know if there is anything else I can do. Thank you*
- *I apologize for missing your class wednesday afternoon, i was stuck in [Name] hall trying to pay my monthly installment for tuition. I will definitely be at mondays lecture.*
- *Thank you for email! I hope to do well on the next 2 exams. I also apologize for my poor performance on the first exam, there was a personal problem I had to deal with the day before and it affected my studying & performance on the exam. Thank you for reaching out, I really appreciate it.*
- *Thank you so much for your concern I have been struggling a bit in the class with chapter 3, I have been trying to keep up with school along with working but I am not making any excuses. I was also not too pleased with my performance with my grade on the first midterm because I had did well on majority of the homework's and attend class daily. I do plan on seeking help and getting a tutor in Brighton Hall that will work with my schedule and spending a little more time focusing on homework's. I appreciate your encouragement in making sure I stay on track in the class and I will be sure to do better the remaining of the semester*
- *I truly appreciate the grade check-in, the bad grade was due to my lack of organization and failure to take it before the deadline. Once again i truly appreciate the check in and i will make sure to be more aware of the upcoming due dates.*

Students—at least those that replied—communicated an effort to try to respond to the suggested actions on the part of the instructor. As an example, in response to one instructor's final email to students in the treatment group, as follows:

I hope you had a great Thanksgiving break! We are approaching the end of the semester.

I want to let you know that I have been looking over your grades. Earlier today I sent an email announcement to the class, where I mentioned that your current grades on the class have been posted on UnivCT under the heading "Grade_Nov27" and explained how this grade was calculated. Your current grade in the class is XX%. I am a bit concerned with your current grade and want to encourage you to study hard for this exam and the final.

I also encourage you to continue coming to class regularly, completing the few remaining assignments on time and seeking help when concepts are unclear.

We have an exam coming up this Friday. To remind you, my office hours are as follows... Please feel free to contact me if you have any questions.

Students' replies include:

- *Thank you for your concern and informing me on my current grade. I intend on focusing my time to study hard for the upcoming exam as well as the final. If I'm unsure about a topic or have any questions I will be sure to come to your office.*
- *Thank you professor, I am trying my best to prepare for this exam, I plan on earning at least a B on this one! Thank you for the encouragement, it helps a lot!*
- *Thank you, I hope that you did too! I'm going to come see you during office hours tomorrow because I know that although I have an 81.25% nearly half of my grade is undetermined yet. I really need to get a passing score on this Exam so I will see you at 2PM tomorrow! Thanks for the update on my grade. I appreciate it.*
- *Thank you for the email and thank you for caring about my grade I really appreciate it and I can say that your efforts have helped me. I will be finishing off the semester the best that I can by performing my best on the exam 3 and final exam I hope to come to one of your office hours tomorrow.*

5.3 Faculty Response to the Intervention

How did faculty respond to this intervention? We surveyed all faculty participants in the scale-up. We asked faculty how they interpreted the nature of the student responses to their emails. Faculty reported students' replies were largely positive, thanking them and suggesting they would try harder. A few also described students' concern over receiving an email, either in a curious way, with some potentially worried. Faculty were both surprised by the gratitude expressed by students: *"It was surprising how thankful they were for such a simple email"* (as reported by one instructor); while other faculty were more skeptical: *"Responses generally came from what I would consider already conscientious students. They weren't defined by grade, but by active involvement. If they were really engaged in the classroom, they were more interested in the emails. Students that didn't care probably ignored them"* (as reported by another instructor). Faculty were asked how long the emails took them to complete; a conservative estimate is approximately 1 minute per email. Faculty believed that increasing interaction with students in their class could improve student outcomes, and while some were enthusiastic by this specific

effort: *“With a class of this size, I think these emails really did serve a useful purpose of establishing some level of one-on-one interaction between myself and the students.”* Others were more skeptical of the effort, *“I think it’s important, but some of them really don’t care. I can’t force them to come to my office.”* These qualitative findings suggest that faculty are by and large receptive to various tools that may increase feedback to students and greater interaction with their students. However, it also suggests that these efforts may be mediated by faculty attitudes and perceptions of the utility of various pedagogical efforts in their classrooms.

5.4 Pilot Replication

Given the heterogeneity in our treatment findings on course performance, we replicated the pilot in the same large, introductory-level course at a selective comprehensive research university in the spring of 2019. We do so to examine the robustness of the finding that motivated the scale up. There was only one major difference regarding the research design from the initial pilot. The professor was unable to draw the sample based on performance on the first homework because only 13 of 720 students failed the first homework assignment. Rather, the professor based the sample from the students who failed (<60%) the first (of four) exams, which occurs within the first three weeks of the course. Two feedback emails were sent, one immediately after exam 1 and the second after exam 3.

Table 7 presents results from the pilot replication. Similar to the initial pilot, we find a significant positive effect from instructor feedback on exam performance. Across all exams, there is a 5 percentage point boost in student performance, though the estimates are only significant at the 10-percent level ($p=0.066$). However, of note is that this effect did not lead to a statistically significant effect on the course grade, which was determined after dropping each

student's lowest exam score and curving. Though, students in the treatment group were more likely to pass the course (0.021 ppts) and earn a grade of B or higher (0.129 ppts), neither of these effects are significant at conventional levels. Finally, similar to both the original pilot and scale-up, students by and large provided positive responses to receiving individualized feedback from the instructor.

[Insert Table 7]

6. Conclusion

College completion and success remains highly uneven by institutional type (i.e. selectivity), and, by students' background characteristics. Despite the robust evidence from K-12 on the role of teachers, to date, we have a much more limited sense of the role of college faculty/instructors in student success. To our knowledge, this study represents the first experiment aimed at altering faculty behaviors in the college classroom. The experiment follows a theoretically grounded and carefully piloted and tested treatment that represents an effort, not to revolutionize the college classroom, but rather to modestly increase faculty engagement through individualized feedback to students in large lecture class.

The results provide experimental evidence that professor feedback to students can have a positive significant effect on all students' perception of support in their college classes, and on course performance for some students and in some settings. Moreover, a compelling set of qualitative evidence suggests that students recognize and appreciate this type of feedback from their instructors.

By conveying beliefs in students' abilities to succeed in a course and in college more generally, college instructors have an important way to directly and indirectly contribute to

college success: directly through the intended transfer of content knowledge and/or skills and indirectly through boosting students' sense of self-efficacy. Students' beliefs about college and how they process early difficulties can influence their postsecondary trajectory. Thus, we hypothesize that feedback and encouragement earlier in an academic transition, particularly from a faculty member, could trigger a host of positive effects (e.g. improved self-efficacy), or avert a downward cycle of self-doubt that may lead to premature departure from college.

Faculty are often an untapped source of (dis)encouragement and/or support for college students aspiring to obtain a degree. Despite considerable conjecture about the role of faculty, we have very limited evidence about their potential influence and virtually no evidence about how they might influence student outcomes. This study affirms that faculty can play a critical role in improving student success through a modest set of activities to reach out to their students. More specifically, having direct feedback from faculty that is both individualized in knowledge of the student's progress in the course and encouraging about their potential success could be a powerful motivator. Future work can and should offer additional experimentation with pedagogical approaches to feedback and support in the college classroom, and that is cognizant of how such efforts may be received differently by different types of students (e.g., demographic background, preparation levels, etc.), by different messengers (i.e. instructor differences), and in different contexts (e.g., institutions, disciplines, course format, etc.).

REFERENCES

- Avery, C. & Kane, T. (2004). Student Perceptions of College Opportunities. The Boston COACH Program. In Caroline Hoxby ed, *College Choices: The Economics of Where to Go, When to Go, and How to Pay For It*, Chicago: University of Chicago Press, pages 355-394.
- Agrist, J. Oreopoulos, P. & Williams, T. (2014). When Opportunity Knocks, Who Answers? New Evidence on College Achievement Awards. Forthcoming, *Journal of Human Resources*.

- Anderson, D. M., Broton, K. M., Goldrick-Rab, S., & Kelchen, R. (2018). Experimental Evidence on the Impacts of Need-Based Financial Aid: Longitudinal Assessment of the Wisconsin Scholars Grant.
- Angrist, J., Lang, D., & Oreopoulos, P. (2009). Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, 1(1), 136-63.
- Athey, S. & Imbens, G. (2017). The Econometrics of Randomized Experiments, *Handbook of Economic Field Experiments*, Edited by Abhijit Vinayak Banerjee, Esther Duflo, Volume 1, 2017, Pages 73-140.
- Bailey, M. J. & Dynarski, S. (2011). Inequality in Postsecondary Education. In Duncan, G. J. and Murnane, R. J. (eds.), *Wither Opportunity*. Russell Sage. Pp. 117-132.
- Balzer, W. K., Doherty, M. E., & O'Connor, R., Jr. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin*, 106(3), 410-433.
- Bandura, A. (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. *Educational Psychologist*, 28(2), 117-148.
- Bandura, A., & Schunk, D. H. (1981). Cultivating Competence, Self-Efficacy, and Intrinsic Interest Through Proximal Self-Motivation. *Journal of Personality and Social Psychology*, 41(3), 586-598.
- Barr, A. and Castleman, B. (2017). The Bottom Line on College Counseling. Working Paper.
- Baum, S., Ma, J., & Payea, K. (2013) Education pays: The benefits of higher education for individuals and society. Retrieved from The College Board website: <http://trends.collegeboard.org/sites/default/files/education-pays-2013-full-report.pdf>
- Bettinger, E. P., & Baker, R. B. (2014). The effects of student coaching: An evaluation of a randomized experiment in student advising. *Educational Evaluation and Policy Analysis*, 36(1), 3-19.
- Bettinger, E. P., Long, B. T., Oreopoulos, P., & Sanbonmatsu, L. (2012) The role of application assistance and information in college decisions: From the H&R Block FAFSA experiment. *The Quarterly Journal of Economics* 127(3), 1205-1242.
- Beleche, T., Fairris, D., & Marks, M. (2012). Do course evaluations truly reflect student learning? Evidence from an objectively graded post-test. *Economics of Education Review*, 31(5): 709-719.
- Bettinger and Long, (2010). Does Cheaper Mean Better? The Impact of Using Adjunct Instructors on Student Outcomes. *Review of Economics and Statistics*, 92(3): 598-613.
- Braga, M., Paccagnella, M. & Pellizzari, M., (2014). Evaluating Students' Evaluations of Professors. *Economic of Education Review*, 41, 71-88.

- Brownback, A. & Sadoff, S. (2019). Improving College Instruction Through Incentives. *Journal of Political Economy*.
- Carlson, D. E., Elwert, F., Hillman, N., Schmidt, A., & Wolfe, B. L. (2019). The effects of financial aid grant offers on postsecondary educational outcomes: New experimental evidence from the fund for Wisconsin scholars. Technical report, National Bureau of Economic Research.
- Carrell, S. E., Page, M. E., & West, J. E. (2010). Sex and Science: How Professor Gender Perpetuates the Gender Gap. *The Quarterly Journal of Economics*, 125(3): 1101-1144.
- Carrell, S., & Sacerdote, B. (2017). "Why Do College Going Interventions Work?" *American Economic Journal: Applied Economics*, 9(3): 124-51, July 2017.
- Carrell, S.E. & West, J. E., (2010). Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors," *Journal of Political Economy*, Volume 118, No. 3, 409-432, June 2010.
- Castleman, B.L., Page, L.C., & Schooley, K. (2014). The forgotten summer: Mitigating summer attrition among college-intending low-income high school graduates. *The Journal of Policy Analysis and Management*.
- Dobronyi, C. R., Oreopoulos, P., & Petronijevic, U. (2019). Goal Setting, Academic Reminders, and College Success: A Large-Scale Field Experiment, *Journal of Research on Educational Effectiveness*, 12(1):38-66.
- Ehrenberg, R. G. & Zhang, L. (2005). Do tenured and tenure-track faculty matter? *Journal of Human Resources*, 40(3):647–659.
- Fairlie, Hoffman, & Oreopoulos, P. (2013). A community college instructor like me: Race and ethnicity interactions in the classroom. *American Economic Review*.
- Figlio, D. N., Schapiro, M. O., & Soter, K. B. (2015). Are tenure track professors better teachers? *Review of Economics and Statistics*, 97(4):715–724.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Hoffmann, F. & Oreopoulos, P. (2009). A Professor Like Me: The Influence of Instructor Gender on College Achievement, *Journal of Human Resources*, 44(2).
- Holzer, H. & Baum, S. (2017). Making College Work: Pathways to Success for Disadvantaged Students. The Brookings Institution, Washington, D.C.
- Hoxby, C., & Turner, S. (2013). Expanding college opportunities. *Education Next*, 13(4). Retrieved from <http://educationnext.org/expanding-college-opportunities/>

- Laibson D. (1997). Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics*, 112(2):443-477.
- Leonhardt, D. (2014, May 27). Is College Worth It? Clearly, New Data Say. *The New York Times*. Retrieved from <http://www.nytimes.com>
- Karabenick, S. A., & Knapp, J. R. (1991). Relationship of academic help seeking to the use of learning strategies and other instrumental achievement behavior in college students. *Journal of Educational Psychology*, 83, 221–230.
- Kluger, A. N., & DeNisi, A. (1998). Feedback interventions: Towards the understanding of a double-edge sword. *Current Directions in Psychological Science*, 7(3) 67–72.
- List, J.A., Shaikh, A.M. & Xu, Y. Multiple hypothesis testing in experimental economics. *Exp Econ* 22, 773–793 (2019). <https://doi.org/10.1007/s10683-018-09597-5>
- Oreopoulos, P. & Petronijevic, U. (2019). The Remarkable Unresponsiveness of College Students to Nudging and What We Can Learn from It. National Bureau of Economic Research, Working Paper 26059.
- Pew Research Center. (2014, February). The Rising Cost of *Not* Going to College. Retrieved from <http://www.pewsocialtrends.org/2014/02/11/the-rising-cost-of-not-going-to-college/>
- Phillips, M. & Reber, S. (2019). Does Virtual Advising Increase College Enrollment? Evidence from a Random Assignment College Access Field Experiment. National Bureau of Economic Research, Working Paper 26509.
- Price, J. (2010). The effect of instructor race and gender on student persistence in STEM fields. *Economics of Education Review*, Volume 29, Issue 6, December 2010, Pages 901-910.
- Ran, X. & Xu, D. (2016). How and why do adjunct instructors affect students' academic outcomes? Evidence from two-year and four-year colleges.
- Snyder, T.D., & Dillow, S.A. (2013). Digest of Education Statistics 2012 (NCES 2014-015). National Center for Education Statistics, Institute of Education Sciences. Washington, D.C.: U.S. Department of Education.
- Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology*. New York: Academic Press.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New London, CN: Yale University Press.
- Turner, S. (2018). The Policy Imperative: Policy Tools Should Create Incentives for College Completion. American Enterprise Institute.

- Umbach, P. D., & Wawrzynski, M. R. (2005). Faculty do matter: The role of college faculty in student learning and engagement. *Research in Higher Education, 46*(2), 153-184.
- Walton, G. & Cohen, G. (2011). A Brief Social-Belonging Intervention Improves Academic and Health Outcomes of Minority Students, *Science, 331*(6023): 1447-1451.
- Yeager, D. S. & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*, 267-301.
- Yeager, D. S., Walton, G.M., Brady, S.T., Akcinar, E.N., Paunesku, D., Keane, L., Kamentz, D., Ritter, G., Duckworth, A. L., Urstein, R., Gomez, E. M., Markus, H. R., Cohen, G.L., Dweck, C.S. (2016). Proceedings of the National Academy of Sciences, June 2016, 113 (24): E3341-E3348; DOI: 10.1073/pnas.1524360113.

Tables

Table 1. Descriptive Statistics

| Variable | No. Obs. | Mean | SD | Min. | Max. |
|--------------------------------------|----------|---------|------|------|-------|
| Midterm 1 (%) | 53 | 0.74 | 0.17 | 0.33 | 1 |
| Midterm 2 (%) | 53 | 0.68 | 0.20 | 0 | 1 |
| Final Exam (%) | 53 | 0.64 | 0.15 | 0 | 0.95 |
| All Exams (%) | 53 | 0.68 | 0.14 | 0 | 0.93 |
| Total Course Score (%) | 53 | 0.72 | 0.13 | 0.21 | 0.94 |
| Course Grade (0-4) | 53 | 2.40 | 0.89 | 0 | 4 |
| Homework Score (%) | 53 | 0.93 | 0.20 | 0 | 1 |
| Homework Points Earned (%) | 53 | 0.55 | 0.21 | 0 | 0.88 |
| Homework Total Time Spent (hours) | 53 | 7.05 | 4.16 | 0 | 14.82 |
| Homework Median Time Spent (hours) | 53 | 0.85 | 0.59 | 0 | 2.38 |
| Professor Cares about my Performance | 51 | 2.39 | 0.87 | 0 | 3 |
| Motivated to Do Well in Course | 52 | 3.02 | 1.02 | 0 | 4 |
| Total Office Hour Visits (number) | 53 | 2.32 | 2.29 | 0 | 9 |
| Dropped Out of Course (%) | 69 | 0.23 | 0.43 | 0 | 1 |
| Male | 53 | 0.68 | 0.47 | 0 | 1 |
| First Generation College Student | 53 | 0.26 | 0.45 | 0 | 1 |
| HS GPA | 53 | 3.77 | 0.37 | 2.87 | 4.24 |
| Under-represented Minority | 53 | 0.23 | 0.42 | 0 | 1 |
| CA Resident | 53 | 0.89 | 0.32 | 0 | 1 |
| Entering Cohort | 53 | 2012.43 | 0.69 | 2011 | 2013 |

Table 2: Pilot Results

| Panel A. Test Score Outcomes on Exams | | | | | | |
|--|------------------|--------------------|------------------|--------------------|--------------------|--------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Outcome | Midterm 1 (pct) | Midterm 2 (pct) | Final Exam (pct) | All Exams (pct) | Total Score (pct) | Course Grade (0-4) |
| No Controls | 0.065 (0.048) | 0.121** (0.054) | 0.022 (0.042) | 0.063 (0.039) | 0.064* (0.034) | 0.431* (0.239) |
| TA Fixed Effects | 0.073 (0.052) | 0.150** (0.057) | 0.042 (0.043) | 0.082** (0.041) | 0.078** (0.036) | 0.521** (0.249) |
| Individual Controls and TA Fixed Effects | 0.057 (0.053) | 0.136** (0.060) | 0.049 (0.042) | 0.076* (0.041) | 0.076** (0.037) | 0.501* (0.254) |
| Observations | 53 | 53 | 53 | 53 | 53 | 53 |

| Panel B. Homework Scores and Time Spent | | | | |
|--|----------------------|------------------------------|-----------------------------------|------------------------------------|
| | 1 | 2 | 3 | 4 |
| Outcome | Homework Score (pct) | Homework Points Earned (pct) | Homework Total Time Spent (hours) | Homework Median Time Spent (hours) |
| No Controls | 0.103* (0.057) | 0.067 (0.055) | 1.804 (1.131) | 0.271* (0.159) |
| TA Fixed Effects | 0.119* (0.062) | 0.052 (0.060) | 1.794 (1.242) | 0.257 (0.175) |
| Individual Controls and TA Fixed Effects | 0.152** (0.062) | 0.075 (0.061) | 1.969 (1.333) | 0.311 (0.186) |
| Observations | 53 | 53 | 53 | 53 |

| Panel C. Mechanisms | | | | | |
|--|---------------------------------------|--|----------------------------------|--------------------------------|-----------------------|
| | 1 | 2 | 3 | 5 | 6 |
| Outcome | Professor Office Hour Visits (number) | "Professor Cares About My Performance" | "Motivated to Do Well in Course" | TA Office Hour Visits (number) | Dropped Out of Course |
| No Controls | 0.131 (0.103) | 0.54 (0.332) | -0.328 (0.310) | 0.967 (0.128) | -0.123 (0.102) |
| TA Fixed Effects | 0.101 (0.105) | 0.540 (0.332) | -0.328 (0.310) | 0.802 (0.672) | NA |
| Individual Controls and TA Fixed Effects | 0.093 (0.104) | 0.535 (0.387) | -0.237 (0.367) | NA | NA |
| Observations | 53 | 51 | 52 | 53 | 69 |

Table 2 Notes: * Significant at the 0.10 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level. Each cell represents the results from regressing the outcome listed on a treatment dummy variable. Specifications 2 & 3 in panel C are estimated using an ordered Probit model. All other specifications are estimated using OLS. Individual control variables include whether the student is male, first generation college, under-represented minority, CA resident, entering cohort, and high school GPA.

Table 3. Descriptive Statistics

| Variable | No. Obs. | Mean | SD | Min. | Max. |
|--|----------|------|------|------|------|
| Treatment Group | 3,930 | 0.46 | 0.50 | 0 | 1 |
| Female | 3,922 | 0.54 | 0.50 | 0 | 1 |
| White | 3,931 | 0.22 | 0.42 | 0 | 1 |
| Asian | 3,931 | 0.25 | 0.44 | 0 | 1 |
| Black | 3,931 | 0.10 | 0.29 | 0 | 1 |
| Latino | 3,931 | 0.35 | 0.48 | 0 | 1 |
| High School GPA | 3,384 | 3.29 | 0.42 | 2 | 4.44 |
| Freshman | 3,931 | 0.40 | 0.49 | 0 | 1 |
| Sophomore | 3,931 | 0.28 | 0.45 | 0 | 1 |
| Junior | 3,931 | 0.19 | 0.39 | 0 | 1 |
| College GPA | 3,642 | 2.90 | 0.60 | 0 | 4.00 |
| Total Units | 3,922 | 0.45 | 0.34 | 0 | 2.16 |
| Course Grade | 3,734 | 2.53 | 1.19 | 0 | 4.0 |
| Passed Course | 3,734 | 0.82 | 0.38 | 0 | 1 |
| Earned A or B | 3,734 | 0.57 | 0.50 | 0 | 1 |
| Dropped Course | 3,930 | 0.04 | 0.19 | 0 | 1 |
| % Points Earned after First Feedback | 1,624 | 0.72 | 0.22 | 0.00 | 1.01 |
| How approachable was the instructor in class? | 1,014 | 4.50 | 1.21 | 1 | 6 |
| How available was the instructor outside of class? | 1,011 | 4.46 | 1.11 | 1 | 6 |
| How useful was the instructor's feedback in helping you learn? | 1,015 | 4.57 | 1.22 | 1 | 6 |
| How much do you believe the instructor cared about your success in the class? | 1,014 | 4.31 | 1.24 | 1 | 6 |
| How well did the instructor keep you informed about your progress in the class? | 1,014 | 4.26 | 1.31 | 1 | 6 |

Table 4: Scale-up Results

| Specification | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|----------------|--|----------------|----------------|--|---|--|---|--|-------------------|
| Outcome | Grade | % Points Earned after First Feedback | Passed | A or B | How approachable was the instructor in class? | How available was the instructor outside of class? | How useful was the instructor's feedback in helping you learn? | How much do you believe the instructor cared about your success in the class? | How well did the instructor keep you informed about your progress in the class? | Dropped Course |
| Panel A. Class Fixed Effects | | | | | | | | | | |
| Full Sample | 0.01 [.660] | 0.01 [.250] | 0.01 [.456] | 0.01 [.746] | 0.22*** [.004] | 0.17** [.014] | 0.12 [.118] | 0.28*** [.000] | 0.28*** [.000] | -0.01 [.356] |
| Obs | 3,735 | 1,624 | 3,735 | 3,735 | 1,015 | 1,012 | 1,015 | 1,016 | 1,015 | 3,930 |
| Panel B. Class Fixed Effects and Controls | | | | | | | | | | |
| Full Sample | 0.04 [.204] | 0.01 [.240] | 0.01 [.254] | 0.01 [.298] | 0.23*** [.006] | 0.17** [.016] | 0.11 [.160] | 0.28*** [.000] | 0.27*** [.002] | -0.01 [.326] |
| Obs | 3,729 | 1,620 | 3,729 | 3,729 | 1,012 | 1,009 | 1,012 | 1,013 | 1,012 | 3,922 |

Notes: Each column reports results from a separate regression. All specifications include course by phase fixed effects. Controls include Female, Total Units Earned, High School GPA, College GPA, indicators for year of study, Black, Latino, Asian, White. Square brackets contain p-values from randomization-based inference using a counterfactual of randomly assigning treatment status within classrooms 500 times. Note: There are eight students in the sample with missing demographics. The results are unchanged when dropping these students from Panel A.

- * Significant at the 10% level
- ** Significant at the 5% level
- *** Significant at the 1% level

Table 5a. Scale-up Results for Latino/Black Students (no controls)

| Specification | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|---|-------------------|---|------------------|------------------|--|--|--|---|--|-------------------|--------------------------|
| Outcome | Grade | % Points Earned after First Feedback | Passed | A or B | How approachable was the instructor in class? | How available was the instructor outside of class? | How useful was the instructor's feedback in helping you learn? | How much do you believe the instructor cared about your success in the class? | How well did the instructor keep you informed about your progress in the class? | Dropped Course | N (Grades) N (Survey) |
| Treatment Effects for Black/Latino Students | | | | | | | | | | | |
| All Latino/Black | 0.07 [.172] | 0.03* [.080] | 0.02 [.372] | 0.03 [.124] | 0.27* [.074] | 0.22** [.034] | 0.17 [.252] | 0.35*** [.000] | 0.33*** [.002] | -0.01 [.236] | 1,619 403 |
| Female | 0.10 [.160] | 0.00 [.970] | 0.03 [.360] | 0.05 [.146] | 0.33* [.078] | 0.29** [.038] | 0.18 [.436] | 0.42*** [.004] | 0.42*** [.006] | 0.00 [1.000] | 938 271 |
| Male | -0.01 [.956] | 0.05** [.046] | -0.01 [.806] | -0.01 [.972] | 0.16 [.332] | 0.02 [.880] | 0.24 [.198] | 0.29 [.136] | 0.40* [.060] | -0.02 [.094] | 681 132 |
| Freshman | 0.31*** [.002] | 0.04 [.104] | 0.09 [.022] | 0.09** [.032] | 0.44** [.012] | 0.35** [.028] | 0.41** [.018] | 0.52** [.012] | 0.58*** [.010] | -0.01 [.236] | 710 180 |
| Upperclass | -0.07 [.608] | 0.02 [.378] | -0.03 [.356] | -0.01 [.886] | 0.16 [.490] | 0.17 [.240] | 0.06 [.744] | 0.28** [.036] | 0.19 [.184] | -0.01 [.490] | 909 223 |
| Freshman Female | 0.28** [.046] | 0.02 [.526] | 0.08 [.138] | 0.07 [.234] | 0.55*** [.006] | 0.44*** [.008] | 0.42** [.036] | 0.63*** [.010] | 0.59 [.016] | 0.01 [.708] | 435 127 |
| Freshman Male | 0.30* [.058] | 0.08 [.140] | 0.08 [.106] | 0.11* [.052] | 0.41 [.506] | 0.17 [.768] | 0.35 [.532] | 0.40 [.532] | 0.76 [.182] | -0.06** [.040] | 275 53 |
| Freshman w/Low HS GPA | 0.43*** [.010] | 0.07* [.070] | 0.14** [.016] | 0.11* [.074] | 0.64** [.014] | 0.52** [.014] | 0.27 [.296] | 0.55 [.214] | 0.48 [.198] | -0.02 [.242] | 402 94 |
| Freshman w/High HS GPA | 0.11 [.400] | 0.02 [.688] | 0.00 [.894] | 0.04 [.430] | 0.34 [.228] | 0.22 [.418] | 0.63** [.038] | 0.68*** [.008] | 0.64** [.03] | 0.00 [1.000] | 308 86 |

Notes: Each column reports results from a separate regression. All specifications include course by phase fixed effects. Square brackets contain p-values from randomization-based inference using a counterfactual of randomly assigning treatment status within classrooms 500 times. *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

- * Significant at the 10% level
- ** Significant at the 5% level
- *** Significant at the 1% level

Table 5b. Scale-up Results for Non-Latino/Black Students (no controls)

| Specification | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|---|------------------|---|------------------|-------------------|--|--|--|---|--|-------------------|--------------------------|
| Outcome | Grade | % Points Earned after First Feedback | Passed | A or B | How approachable was the instructor in class? | How available was the instructor outside of class? | How useful was the instructor's feedback in helping you learn? | How much do you believe the instructor cared about your success in the class? | How well did the instructor keep you informed about your progress in the class? | Dropped Course | N (Grades) N (Survey) |
| Treatment Effects for non-Black/Latino Students | | | | | | | | | | | |
| All non-Black/Latino | -0.03 [.512] | -0.01 [.662] | 0.01 [.742] | -0.02 [.396] | 0.22** [.016] | 0.16* [.066] | 0.12 [.154] | 0.24** [.03] | 0.25*** [.008] | 0.00 [.902] | 2,110 609 |
| Female | 0.02 [.576] | -0.02 [.186] | 0.01 [.774] | 0.01 [.784] | 0.31** [.018] | 0.18** [.046] | 0.18* [.062] | 0.30*** [.000] | 0.28*** [.006] | 0.01 [.392] | 1,100 374 |
| Male | -0.10 [.156] | 0.01 [.654] | 0.00 [.890] | -0.05 [.110] | 0.09 [.426] | 0.14 [.334] | -0.04 [.776] | 0.21 [.272] | 0.27** [.040] | -0.01 [.410] | 1,010 235 |
| Freshman | 0.03 [.680] | 0.01 [.668] | 0.01 [.548] | 0.00 [.986] | 0.26** [.026] | 0.14 [.214] | 0.24* [.054] | 0.31** [.038] | 0.51*** [.004] | 0.00 [1.000] | 802 223 |
| Upperclass | -0.05 [.338] | -0.01 [.332] | 0.00 [.906] | -0.02 [.368] | 0.18 [.114] | 0.13 [.296] | 0.04 [.710] | 0.18 [.146] | 0.11 [.248] | 0.00 [.874] | 1,308 386 |
| Freshman Female | 0.24** [.048] | 0.02 [.432] | 0.08** [.05] | 0.08* [.06] | 0.40** [.018] | 0.24 [.126] | 0.33** [.048] | 0.33* [.06] | 0.56*** [.006] | 0.02 [.264] | 433 152 |
| Freshman Male | -0.22* [.074] | -0.01 [.778] | -0.07* [.060] | -0.09** [.036] | -0.05 [.836] | -0.26 [.256] | -0.23 [.354] | 0.42 [.062] | 0.33 [.356] | -0.03 [.250] | 369 71 |
| Freshman Low GPA (HS) | 0.00 [.988] | 0.03 [.406] | 0.06 [.216] | -0.09* [.060] | 0.44* [.090] | 0.30 [.176] | 0.33 [.126] | 0.46* [.090] | 0.56* [.056] | 0.00 [1.000] | 366 87 |
| Freshman High GPA (HS) | 0.02 [.788] | -0.01 [.856] | -0.02 [.386] | 0.06 [.124] | 0.01 [.940] | 0.02 [.896] | 0.11 [.540] | 0.15 [.408] | 0.35 [.112] | 0.00 [1.000] | 436 136 |

Notes: Each column reports results from a separate regression. All specifications include course by phase fixed effects. Square brackets contain p-values from randomization-based inference using a counterfactual of randomly assigning treatment status within classrooms 500 times. *Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level.

- * Significant at the 10% level
- ** Significant at the 5% level
- *** Significant at the 1% level

Table 6: Treatment Effects on Other (Non-Treated) Courses

| Specification | 1 | 2 | 3 | 4 |
|------------------------|-------------------------|-------------------|---------------------------|-----------------|
| | Black/Latino Students | | Non-Black/Latino Students | |
| | | Grades in Non- | | Grades in Non- |
| Outcome | Grade in Treated Course | Treated Courses | Grade in Treated Course | Treated Courses |
| All Students | 0.07 [.172] | 0.05 [.222] | -0.03 [.512] | 0.02 [.608] |
| Female | 0.10 [.160] | 0.06 [.276] | 0.02 [.576] | 0.04 [.536] |
| Male | -0.01 [.956] | 0.01 [.900] | -0.10 [.156] | -0.01 [.810] |
| Freshman | 0.31*** [.002] | 0.15** [.024] | 0.03 [.680] | 0.06 [.334] |
| Upperclass | -0.07 [.608] | 0.00 [.982] | -0.52 [.338] | 0.18 [.832] |
| Freshman Female | 0.28** [.046] | 0.08 [.342] | 0.24** [.048] | 0.11 [.266] |
| Freshman Male | 0.30* [.058] | 0.23** [.034] | -0.22* [.074] | 0.05 [.612] |
| Freshman w/Low HS GPA | 0.43*** [.010] | 0.24*** [.010] | 0.00 [.988] | 0.07 [.508] |
| Freshman w/High HS GPA | 0.11 [.400] | 0.03 [.768] | 0.02 [.788] | 0.05 [.526] |

Notes: Each column reports results from a separate regression. All specifications include course by phase fixed effects. Square brackets contain p-values from randomization-based inference using a counterfactual of randomly assigning treatment status within classrooms 500 times. *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

* Significant at the 10% level

** Significant at the 5% level

*** Significant at the 1% level

Table 7: Pilot Replication Results

| Panel A. Test Score Outcomes | | | | | | |
|------------------------------|--------------------|--------------------|--------------------|--------------------|---------------------|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Outcome | Midterm 1 (pct) | Midterm 2 (pct) | Midterm 3 (pct) | Midterm 4 (pct) | Exams 2- 4 (pct) | Course Grade Post Exam 1 (pct) |
| Treatment | 0.008 (0.031) | 0.049* (0.029) | 0.050 (0.031) | 0.049 (0.037) | 0.049* (0.027) | 0.044* (0.026) |
| Observations | 138 | 138 | 138 | 138 | 138 | 138 |
| Control Mean | 0.542 | 0.628 | 0.490 | 0.598 | 0.572 | 0.624 |

| Panel B. Homework and Grade Outcomes | | | | | | |
|--------------------------------------|----------------------------|---------------------------|------------------|------------------|----------------------------|-------------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Outcome | Homework Score (pts) | Homework Time (min) | Letter Grade | Pass Course | Grade of C or higher | Grade of B or higher |
| Treatment | 1.264 (2.941) | 7.975 (49.770) | 0.163 (0.126) | 0.021 (0.042) | 0.082 (0.070) | 0.129 (0.081) |
| Observations | 138 | 114 | 138 | 138 | 138 | 138 |
| Control Mean | 70.143 | 440.690 | 2.097 | 0.890 | 0.643 | 0.271 |

Notes: Each cell represents the results from regressing the outcome listed on a treatment dummy variable. All specifications control for midterm 1 score.

* Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level.

Appendix A

Table A1: Pilot Randomization Checks

| Variable | 1 | 2 |
|--|-------------------|-------------------|
| Male | -0.155 (0.173) | -0.091 (0.173) |
| First Generation College Goer | 0.058 (0.166) | 0.05 (0.169) |
| HS GPA | 0.153 (0.210) | 0.081 (0.214) |
| Under-represented Minority | -0.227 (0.196) | -0.296 (0.196) |
| CA Resident | 0.2 (0.244) | 0.139 (0.243) |
| Entering Cohort==2012 | -0.091 (0.252) | -0.162 (0.249) |
| Entering Cohort==2012 | 0.083 (0.239) | -0.042 (0.240) |
| Observations | 53 | 53 |
| P-value: Joint Significance of all individual covariates | 0.8905 | 0.8449 |
| Includes TA Fixed Effects | No | Yes |

Notes: * Significant at the 0.10 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level. Each specification represents results for a regression where the dependent variable is an indicator for treatment status.

Table A2: Professor Characteristics

| Variable | Summary Stats | Selection Regressions | | | |
|---|----------------|-----------------------|------------------|-------------------|------------------|
| | (mean/std) | 1 | 2 | 3 | 4 |
| Rate My Professor Overall Difficulty Rating | 3.82 (0.49) | -0.12 (0.10) | | | -0.07 (0.10) |
| Rate My Professor Overall Rating | 2.85 (0.60) | -0.07 (0.09) | | | -0.06 (0.09) |
| Full Professor | 0.38 (0.49) | | 0.24** (0.11) | | 0.17 (0.11) |
| Associate Professor | 0.11 (0.32) | | 0.14 (0.17) | | 0.12 (0.17) |
| Assistant Professor | 0.08 (0.27) | | 0.17 (0.20) | | 0.33 (0.21) |
| Female | 0.35 (0.48) | | | 0.18* (0.10) | 0.20* (0.11) |
| Black | 0.06 (0.23) | | | 0.51** (0.21) | 0.49** (0.22) |
| Asian | 0.12 (0.33) | | | 0.45*** (0.15) | 0.37** (0.16) |
| Latino | 0.07 (0.25) | | | 0.28 (0.20) | 0.24 (0.21) |
| Observations | 89 | 88 | 89 | 89 | 88 |
| P-value: Joint Significance covariates | NA | 0.481 | 0.232 | 0.004 | 0.024 |

Table Notes: * Significant at the 10% level; ** Significant at the 5% level;*** Significant at the 1% level. Each specification represents results for a regression where the dependent variable is an indicator for participating in the experiment.

Table A3: Courses in Scale-up

Accountancy 1
Art 1B
Art 3B
Astronomy 4A
Astronomy 4B
Biology 22
Chemistry 1A (2 courses)
Criminal Justice 1
Criminal Justice 5
Economics 1B
Engineering 45 (3 courses)
English 50B
Ethnic Studies 11
Family and Consumer Sciences 10
Geology 7 (2 courses)
Health Science 50
History 5
History 7
History 17B
History 50
History 51
Math 24 (3 courses)
Music 8
Music 18
Philosophy 4 (2 courses)
Psychology 2 (3 courses)

Table A4: Scale-up Randomization Checks

| Specification | 1 | 2 | 3 |
|---------------------|------------------|------------------|-------------------|
| | All Students | Phase 1 | Phase 2 |
| Female | 0.04* (0.02) | 0.05* (0.02) | 0.02 (0.04) |
| White | -0.01 (0.03) | -0.02 (0.04) | 0.01 (0.04) |
| Asian | 0.01 (0.02) | 0.03 (0.04) | -0.02 (0.03) |
| Black | -0.03 (0.03) | -0.01 (0.04) | -0.06 (0.04) |
| Latino | 0.02 (0.02) | 0.03 (0.03) | 0.00 (0.02) |
| High School GPA | 0.02 (0.02) | 0.01 (0.03) | 0.03 (0.03) |
| Freshman | 0.03 (0.10) | -0.05 (0.15) | 0.13 (0.15) |
| Sophomore | 0.00 (0.08) | -0.04 (0.10) | 0.03 (0.13) |
| Junior | 0.00 (0.07) | -0.06 (0.06) | 0.08 (0.15) |
| College GPA | -0.03* (0.02) | -0.01 (0.02) | -0.07** (0.03) |
| Total College Units | 0.05 (0.12) | (0.05) (0.17) | 0.15 (0.20) |
| Observations | 3,922 | 2,350 | 1,572 |

Table A5: Survey Response Checks

| Specification | 1 | 2 | 3 |
|---------------------|-------------------|-------------------|-------------------|
| | All Students | Phase 1 | Phase 2 |
| Treatment | 0.02 (0.02) | 0.02 (0.02) | 0.01 (0.02) |
| Female | 0.08*** (0.01) | 0.09*** (0.01) | 0.08*** (0.02) |
| White | 0.03 (0.02) | 0.01 (0.03) | 0.06 (0.04) |
| Asian | 0.05** (0.02) | 0.05** (0.02) | 0.05 (0.04) |
| Black | 0.00 (0.02) | 0.01 (0.03) | -0.01 (0.04) |
| Latino | 0.02 (0.02) | 0.01 (0.03) | 0.03 (0.04) |
| High School GPA | 0.05*** (0.02) | 0.00 (0.01) | 0.11*** (0.02) |
| Freshman | 0.08 (0.07) | 0.09 (0.09) | 0.09 (0.13) |
| Sophomore | 0.08 (0.06) | 0.09 (0.06) | 0.08 (0.11) |
| Junior | 0.04 (0.03) | 0.03 (0.04) | 0.05 (0.04) |
| College GPA | 0.07*** (0.01) | 0.06*** (0.01) | 0.09*** (0.02) |
| Total College Units | 0.01 (0.07) | 0.03 (0.09) | 0.02 (0.13) |
| Observations | 3,922 | 2,350 | 1,572 |

Table A6-Panel A: Scale-up Results for Latino/Black Students (with controls)

| Specification | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|---|-------------------|---|------------------|-----------------|--|--|--|---|--|-------------------|--------------------------|
| Outcome | Grade | % Points Earned after First Feedback | Passed | A or B | How approachable was the instructor in class? | How available was the instructor outside of class? | How useful was the instructor's feedback in helping you learn? | How much do you believe the instructor cared about your success in the class? | How well did the instructor keep you informed about your progress in the class? | Dropped Course | N (Grades) N (Survey) |
| Treatment Effects for Black/Latino Students | | | | | | | | | | | |
| All Latino/Black | 0.08 [.116] | 0.03** [.062] | 0.02 [.378] | 0.04* [.054] | 0.27* [.088] | 0.22* [.06] | 0.17 [.272] | 0.35*** [.004] | 0.35*** [.002] | -0.01 [.164] | 1,619 403 |
| Female | 0.12* [.046] | 0.00 [.862] | 0.03 [.274] | 0.06* [.066] | 0.36** [.048] | 0.29* [.056] | 0.18 [.422] | 0.45*** [.002] | 0.45** [.014] | 0.00 [.994] | 938 271 |
| Male | -0.02 [.858] | 0.06** [.040] | -0.02 [.630] | -0.01 [.934] | 0.20 [.288] | 0.02 [.898] | 0.16 [.368] | 0.31 [.160] | 0.48** [.040] | -0.02* [.068] | 681 132 |
| Freshman | 0.25*** [.006] | 0.03 [.166] | 0.07* [.084] | 0.08* [.056] | 0.39** [.036] | 0.32** [.012] | 0.38*** [.010] | 0.54* * [.028] | 0.62*** [.008] | -0.02 [.198] | 710 180 |
| Upperclass | -0.03 [.718] | 0.04 [.144] | -0.02 [.458] | 0.01 [.848] | 0.20 [.386] | 0.18 [.304] | 0.07 [.732] | 0.25* [.078] | 0.19 [.166] | -0.01 [.418] | 909 223 |
| Freshman Female | 0.24** [.034] | 0.01 [.800] | 0.08 [.142] | 0.06 [.242] | 0.53** [.016] | 0.45*** [.010] | 0.45** [.030] | 0.74*** [.002] | 0.67*** [.008] | 0.01 [.536] | 435 127 |
| Freshman Male | 0.25 [.110] | 0.08 [.138] | 0.06 [.226] | 0.10 [.132] | 0.40 [.414] | 0.18 [.666] | 0.30 [.476] | 0.42 [.426] | 0.72 [.148] | -0.05* [.07] | 275 53 |
| Freshman w/Low HS GPA | 0.38*** [.002] | 0.05 [.144] | 0.12** [.028] | 0.11* [.062] | 0.51** [.022] | 0.47** [.016] | 0.20 [.392] | 0.49 [.160] | 0.49 [.124] | -0.02 [.172] | 402 94 |
| Freshman w/High HS GPA | 0.09 [.466] | 0.01 [.694] | 0.00 [.966] | 0.04 [.486] | 0.43 [.124] | 0.26 [.244] | 0.64** [.042] | 0.86*** [.002] | 0.87*** [.008] | 0.00 [.844] | 308 86 |

Notes: Each column reports results from a separate regression. All specifications include course by phase fixed effects. Square brackets contain p-values from randomization-based inference using a counterfactual of randomly assigning treatment status within classrooms 500 times. *Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level.

- * Significant at the 10% level
- ** Significant at the 5% level
- *** Significant at the 1% level

Table A6-Panel B: Scale-up Results for Non-Latino/Black Students (with controls)

| Specification | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|---|-----------------|---|-----------------|-----------------|--|--|--|---|--|-------------------|--------------------------|
| Outcome | Grade | % Points Earned after First Feedback | Passed | A or B | How approachable was the instructor in class? | How available was the instructor outside of class? | How useful was the instructor's feedback in helping you learn? | How much do you believe the instructor cared about your success in the class? | How well did the instructor keep you informed about your progress in the class? | Dropped Course | N (Grades) N (Survey) |
| Treatment Effects for non-Black/Latino Students | | | | | | | | | | | |
| All non-Black/Latino | 0.00 [.894] | 0.00 [.792] | 0.01 [.406] | 0.00 [.778] | 0.23** [.014] | 0.17** [.042] | 0.12 [.190] | 0.23*** [.008] | 0.23*** [.002] | 0.00 [.788] | 2,110 609 |
| Female | 0.03 [.576] | -0.02 [.220] | 0.01 [.608] | 0.01 [.662] | 0.35*** [.002] | 0.24** [.012] | 0.18* [.076] | 0.31*** [.008] | 0.28*** [.008] | 0.01 [.440] | 1,100 374 |
| Male | -0.03 [.570] | 0.02 [.354] | 0.01 [.598] | -0.02 [.358] | 0.12 [.294] | 0.14 [.314] | -0.02 [.894] | 0.24 [.232] | 0.26* [.056] | -0.01 [.216] | 1,010 235 |
| Freshman | 0.03 [.602] | 0.00 [.968] | 0.02 [.468] | 0.00 [.950] | 0.29** [.018] | 0.15 [.198] | 0.25* [.062] | 0.29** [.034] | 0.50*** [.002] | 0.00 [.900] | 802 223 |
| Upperclass | 0.00 [.914] | 0.00 [.792] | 0.01 [.438] | 0.00 [.990] | 0.22* [.060] | 0.15 [.246] | 0.06 [.610] | 0.23* [.074] | 0.09 [.314] | 0.00 [.926] | 1,308 386 |
| Freshman Female | 0.16* [.092] | 0.01 [.662] | 0.06* [.072] | 0.05 [.120] | 0.39** [.026] | 0.26 [.130] | 0.29 [.118] | 0.25 [.150] | 0.53** [.014] | 0.02 [.124] | 433 152 |
| Freshman Male | -0.11 [.308] | -0.01 [.700] | -0.04 [.278] | -0.05 [.266] | 0.22 [.346] | -0.09 [.720] | -0.15 [.536] | 0.62** [.012] | 0.34 [.348] | -0.03* [.080] | 369 71 |
| Freshman Low GPA (HS) | 0.02 [.846] | 0.02 [.596] | 0.07* [.074] | -0.08 [.104] | 0.55** [.044] | 0.39* [.092] | 0.40* [.094] | 0.48* [.066] | 0.63** [.042] | 0.00 [.844] | 366 87 |
| Freshman High GPA (HS) | 0.00 [.958] | 0.00 [.882] | -0.03 [.284] | 0.05 [.174] | 0.10 [.486] | 0.03 [.848] | 0.16 [.342] | 0.18 [.312] | 0.41* [.054] | 0.00 [.830] | 436 136 |

Notes: Each column reports results from a separate regression. All specifications include course by phase fixed effects. Square brackets contain p-values from randomization-based inference using a counterfactual of randomly assigning treatment status within classrooms 500 times. *Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level.

- * Significant at the 10% level
- ** Significant at the 5% level
- *** Significant at the 1% level

Appendix B—Research Protocols

Recruitment Letter

Dear [Faculty member’s first and last name]:

We are pleased to invite you to participate in an exciting opportunity on our campus. Recently, our Center for College & Career Readiness (CCR) was awarded a grant from the College Futures Foundation to explore innovations to improve college persistence and completion. As a result of this funding, the Center, in partnership with collaborators from University of California Davis, is continuing their work on an exciting faculty-based project. The project aims to carry out a “light-touch” intervention where faculty provide students with a few structured individualized emails about their progress in the course throughout the term.

The Co-PIs for the project are XXX, Director of the Center for College & Career Readiness and Professors Michal Kurlaender and Scott Carrell from University of California Davis. Based on initial review of classes, the Co-PIs have determined that your course (**Class Number – Course Title**), scheduled for fall semester, meets the criteria for inclusion in the project. Participation in the study is voluntary, but faculty who participate will be compensated for their time at a rate of \$500 for the term.

Please let us know if you are interested in participating by *replying all* to this email or by contacting XXX directly using the contact information below. We will schedule a follow up phone call with you to share additional details and to answer questions you may have.
[INSERT CONTACT INFO]

In keeping with the goals of our campus Strategic Plan, we are eager to find promising innovations to provide support to our students, and to increase our persistence and degree completion rates. We believe supporting all of our students through degree completion demands new innovations across our campus, and this effort to increase information about college success at the most micro level – the faculty member in the classroom – has great promise. We look forward to your participation.

Sincerely,

XXX

Provost & Vice President for Academic Affairs

XXX

Dean of Undergraduate Studies

FAQs with Participating Faculty

How are students chosen for additional email correspondence (the intervention) from faculty?

As required by our grant funding, the research team will randomly select students to receive the correspondence. The random selection methods differ based on the class size. For smaller classes with multiple sections taught by the same professor, we will randomly select one section where all students receive the correspondence. The second section will serve as the control group where no students will receive the correspondence. For large courses, we will randomly select a subset of students within the course to receive the correspondence. Randomization of students will enable us to determine whether the intervention has meaningful effects on student outcomes. Faculty will be notified which sections or students are to receive correspondence (treatment group) or not receive the additional correspondence (control group).

How many students will receive the intervention?

For small classes, all students will receive the correspondence. For large courses, no more than 70 students will be randomly selected to receive the emails.

How will you assess the outcome of the intervention?

Because the intervention is randomly assigned, we will compare outcomes of students who receive the correspondence to those who do not receive the additional emails.

How will student progress/performance be measured?

We will measure both short-run and medium-run outcomes. Short-run outcomes include completion of the course and course grade. The medium-run outcomes we plan to investigate include persistence into the subsequent term(s), and entry or persistence into specific college majors.

How many emails do you expect faculty to send during the semester as part of this program?

We request a minimum of three emails per student throughout the semester: 1) after the first assignment, 2) midway through the term, ideally after the first exam, and 3) end of the term.

Will the emails be structured by the researchers or faculty? Will faculty be able to tailor each message to each individual student?

Emails will be sent by the course instructor to students in the treatment group. Research assistance will be provided to work with faculty on email content that is tailored to specific course structure and instructor style. Below are examples:

Example of Email 1 from professor to student after submission of first assignment/exam

STUDENT PROFILE 1 (DID NOT TURN IN FIRST ASSIGNMENT)

Dear [Student Name],

I see you didn't turn in HW#1; although you only have to submit 5 out of the 7 homework assignments for full credit, the material in every homework is important for doing well in this class. I want to emphasize that doing well in this class requires coming to class regularly, completing the homework assignments, and seeking help when the concepts are unclear. To remind you, my office hours are: Monday and Wednesday 3:15 - 4:15 pm. Please feel free to come see your TA or me if you have any questions.

Sincerely, Professor X

STUDENT PROFILE 2 (TURNED IN FIRST ASSIGNMENT)

Dear [Student Name],

I noticed you did well [struggled with] on HW#1. Keep up the good work [Please make sure you] by coming to class regularly, completing future homework assignments, and seeking help when the concepts are unclear. To remind you, my office hours are: Monday and Wednesday 3:15 - 4:15 pm. Please feel free to come see your TA or me if you have any questions.

Sincerely, Professor X

Example of Email 2:

STUDENT PROFILE 1 (< C ACCEPTABLE PERFORMANCE)

Dear [Student Name],

We are approaching the mid point in the semester. I am concerned that based on your performance on the [Fill-in: Midterm, quiz, homework/?] you may be struggling in this course. However, don't be discouraged, there is still plenty of time to recover. To do well in the upcoming [Fill-in Final, next assignment, HW, ??] I encourage you to [fill-in: come to class regularly, review lecture notes, go to office hours].

Sincerely,

Professor X

STUDENT PROFILE 2 (B/ C PERFORMANCE)

Dear [Student Name],

We are approaching the mid point in the semester. You've done well so far on [Fill-in: Midterm, quiz, homework/?]. To strengthen your grade in the course and do well in the upcoming [Fill-in Final, next assignment, HW, ??], I want to encourage you to [fill-in: come to class regularly, review lecture notes, go to office hours].

Sincerely,

Professor X

STUDENT PROFILE 3 (B+ OR HIGHER PERFORMANCE)

Dear [Student Name],

We are approaching the mid point in the semester. You've done very well so far on [Fill-in: Midterm, quiz, homework/?]. To keep up your grade and do well in the upcoming [Fill-in Final, next assignment, HW, ??], I encourage you to [fill-in: come to class regularly, review lecture notes, go to office hours].

Sincerely,

Professor X

Email 3:

This will be structured as E-mail 2 with a focus on passing the course for students in the <C group.

What will I tell a student who didn't receive an email (but may be aware that their peer did) and inquires as to why?

If a student inquires about why they did not receive correspondence from you, please indicate that you are randomly emailing students regarding their performance in the course. Then proceed to advise the student as you usually would. Please make note of this if it occurs, as it will be tracked by the research team.

What else is required for participation?

We would like you to forward students' responses to the emails (if they occur). We would also appreciate access to your course grade book, any records of office hour attendance, and any other communication/interactions that may occur. We will also request your participation in a short survey at the end of the term.