American Economic Journal: Economic Policy 2023, 15(4): 113–141
https://doi.org/10.1257/pol.20210699

My Professor Cares: Experimental Evidence on the Role of Faculty Engagement†

By Scott E. Carrell and Michal Kurlaender*

We provide experimental evidence on the impact of specific faculty behaviors aimed at increasing student success for college students from historically underrepresented groups. The intervention was developed after conducting in-person focus groups and a pilot experiment. We find significant positive treatment effects across a multitude of short- and longer-run outcomes. Specifically, underrepresented students in the treatment report more positive perceptions of the professor and earned higher course grades. These positive effects persisted over the next several years, with students in the treatment more likely to persist in college, resulting in increased credit accumulation and degree completion. (*JEL* I22, I23, I28, J15, J44)

The rising value of a college degree has been well documented among social scientists (Pew Research Center 2014; Baum, Ma, and 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 and Baum 2017; Pew Research Center 2014; Snyder and Dillow 2013). Only 63 percent of first-time, full-time students finish a bachelor's degree within six years (NCES 2020). Moreover, many disparities by social origin and by institutional type exist in college access and college completion (Holzer and Baum 2017; Hoxby and Avery 2013; Bailey and Dynarski 2011). Despite a growing number of randomized control trials on improving college access, particularly for low-income and other underrepresented groups (Phillips and Reber 2019; Castelman, Page, and Schooley 2014; Carrell and Sacerdote 2017; Hoxby and Turner 2013; Bettinger, Long, Oreopoulos, and Sanbonmatsu 2012; Avery and Kane 2004; Barr and Castleman 2017), the research base is decidedly thin on how to keep students in college, and on improving college success and degree completion (Broda et al. 2018; Murphy et al. 2020). To help fill

* Carrell: University of Texas at Austin (email: scott.carrell@austin.utexas.edu); Kurlaender: UC Davis School of Education (email: mkurlaender@ucdavis.edu). C. Kirabo Jackson was coeditor for this article. We are grateful for funding from the College Futures Foundation. 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. This study is registered in the AEA RCT Registry (#0005875), see Carrell and Kurlaender (2020a). We obtained IRB approval from the University of California, Davis (IRB ID: 808387-1). All errors are our own.
† Go to https://doi.org/10.1257/pol.20210699 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

this gap, this study provides experimental evidence on the impact of specific faculty behaviors aimed at increasing student success in the classroom for college students from historically underrepresented groups.

To date, 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 performance, 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 and Petronijevic 2019; Bettinger and Baker 2014; Angrist et al. 2009), and relative performance feedback (Azmat, Bagues, Cabrales, Iriberri 2019). 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 (Broda et al. 2018; Murphy et al. 2020; Walton and Cohen 2011; Yaeger et al. 2016); yet similar interventions fail to replicate in other settings (Dobronyi et al. 2019; Oreopoulos and Petronijevic 2019; Broda et al. 2018). Other studies have explored grade incentives on additional skill development in introductory courses (Pozo and Stull 2006) and technological innovations in the classroom (Ball, Eckel, and Rojas 2006), and their impacts on student learning, course performance and course satisfaction. Notably, nearly all of the college performance and persistence interventions have been targeted at students, neglecting a potential key input in the education production function—faculty. This is critical, since short of changing the professor labor supply at scale, improving effectiveness among faculty may be key to increasing student persistence.

A growing body of literature suggests that college instructors matter (Braga et al. 2014; Carrell and West 2010). In particular, prior work has demonstrated that demographic characteristics of professors such as gender (Carrell et al. 2010; Price 2010; Hoffmann and Oreopoulos 2009) and race (Fairlie et al. 2014; Price 2010) can influence student performance and attainment in particular courses. Instructor status (i.e., adjunct employment or academic rank) (Ran and Xu 2018; Figlio, Schapiro, and Soter 2015; Bettinger and Long 2006; Carrell and West 2010; Ehrenberg and Zhang 2005) as well as student evaluations (Braga et al. 2014; Beleche et al. 2012; Carrell and West 2010) have been established as predictors of both contemporaneous and longer run outcomes of students. Yet, prior studies on college instructors have focused almost exclusively on their innate traits, job features, or unobservable characteristics. 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, even the Brownback and Sadoff (2019) study did not incentivize particular types of actions on the part of faculty. In fact, the literature, as a whole, leaves the question about *how* faculty could improve their effectiveness largely unanswered.

In this paper 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, economics courses, 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 change in faculty *behavior* toward students.

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; and the full-scale implementation.[1] The "light-touch" intervention consisted of two to three strategically timed personalized emails 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 for additional supports through office hours. 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 a large, broad-access, four-year institution where we conducted the focus groups. We implemented the intervention with 22 faculty members teaching large classes in 20 different course subjects and nearly 3,000 students during the spring of 2016 and fall of 2017. Results show significant positive treatment effects for the target population of students across a multitude of short- and longer-run outcomes. Underrepresented students in the treatment reported more positive perceptions of the professor and course, and they earned higher grades. The intervention also resulted in positive spillovers on grades in other courses during the same term. These positive effects persisted over the next several years, with underrepresented students in the treatment group more likely to persist in college, resulting in increased credit accumulation and graduation. We conclude that targeted feedback from professors can lead to meaningful gains in achievement for historically underrepresented students.

Our study provides three important contributions to the literature. First, we design and test an intervention specifically focused on underrepresented college students developed through targeted focus groups and a pilot intervention prior to launching a full-scale experiment. As such, our positive findings may provide a potential roadmap for future education research intervention design at scale that aims to address inequalities in college access and persistence. Second, our intervention deployed specific changes to instructor *behaviors* shifting some of the onus of course success from students to faculty. Our initial focus groups revealed that students of color report increased success when they felt more connected to their instructor and when they understood class expectations. Results from our experiment confirm these convictions as students in the treatment group earn higher grades and

[1] In NBER working paper 27312 (Carrell and Kurlaender 2020), we also show results from a replication of the pilot, which also showed positive treatment effects.

report more positive perceptions of the instructor and course when faculty increased their personalized engagement. Finally, and perhaps most importantly, we show that a light-touch intervention aimed at increasing achievement in one course contributed to positive spillovers in other courses, and it had a lasting impact, resulting in increased persistence, credit accumulation, and eventual graduation.

I. Faculty Feedback and Student Engagement

A. *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 their college instructors both in and outside of class (i.e., asking questions during or after class and coming to office hours). Interestingly, virtually all of the students we interviewed reported a close rapport with their high school teachers, but when reflecting on their experience in college, many described such a connection to be rare. Second, students felt unsure of what they needed to do to be more successful in their courses, something they also described as a departure from the success they felt as high school students where they believed expectations in their courses were clearer. In short, students did not believe most college instructors were accessible, clear about their expectations of students, or supportive of their learning.

The findings from our focus groups were consistent with extant social-psychological theory on the role of uncertainty in predicting college success (Murphy et al. 2020). Students expressed the most academic success in courses where they felt instructors were accessible to them in and outside of class and when faculty communicated expectations clearly. Students' beliefs about college, and how they interpret early difficulties, can have important consequences on postsecondary success (Murphy et al. 2015; Murphy et al. 2020; Good et al. 2012). As such, the students we aimed to target—Black and Latinx first-year students—may be particularly responsive to efforts to reduce uncertainty and improve faculty engagement in the classroom.

Moreover, a significant body of theoretical and descriptive work in higher education has focused on the role of faculty-student interactions, noting their positive association with a host of college outcomes (e.g., persistence, performance, graduate school enrollment) (Tinto 1993; Astin 1993, 1999; Kuh and Hu 2001; Kim and Sax 2017). These studies posit that faculty-student interactions are critical forms of students' development and socialization, particularly in the first year, as they navigate the higher education context (Kim and Sax 2017). In particular, Tinto's model of student departure theorizes that student-faculty interaction is key for academic integration, which can support student persistence (Tinto 1993). Finally, Rendon's

(1994; 2002) theory on student validation proffers that such validation comes from faculty that affirm students' capacity to succeed. These theories are applied largely through correlational work that relies on student surveys of their experiences in college, including self-reported interactions with faculty.

B. *Developing an Intervention—Theory and Prior Research*

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, 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, affirmation, and belonging. More explicitly, the intervention rests upon a key premise: faculty are an important and (potentially) underutilized resource to increase student success more generally and for historically underrepresented students 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 belonging in the college classroom, and also influence their decision to persist toward, 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—Black and Latinx students.

Students in the treatment condition received emails 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 grades. We also tested potential mechanisms 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 will affirm their sense of self-efficacy and belonging and improve college success. 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 and

Kane 2004).[2] However, recent work in behavioral economics is more critical of rational choice, and it 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 and 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 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 and 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 (Hattie and Timperley 2007; Kluger and DeNisi 1998). Feedback at the "process level" has been found to be particularly effective (Balzer, Doherty, and 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 and Knapp 1991). A critical mediator to feedback is the perception of self-efficacy (Hattie and Timperley 2007; Kluger and DeNisi 1998). That is, feedback is principally 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' self-efficacy, help-seeking behavior, and increased belonging. 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 and performance (Bandura 1993; Yeager and Walton 2011), and improve social belonging (Murphy et al. 2020). 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 and persistence in college.

Research from social psychology has also established the important role of social belonging and other social-psychological determinants of students' educational

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

success in college (Murphy et al. 2020; Walton and Brady 2017; Strayhorn 2012). A key challenge for first-generation and other minoritized populations in higher education is to feel successful in navigating what can be referred to as the "hidden curriculum" of college and universities and to feel like they belong. Using nationally representative data, Gopalan and Brady (2019) find that underrepresented minority and first-generation students enrolled at four-year colleges report lower levels of belonging in college when compared to their White, Asian, and continuing generation peers.

Several interventions have been successful in improving social belonging to support student persistence and success for historically marginalized groups. Treatment conditions focus on normalizing the common challenges (academic and social-psychological) that students may face in college through stories from more advanced students, and in some cases through a writing activity allowing students to connect these stories to their own lived experiences (Murphy et al. 2020). Nevertheless, results from replication of social belonging interventions have been met with mixed results (Broda et al. 2018; Oreopoulos et al. 2020).

II. Piloting the Intervention

A. *Pilot Design*, *Data*, *and Methods*

To test the efficacy of our designed intervention, we piloted the concept in 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 five of seven 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.[3]

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 emails 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.

The first email to the treatment group was sent as a result of failing the first homework assignment. The second email to the treatment group was sent after the first midterm exam, and feedback to students was based on their exam performance.[4]

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

[4] Students that received a B+ or higher received an email commending them on a job well done and reminding them of the professor's office hours. Students that received between a C− and B received an email telling them what their grade in the course is likely to be based on this midterm performance, and highlighting that it is not too late to improve their grade and the set of behaviors that will help them be successful in the course, as well as reminding the student of the professor's office hours. Students that received lower than a C− on the midterm received an email warning that based on his/her trajectory, the student may be at risk of failing the course, but also reminding them there is time to recover, and providing details on the behaviors that would allow them to pass the course successfully, along with a reminder to seek additional supports and the professor's office hours. A fourth group of five students, who had dropped out of the course from the treatment group at the time the second email was sent, received no email.

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.

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.[5]

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 percent male, 89 percent California residents, 26 percent of students are first-generation college students, and 23 percent of students are underrepresented minorities (Table 1). 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 provide evidence that randomization created groups that were equal in expectation for receipt of the treatment.

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. We investigate several outcomes: exam grades, total course score and grade, homework score, time spent on homework, office hour attendance, attitudinal measures toward the course and professor, and course completion.[6]

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 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.[7] 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 allow 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, underrepresented minority status, residency status, entering cohort year, and high school GPA.

[5] Data and replication files for this project can be found at https://doi.org/10.38886/E169341V1.

[6] 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.

[7] Importantly, students do not choose their TA, as the TA's are assigned to sections after the student's primary registration period ends.

Table 1—Pilot Descriptive Statistics

Variable	Number of observations	Mean	SD	Min.	Max.
Midterm 1 (pps)	53	0.74	0.17	0.33	1
Midterm 2 (pps)	53	0.68	0.20	0	1
Final exam (pps)	53	0.64	0.15	0	0.95
All exams (pps)	53	0.68	0.14	0	0.93
Total course score (pps)	53	0.72	0.13	0.21	0.94
Course grade (0–4)	53	2.40	0.89	0	4
Homework score (pps)	53	0.93	0.20	0	1
Homwork points earned (pps)	53	0.55	0.21	0	0.88
Homwork 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
Professsor cares about my performance	51	2.39	0.87	0	3
Motivated to do well in course	50	3.14	0.83	1	4
Total office hour visits (number)	53	2.32	2.29	0	9
Dropped out of course (pct)	69	0.23	0.43	0	1
Male	53	0.68	0.47	0	1
First-generation college goer	53	0.26	0.45	0	1
HS GPA	53	3.77	0.37	2.87	4.24
Underrepresented 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

B. *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 drop out 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 email 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.[8]

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 cared about their performance but less likely to report that they are

[8] The grade effects are conditional on course completion. To bound this estimate, we estimated the effect assuming all students who dropped the course would have failed. Doing so increases the estimate from 0.431 to 0.622 ($p = 0.043$).

Table 2—Pilot Results

Outcome	Midterm 1 (pps) 1	Midterm 2 (pps) 2	Final exam (pps) 3	All exams (pps) 4	Total course score (pps) 5	Course grade (0–4) 6
Panel A. Test score outcomes on exams						
No controls	0.065	0.121	0.022	0.063	0.064	0.431
	(0.048)	(0.054)	(0.042)	(0.039)	(0.034)	(0.239)
TA fixed effects	0.073	0.150	0.042	0.082	0.078	0.521
	(0.052)	(0.057)	(0.043)	(0.041)	(0.036)	(0.249)
Individual controls and TA fixed effects	0.057	0.136	0.049	0.076	0.076	0.501
	(0.053)	(0.060)	(0.042)	(0.041)	(0.037)	(0.254)
Observations	53	53	53	53	53	53

Outcome	Homework score (pps) 1	Homwork points earned (pps) 2	Homwork total time spent (hours) 3	Homework median time spent (hours) 4
Panel B. Homework scores and time spent				
No controls	0.103	0.067	1.804	0.272
	(0.057)	(0.055)	(1.131)	(0.159)
TA fixed effects	0.119	0.052	1.794	0.257
	(0.062)	(0.060)	(1.242)	(0.175)
Individual controls and TA fixed effects	0.152	0.075	1.969	0.311
	(0.062)	(0.061)	(1.333)	(0.186)
Observations	53	53	53	53

Outcome	Professor office hour visits (number) 1	"Professsor cares about my performance" 2	"Motivated to do well in course" 3	TA office hour visits (number) 4	Dropped out of course 5
Panel C. Mechanisms					
No controls	0.131	0.540	−0.422	0.967	−0.123
	(0.103)	(0.332)	(0.323)	(0.128)	(0.102)
TA fixed effects	0.101	0.549	−0.405	0.802	NA
	(0.105)	(0.360)	(0.347)	(0.672)	
Individual controls and TA fixed effects	0.093	0.535	−0.237	NA	NA
	(0.104)	(0.387)	(0.367)		
Observations	53	51	50	53	69

Notes: Each cell represents the results from regressing the outcome listed on a treatment dummy variable. Specifications 2 and 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, underrepresentated minority, CA resident, entering cohort, and high school GPA.

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 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 indicates 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 toward this individual attention are instructive, examples of this feedback can be found in Carrell and Kurlaender (2020b). It is worth noting that these comments suggest that students are appreciative primarily of the contact between them and the professor, rather than the information provided itself. Moreover, these emails 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. Importantly, we can rule out potential concerns with endogenous subjective grading from the professor or TAs. Homework assignments were graded electronically, while the TAs graded the exams. Importantly, the TAs were not made aware of the pilot experiment, so there was no chance of preferential treatment based on treatment status.

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.

III. Scale-Up

The scale-up was implemented in two separate waves during the spring of 2016 and fall of 2017 at the same broad-access, four-year institution where the original focus groups took place.[9] 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.[10] In total, 22 faculty members across 20 different course subjects participated in the study, with nearly 3,000 total students in the treatment and control groups. All participating faculty were given templates of emails that they were encouraged to personalize to their own courses (available in online 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 instructor 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

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

[10] All recruitment materials available from the authors upon request.

of the experiment, with an additional $100 gift card for completing a survey at the end of the semester.[11]

A. *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, data availability, and the overall relatively low graduation rate at the institution we study, rather than condition on the first assignment (or sample within our target population), we chose to randomly select students from the entire course. Hence, although our intervention was specifically designed for and tailored based on feedback from underrepresented students, we sampled among all students within each course using a blocked randomization research design at the course level. Doing so allows to estimate average treatment effects for not only the targeted group, but the entire population of students in the selected courses.

In the spring of 2016 the treatment group comprised of a randomly selected one-half of the students in 20 large undergraduate classes and in the fall of 2017 five additional large classes were added to the study, where we randomly selected one-third of students into treatment.[12] 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 a 16-week semester system courses).[13]

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).[14] Comparing professor characteristics in Table A2 shows there are no significant differences in Rate My Professor ratings of participating versus nonparticipating professors. However, participating professors are significantly *more* likely to be Asian. Additionally, though not precisely estimated, the coefficients on female, Black and

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

[12] When examining the results of our experiment, we find no differential effects (statistically or economically) across the two phases. Additionally, in the Fall of 2017, we randomly selected the entire class to receive treatment in the nine cases where the professor taught two small sections of the identical course. We implemented this "matched-pair" design in an attempt to examine whether potential spillovers within a class bias the treatment effects. However, this design was severely underpowered due to large intra-cluster correlations (e.g., classroom level common shocks). Power analysis for a blocked design while modeling the intra-cluster correlation coefficient estimated in our data (0.009) and class sizes of 35 students indicates that a total of at least 34 professors teaching two classes of the same course would be required to detect a moderate effect size of 0.15 grade points with a power of 0.80. Given this, we have excluded these nine courses from our main tables. These results in including these courses can be found in online Appendix Table A8 and A9.

[13] The timing of these 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 feedback 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.

[14] Similarly, the external validity of our estimates could be affected by the fact that professors were paid to participate in the study.

Latino are positive and economically meaningful. It is unclear, however, how this selection may affect external validity.

B. *Scale-Up Data and Methods*

Our sample consists of a broad set of academic subjects from art to engineering (Table A3 in the online 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 35 percent of the sample. Fifty-five percent of students are female and 43 percent are first-time freshman. The average student is midway through their sophomore year, having completed just under 44 units, with an average high school GPA of 3.29.

To test balance across treatment and control groups, we present results from models regressing treatment status on our full set of observable pretreatment characteristics (panel A) and when regressing predicted grades[15] on treatment status (panel B). Additionally, because the focus of the study is underrepresented students, we show results from these balancing tests for the entire sample as well as for our specific subgroups of interest: all underrepresented minority students (URM), underclass (freshman and sophomore) URM students, and upperclass (junior and senior) URM students. Results of these balancing tests are presented in online Appendix Table A4.[16] Panel A shows that (pretreatment) student characteristics are largely uncorrelated with treatment status in both the full sample as well as for our subgroups of URM students, with only 3 of 31 coefficients significant at the 10-percent level. Notably, all three statistically significant coefficients are on the indicator variable for being female, indicating that female students are overrepresented in the treatment group. Results in panel B show that predicted grades are not significantly correlated with treatment status for either the full sample or the subsample of URM students. We do note that there is a relatively small and marginally significant relationship ($p = 0.100$) between treatment status and predicted grades for our subgroup of underclass URM students, as shown in column 3. As such, for all of our estimated treatment effects we present our main findings controlling for the full set of pretreatment characteristics. We also show unconditional estimates in the online Appendix tables.

To assess plausible mechanisms of treatment effects, we administered a survey at the end of the semester on student perceptions of the professor and course. The survey was administered primarily by email as well as in-person by our research assistants in select classes. Students were incentivized to complete the survey by being entered into a lottery to win an iPad or Amazon gift cards. 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 26.4 percent, which is quite similar to other college surveys (Carrell and Sacerdote 2017). Table A5 in the online

[15] We predict course grade for the control group on observable pretreatment characteristics and classroom fixed effects.

[16] As discussed in our methods section, for statistical inference in addition to reporting robust standard errors in parentheses clustered by classroom, square brackets contain empirical p-values from randomization-based inference using a counterfactual of randomly assigning treatment status within classrooms 500 times.

Table 3—Scale-Up Descriptive Statistics

	Full sample	All students treatment	All students control	All URM students treatment	All URM students control	Underclass URM students treatment	Underclass URM students control	Upperclass URM students treatment	Upperclass URM students control
Black	0.10	0.09	0.10	0.20	0.25	0.20	0.25	0.23	0.23
	(0.30)	(0.29)	(0.31)	(0.40)	(0.43)	(0.40)	(0.44)	(0.42)	(0.42)
Latino	0.35	0.36	0.33	0.82	0.79	0.82	0.79	0.79	0.79
	(0.48)	(0.48)	(0.47)	(0.39)	(0.41)	(0.38)	(0.41)	(0.41)	(0.41)
Asian	0.26	0.26	0.25	0.02	0.02	0.02	0.01	0.04	0.03
	(0.44)	(0.44)	(0.43)	(0.15)	(0.13)	(0.13)	(0.11)	(0.20)	(0.16)
Female	0.55	0.58	0.53	0.62	0.58	0.64	0.58	0.57	0.56
	(0.50)	(0.49)	(0.50)	(0.49)	(0.49)	(0.48)	(0.49)	(0.50)	(0.50)
High school GPA	3.29	3.30	3.28	3.25	3.21	3.26	3.20	3.17	3.25
	(0.42)	(0.42)	(0.42)	(0.38)	(0.40)	(0.38)	(0.40)	(0.39)	(0.42)
Prior college GPA	2.87	2.86	2.88	2.78	2.76	2.79	2.74	2.76	2.83
	(0.65)	(0.65)	(0.66)	(0.64)	(0.66)	(0.67)	(0.70)	(0.55)	(0.51)
Total college units	43.79	43.27	44.22	39.84	41.41	25.09	25.88	85.56	84.14
(pre-treatmnet)	(32.99)	(33.02)	(32.97)	(30.81)	(31.43)	(14.79)	(15.83)	(21.11)	(22.92)
Freshman	0.43	0.44	0.42	0.47	0.43	0.62	0.59	0.00	0.00
	(0.49)	(0.50)	(0.49)	(0.50)	(0.50)	(0.49)	(0.49)	0.00	0.00
Sophomore	0.27	0.27	0.27	0.29	0.30	0.38	0.41	0.00	0.00
	(0.45)	(0.45)	(0.45)	(0.45)	(0.46)	(0.49)	(0.49)	0.00	0.00
Junior	0.19	0.17	0.20	0.16	0.18	0.00	0.00	0.64	0.67
	(0.39)	(0.38)	(0.40)	(0.36)	(0.38)	0.00	0.00	(0.48)	(0.47)
Course grade	2.51	2.52	2.49	2.40	2.27	2.40	2.23	2.39	2.35
	(1.19)	(1.19)	(1.19)	(1.25)	(1.23)	(1.25)	(1.25)	(1.26)	(1.16)
Percent points earned after	0.70	0.71	0.69	0.68	0.65	0.67	0.64	0.70	0.69
first feedback	(0.23)	(0.21)	(0.25)	(0.22)	(0.26)	(0.22)	(0.26)	(0.22)	(0.25)
Passed course (>D)	0.81	0.81	0.81	0.77	0.76	0.77	0.74	0.77	0.80
	(0.39)	(0.39)	(0.39)	(0.42)	(0.43)	(0.42)	(0.44)	(0.42)	(0.40)
Course grade A/B	0.56	0.57	0.56	0.53	0.47	0.52	0.47	0.55	0.48
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Dropped course	0.04	0.04	0.04	0.03	0.04	0.02	0.04	0.04	0.04
	(0.19)	(0.19)	(0.20)	(0.16)	(0.20)	(0.14)	(0.20)	(0.20)	(0.19)
"Quit"	0.07	0.04	0.09	0.06	0.11	0.06	0.11	0.05	0.09
	(0.25)	(0.21)	(0.28)	(0.23)	(0.31)	(0.24)	(0.31)	(0.22)	(0.29)
Grades in other courses	2.78	2.79	2.76	2.70	2.63	2.68	2.59	2.77	2.77
	(0.94)	(0.92)	(0.96)	(0.91)	(0.96)	(0.93)	(0.97)	(0.85)	(0.95)
Persist 1 semester later	0.93	0.94	0.92	0.96	0.91	0.96	0.89	0.95	0.96
(or graduate)	(0.26)	(0.23)	(0.27)	(0.20)	(0.29)	(0.19)	(0.32)	(0.22)	(0.21)
Persist 2 semesters later	0.87	0.88	0.86	0.90	0.84	0.88	0.81	0.93	0.92
(or graduate)	(0.34)	(0.32)	(0.35)	(0.31)	(0.37)	(0.32)	(0.39)	(0.26)	(0.27)
Persist 3 semesters later	0.81	0.82	0.80	0.83	0.77	0.80	0.73	0.92	0.89
(or graduate)	(0.39)	(0.38)	(0.40)	(0.37)	(0.42)	(0.40)	(0.44)	(0.28)	(0.32)
Persist 4 semesters later	0.80	0.81	0.79	0.83	0.76	0.80	0.73	0.89	0.86
(or graduate)	(0.40)	(0.40)	(0.41)	(0.38)	(0.43)	(0.40)	(0.45)	(0.31)	(0.35)
Persist 5 semesters later	0.77	0.77	0.76	0.78	0.74	0.76	0.70	0.85	0.86
(or graduate)	(0.42)	(0.42)	(0.43)	(0.41)	(0.44)	(0.43)	(0.46)	(0.36)	(0.35)
Persist 6 semesters later	0.76	0.76	0.75	0.76	0.72	0.73	0.68	0.85	0.83
(or graduate)	(0.43)	(0.43)	(0.43)	(0.43)	(0.45)	(0.44)	(0.47)	(0.36)	(0.37)
Persist 7 semesters later	0.75	0.75	0.74	0.76	0.71	0.73	0.67	0.84	0.82
(or graduate)	(0.44)	(0.43)	(0.44)	(0.43)	(0.45)	(0.44)	(0.47)	(0.37)	(0.38)
Total units earned as	114.78	116.36	113.47	115.08	109.52	107.23	99.51	139.42	137.05
of Fall 2020	(41.03)	(40.58)	(41.36)	(38.69)	(43.59)	(39.30)	(44.78)	(23.92)	(24.11)
Graduate	0.63	0.65	0.62	0.64	0.58	0.59	0.51	0.80	0.79
	(0.48)	(0.48)	(0.48)	(0.48)	(0.49)	(0.49)	(0.50)	(0.40)	(0.41)
Student-year observations	2,918	1,322	1,596	582	675	440	495	142	180

Appendix provides results from models regressing the probability of response on student background characteristics. Unsurprisingly, response is positively correlated with college GPA and gender. Importantly, response is not significantly correlated with treatment status.[17]

Our outcome measures consist of both short-run academic performance measures during the semester of the interventions as well as longer-run outcomes measuring persistence and graduation through the fall semester of 2020. The primary short-run 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 2.51, with 81 percent 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 70 percent.

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:

$$Y_{ict} = \alpha + \beta\, treat_{ict} + \gamma X_i + \lambda_{ct} + \varepsilon_{ict}, \tag{1}$$

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. Our main specifications include a dummy indicator for treatment status classroom fixed effects (λ_{ct}), as well as student-level controls.[18] The classroom fixed effects are used to account for unobserved differences across classes/instructors. Individual control variables include: cumulative units earned; college GPA; high school GPA; and indicators for gender, race/ethnicity, and year in college. 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 the null hypotheses in which random sampling is used to assign treatment status. Hence, in addition to reporting robust standard errors in parentheses clustered by classroom, square brackets contain empirical p-values from randomization-based inference using a counterfactual of randomly assigning treatment status within classrooms 500 times.[19]

[17] We find qualitatively similar results for our estimated treatment effects on survey outcomes when reweighting our estimates by the inverse probability of response.

[18] Online Appendix Tables A6 and A7 report results when excluding student-level controls.

[19] 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."

C. *Scale-Up Results*

Short-Run Outcomes.—Findings of our short-run treatment effects for the scale-up intervention are presented in Table 4. Each column presents results for different outcomes. Our focus groups targeted 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, results in row 1 include the entire sample of students, while rows 2–4 present results for our targeted samples of students—URM students, underclass URM students, and upperclass URM students.[20]

Results in row 1, columns 1–4 show the treatment had small positive, but statistically insignificant, effects on *overall* student performance in the course. Across all measures of achievement, including course grade, percentage of points earned in the course,[21] passing the course, or earning an A or B grade, we find positive and insignificant effects. However, when examining results for the target group of students in rows 2–4, we find moderately sized positive and significant treatment effects for several of our short-run academic outcomes. Here we first note positive average treatment effects on course grades, which is largely driven by underclass (freshman and sophomore) URM students. For this group, students in the treatment significantly outperformed control students by nearly one-sixth a letter grade (0.143).[22] Additionally, URM students in the treatment are 4.9 percentage points ($p = 0.046$) more likely to earn an A or B in the course.

In columns 5–6 we examine two measures of student effort/participation in the course. Column 5 shows results for a measure of "giving-up" in the class, which is an indicator for whether the student earns less than 25 percent of the points in the course after the initial feedback.[23] Column 6 shows results for dropping the course. While we find no significant average treatment effects on dropping the course for the full sample, we find that underclass students in our target group are 2.8 percentage points less likely to drop the course ($p = 0.012$). Additionally, we also find a large and statistically significant negative effect of 4.4 percentage points on our measure of "giving up" for the entire sample, which represents a nearly 50 percent decrease from the control mean of 9 percent. This effect is relatively consistent in magnitude across our targeted subgroups, though measured with less precision.

[20] We have also examined treatment effects by gender and find no appreciable differences. These results can be found in the NBER working paper version of the paper.

[21] Estimates for the percentage of points earned in the course are likely a lower bound. This is due to the fact that the effects on course grade for the sample of courses where we have grade books is smaller than that of the full sample.

[22] Course grades are conditional on course completion. To bound these estimates, in results available upon request, we estimated the effects assuming all students who dropped the course would have failed. Under this assumption our positive treatment effects are larger across all subgroups. The effect for all students increases from 0.047 to 0.075 grade points ($p = 0.042$), for URM students the effect increases from 0.112 to 0.156 grade points ($p = 0.006$), and the effect for underclass URM students increases from 0.143 to 0.2981 grade points ($p = 0.014$).

[23] Our preregistration did not specify the outcome of "giving up." Specifically, we prespecified the following short-run outcomes: persistence in the course, completion of course with a passing grade, actual grade in the course.

Table 4—Scale-Up Results: Short-Run Outcomes

Outcome Specification	Grade 1	Percent points earned after first feedback 2	Passed (>D) 3	A or B 4	"Gave up" 5	Dropped course 6	Grades in other courses 7	N (grades) N (*dropped*)
Panel A. Grade outcomes								
All students	0.049	0.015	0.010	0.019	−0.044	−0.010	0.057	
	(*0.047*)	(*0.011*)	(*0.017*)	(*0.020*)	(*0.011*)	(*0.009*)	(*0.030*)	2,768
	[0.166]	[0.270]	[0.454]	[0.228]	[0.006]	[0.148]	[0.052]	*2,914*
All URM students	0.112	0.029	0.015	0.049	−0.043	−0.016	0.083	
	(*0.061*)	(*0.017*)	(*0.026*)	(*0.027*)	(*0.022*)	(*0.013*)	(*0.054*)	1,197
	[0.054]	[0.136]	[0.498]	[0.046]	[0.082]	[0.100]	[0.062]	*1,257*
Underclass URM students	0.143	0.030	0.033	0.045	−0.043	−0.028	0.090	
	(*0.066*)	(*0.019*)	(*0.023*)	(*0.032*)	(*0.026*)	(*0.012*)	(*0.071*)	891
	[0.040]	[0.168]	[0.208]	[0.108]	[0.142]	[0.010]	[0.110]	*935*
Upperclass URM students	0.004	−0.019	−0.033	0.053	−0.015	0.009	0.052	
	(*0.126*)	(*0.043*)	(*0.058*)	(*0.057*)	(*0.045*)	(*0.028*)	(*0.100*)	306
	[0.968]	[0.672]	[0.478]	[0.332]	[0.808]	[0.738]	[0.606]	*322*

Outcome Specification	How **approachable** was the instructor in class? 8	How **available** was the instructor outside of class? 9	How useful was the instructor's **feedback** in helping you learn? 10	How much do you believe the instructor **cared** about your success in the class? 11	How well did the instructor keep you **informed** about your progress in the class? 12	N (survey)
Panel B. Survey outcomes						
All students	0.195	0.144	0.081	0.294	0.311	
	(*0.077*)	(*0.048*)	(*0.078*)	(*0.107*)	(*0.103*)	
	[0.010]	[0.026]	[0.326]	[0.000]	[0.000]	732
All URM students	0.176	0.160	0.083	0.410	0.360	
	(*0.149*)	(*0.129*)	(*0.152*)	(*0.194*)	(*0.219*)	
	[0.186]	[0.126]	[0.532]	[0.002]	[0.010]	294
Underclass URM students	0.192	0.166	0.085	0.463	0.396	
	(*0.190*)	(*0.161*)	(*0.212*)	(*0.237*)	(*0.275*)	
	[0.168]	[0.142]	[0.588]	[0.004]	[0.028]	220
Upperclass URM students	0.370	0.350	0.343	0.326	0.367	
	(*0.315*)	(*0.270*)	(*0.242*)	(*0.345*)	(*0.293*)	
	[0.258]	[0.202]	[0.292]	[0.374]	[0.234]	74

Notes: Each column reports results from a separate regression. All specifications include classroom fixed effects and individual controls race, gender, high school and pretreatment college GPA, pretreatment units earned, and year of schooling (e.g., freshman, sophomore, or junior). Standard errors in parentheses are clustered at the course by phase level. Square brackets contain p-values from randomization-based inference using a counterfactual of randomly assigning treatment status within classrooms 500 times.

Next, in column 7 we present results for student grades in courses other than those in the experimental study taken during the same semester. We do so to examine whether professor engagement in one course affects academic performance in treated students' other (non-treated) courses. On the one hand, the positive treatment effects we observe for the targeted group 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 (or even a negative impact). 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. Overall, the pattern of results shows evidence of potential positive spillovers from treated courses to non-treated courses. Moreover, for URM students, the magnitude of the treatment effect on grades in non-treated courses is roughly three-quarters the size of the effects in treated courses and statistically significant at the 0.10-level.

Finally, in columns 8–12 we present results from survey responses measuring perceptions of the professor and course. These results show evidence of a strong positive average 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, and the extent to which a student felt supported and informed. The largest treatment effects for both the full sample and the target population 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 URM students in the treatment group, on average, responding over a third of a standard deviation higher than students in the control condition.

Given the positive treatment effects on course performance, a natural question is whether these effects are driven by: (1) students feeling the professor cares about them, and/or (2) the specific information provided by the instructor? Although our experiment was not designed to distinguish between the two, we note that the question asking about the *usefulness* of professor feedback is not statistically significant and of much smaller magnitude for both the full sample and the targeted population. Whereas the question asking about whether you believed the instructor cared about your success in the class was statistically significant for the full sample and the targeted population.

Overall, these results suggest that the treatment had a positive effect on students' perceptions of instructor support. However, this only translated to significantly higher course grades for students in the target population who are early in their college career,[24] suggesting that students may interpret targeted emails from the professor differently depending on their background and previous college experience.

Longer-Run Treatment Effects.—A natural question is whether the positive treatment effects persisted to longer-run outcomes? Evidence of positive spillovers to other course grades suggests this relatively light-touch intervention may result in improved longer-run outcomes for students in the treatment. Table 5 presents results for measures of persistence, credit accumulation, and graduation through the fall of 2020. Similar to Table 4, each column represents results from a different outcome, while rows show results for our different samples of students.

[24] In results not presented, when examining treatment effects separately for White and Asian students, we find null effects on course grades and our other measures of academic performance; we find a negative and marginally significant effect on our measure of "giving up;" and we find large, positive, and significant effects on the survey responses measuring perceptions of the professor and course.

Table 5—Scale-Up Results: Long-Run Outcomes

Outcome	Persist 1-semester later (or graduate)	Persist 3-semesters later (or graduate)	Persist 5-semesters later (or graduate)	Persist 7-semesters later (or graduate)	Total units earned as of Fall 2020	Graduate by Fall 2020
Specification	1	2	3	4	5	6
All students	0.027	0.022	0.013	0.012	2.771	0.009
	(*0.007*)	(*0.010*)	(*0.012*)	(*0.017*)	(*1.311*)	(*0.015*)
	[0.002]	[0.118]	[0.408]	[0.426]	[0.020]	[0.528]
Observations	2,914	2,914	2,914	2,914	2,914	2,914
All URM students	0.050	0.056	0.043	0.047	5.091	0.040
	(*0.013*)	(*0.021*)	(*0.021*)	(*0.022*)	(*1.834*)	(*0.020*)
	[0.000]	[0.008]	[0.054]	[0.042]	[0.006]	[0.100]
Observations	1,257	1,257	1,257	1,257	1,257	1,257
Underclass URM students	0.073	0.060	0.053	0.049	5.592	0.042
	(*0.014*)	(*0.025*)	(*0.025*)	(*0.027*)	(*2.309*)	(*0.027*)
	[0.000]	[0.034]	[0.038]	[0.070]	[0.016]	[0.154]
Observations	935	935	935	935	935	935
Upperclass URM students	−0.015	0.020	−0.005	0.021	1.914	0.022
	(*0.025*)	(*0.030*)	(*0.039*)	(*0.041*)	(*2.320*)	(*0.046*)
	[0.574]	[0.550]	[0.892]	[0.648]	[0.464]	[0.654]
Observations	322	322	322	322	322	322

Notes: Each column reports results from a separate regression. All specifications include classroom fixed effects and individual controls race, gender, high school and pretreatment college GPA, pretreatment units earned, and year of schooling (e.g., freshman, sophomore, or junior). Standard errors in parentheses are clustered at the course by phase level. Square brackets contain p-values from randomization-based inference using a counterfactual of randomly assigning treatment status within classrooms 500 times.

Results in columns 1–4 show treatment effects on semester-by-semester persistence. For the entire sample, we find positive and significant treatment effect of 2.7 percentage points ($p = 0.002$) on persistence one-semester later. This effect slowly decays by about 50 percent and is no longer statistically significant within three semesters. Notably, the effects on persistence are larger and more robust for our target group, particularly for those students in their first two years of college. For this latter group, the treatment results in a 7.3 percentage point ($p = 0.000$) increase in persistence one semester later, with the effect persisting over the next several years.

In columns 5 and 6, we examine treatment effects on course credit accumulation and graduation through the fall semester of 2020. For both the entire sample and our target population, we find positive and significant treatment effects on total course credit accumulation. For the entire sample, students in the treatment earned 2.8 more course credits ($p = 0.020$). Similar to our previous findings, these effects are larger for the target population of URM students (5.1 credits, $p = 0.006$) and, in particular, underclass URM students (5.6 credits, $p = 0.016$). For our graduation outcome, we find a small and insignificant positive treatment effect for the entire sample. Whereas, for our target population of URM students, we find a positive treatment effect of 4.0 percentage points ($p = 0.100$), with a similar estimate of 4.2 percentage points ($p = 0.154$) for underclass URM students.[25]

Faculty and Student Race/Ethnicity Interactions.—Given the significant treatment effects for URM students, we next explore whether these heterogenous treatment effects for URM students differ by the race/ethnicity of the faculty member teaching the course. Results of this exercise are shown in Table 6, which estimates models of our various outcomes, while including an interaction between treatment status and whether the instructor is URM. Though a majority of the coefficients on the interaction are positive, we find no evidence of a statistically significant differential treatment effect for URM students when taught by URM faculty. We interpret these finding as evidence that fidelity in treatment was similar irrespective of faculty race/ethnicity. Additionally, these results provide some evidence that all professors, regardless of their race/ethnicity, may be a catalyst for improving outcomes for underrepresented students.[26]

IV. Discussion

Overall, our results suggest that a light-touch intervention that increased professor engagement significantly improved students' perceptions of the professor and course, and the course performance of underrepresented minority students in their early years of college. Moreover, we find that these positive benefits are lasting; underrepresented students in the treatment group were more likely to persist in college and graduate. To better understand why this particular intervention was effective, we further examine qualitative evidence from both the students who received the feedback as well as the faculty giving the feedback.

A. *Student Response to Feedback*

In addition to the previous analysis of survey responses from students, we asked all participating faculty to collect student replies to their emails, which we analyze qualitatively.[27] Though we recognize that students may endogenously respond to professors in a strategic manner, overall, we identify several themes in the qualitative coding of these data, which confirms that the intervention at least partially helped overcome some of the concerns raised by students in our focus groups regarding faculty and student interactions.

[25] In results not presented, we find qualitatively similar results when examining persistence two, four, and six semesters later. Additionally, when examining longer-run treatment effects separately for White and Asian students, we find null effects across all our outcomes measuring persistence, graduation and credit accumulation.

[26] We also examined differences in treatment effects differed by course discipline, class size, faculty gender, and academic rank. We find no systematic pattern of treatment effects by course discipline, class size, or instructor type, with the exception of academic rank, where we find smaller treatment effects on course grades for full professors, we find no discernable differences in the treatment effects on longer-run outcomes across academic rank.

[27] We employed an open qualitative coding scheme, using two readers to confirm themes found in the student email data.

TABLE 6—INTERACTIONS BY PROFESSOR AND STUDENT

Group Specification	Grade 1	How much do you believe the instructor **cared** about your success in the class? 2	How well did the instructor keep you **informed** about your progress in the class? 3	Persist 1-semester later (or graduate) 4	Persist 3-semesters later (or graduate) 5	Persist 5-semesters later (or graduate) 6	Total units earned as of Fall 2020 7	Graduate by Fall 2020 8
Panel A. All URM students								
Treatment	0.072	0.344	0.323	0.043	0.058	0.046	5.010	0.042
	(0.067)	(0.233)	(0.257)	(0.015)	(0.023)	(0.021)	(1.930)	(0.021)
	[0.312]	[0.012]	[0.040]	[0.000]	[0.014]	[0.072]	[0.014]	[0.112]
Treatment × URM faculty	0.223	0.435	0.248	0.040	−0.011	−0.015	0.459	−0.011
	(0.123)	(0.266)	(0.343)	(0.031)	(0.062)	(0.062)	(5.503)	(0.060)
	[0.140]	[0.332]	[0.542]	[0.296]	[0.818]	[0.796]	[0.936]	[0.854]
Panel B. Underclass URM students								
Treatment	0.099	0.416	0.407	0.066	0.063	0.057	5.210	0.043
	(0.072)	(0.293)	(0.324)	(0.015)	(0.024)	(0.023)	(2.353)	(0.029)
	[0.196]	[0.018]	[0.034]	[0.004]	[0.036]	[0.072]	[0.054]	[0.196]
Treatment × URM faculty	0.222	0.283	−0.070	0.032	−0.018	−0.018	1.919	−0.006
	(0.125)	(0.365)	(0.390)	(0.039)	(0.076)	(0.081)	(7.104)	(0.074)
	[0.232]	[0.566]	[0.902]	[0.506]	[0.798]	[0.796]	[0.746]	[0.928]
Panel C. Upperclass URM students								
Treatment	−0.027	0.271	0.248	−0.018	0.011	−0.008	2.505	0.028
	(0.138)	(0.412)	(0.328)	(0.029)	(0.035)	(0.044)	(2.605)	(0.051)
	[0.844]	[0.518]	[0.468]	[0.508]	[0.760]	[0.852]	[0.370]	[0.596]
Treatment × URM faculty	0.258	0.328	0.704	0.027	0.070	0.019	−4.636	−0.041
	(0.265)	(0.480)	(0.339)	(0.042)	(0.044)	(0.063)	(3.449)	(0.075)
	[0.552]	[0.756]	[0.294]	[0.696]	[0.404]	[0.844]	[0.502]	[0.736]
Number of URM professors in sample	5	5	5	5	5	5	5	5
Number of URM students in sample	1,197	293	294	1,257	1,257	1,257	1,257	1,257
Fraction of URM students taught by URM faculty	0.176	0.147	0.146	0.176	0.176	0.176	0.176	0.176

Notes: Each column reports results from a separate regression. All specifications include classroom fixed effects and demographic controls. Standard errors in parentheses are clustered at the course by phase level. Square brackets contain *p*-values from randomization-based inference using a counterfactual of randomly assigning treatment status within classrooms 500 times.

First, many students from the treatment group wrote emails expressing their appreciation and gratitude toward this individualized 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.
- Hello Professor, It means a great deal to receive feedback and I am appreciative of your time and help. I love what I'm 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!

A second theme that emerged in the qualitative data is that students were apologetic, often expressing regret for their actions, and 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 definitely 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 two 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 and 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 did well on the majority of the homework 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. 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.

A third dominant theme in the return emails from students is 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.

The replies from students in this instructor's course 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 percent, 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 2 PM 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.

B. *Faculty Response to the Intervention*

To assess how faculty responded to the intervention, we surveyed all faculty participants after the first phase. Specifically, we asked faculty how they interpreted the nature of the student responses to their emails. Faculty responses largely mirrored our analysis of the qualitative data; they 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).

We also asked faculty about their efforts at implementing the intervention, and beliefs about the outcome of such efforts. Faculty were asked how long the emails took them to complete; a conservative estimate is approximately one minute per email. Faculty believed that increasing interaction with students in their class could improve student outcomes; and most 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 efforts in their classrooms. Future experimental studies in higher education can thus consider targeting other types of faculty behaviors in the classroom, such as particular instructional strategies, specific course participation activities, or alternative student engagement efforts such as required office hours.[28]

We followed up with participating faculty to share the findings from the study, and to learn about whether they have continued with any of the behaviors initiated by the experiment following the end of the formal intervention period. Specifically, we asked: "*since participating in this study five years ago, have you continued to provide students with a similar set of individualized feedback messages about their performance throughout the semester?*"

A little over half of the faculty in the study (58 percent) indicated that they have continued, while 42 percent said no. When we asked why or why not, faculty who continued the practice offered a variety of reasons connected to enhancing student connection and support. For example, "I feel that it might help to provide students with individualized feedback and it could not hurt. Sometimes I get replies from students of assistance they need and am able to provide them with resources which I think makes it even more worthwhile." Those who have stopped, largely identified time constraints as the primary reason, and suggested that they still reach out to some students; "*I haven't taken the time to execute the practice across my whole class, but I have at least stuck with individualized follow up emails to struggling students.*" Finally, we asked: "*Now that you know that the faculty feedback intervention was effective at improving underrepresented minority students' course grades and longer-term outcomes such as graduation, how likely are you to implement this strategy in your courses in the future?*" On a five-point Likert scale, 75 percent indicated a "5" (highly likely), the remaining 25 percent reported a 4. Whether knowledge of the success of this intervention ultimately institutionalizes faculty behaviors toward providing more directed student feedback is an important question to consider, albeit outside the scope of this study. Nevertheless, as this faculty member's comments suggest, there is clearly potential in engaging faculty on instructional strategies in the college classroom to increase student success for historically marginalized groups; "[*I*] *didn't know how effective it was. Now since the results are positive, I am willing to try it if time permits. Thank you for letting us know the results of the study.*"

V. Conclusion

College completion and success remains highly uneven by institutional 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

[28] We also examined the fidelity of the implementation of the experiment. To increase fidelity and consistency in implementation, we assigned a research assistant to all participating faculty to assist with email drafts and ensure no contamination of the control group. Second, we examined 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, or lack thereof; and/or encouragement), and timing of the feedback (i.e., in conjunction with key course assignments or exams), all participating instructors met the three criteria required by the emails—repeated personalized feedback responsive to student performance. Thus, we can confidently rule out fidelity as a potential mechanism. As such, given the clean implementation of our experiment, results should be viewed as treatment on the treated estimates.

of the role of college faculty/instructors in student success. Results from a 2014 Gallup-Purdue study on the undergraduate college experience reveals that only 27 percent of students strongly agree with the statement "My professors cared about me as a person," (Ray and Marken 2014). To our knowledge, this study represents the first experiment aimed at altering specific faculty behaviors in the college classroom to enhance faculty-student engagement. 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 a large lecture class. As such, our research was designed with an explicit focus on what might work at scale (Banerjee et al. 2017) and generated with the population we aimed to target in mind.

The results provide experimental evidence that professor feedback to students can have a positive significant effect on all students' perceptions of support in their college classes, and on course performance and college persistence and completion for underrepresented students, a target population with an increased risk of dropping out. In addition, 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 of directly and indirectly contributing to college success: directly through the intended transfer of content knowledge and/or skills and indirectly through boosting students' sense of self-efficacy and belonging. Students' beliefs—especially those from historically marginalized groups—about college and how they process early difficulties can influence their postsecondary trajectory. Thus, 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, particularly for underrepresented minority students.

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 and, importantly, in attenuating equity gaps in college success through a modest set of activities to reach out to their students. Moreover, our study shows that a relatively low-cost intervention can have high returns. We conducted a back-of-the-envelope cost-benefit calculation and found that our intervention induced 24 additional college graduates in the treatment group at a cost of approximately $2,500 per graduate. Although measuring the total benefit of these 24 additional graduates is difficult, even under the most conservative estimates, the intervention passes the cost-benefit test.[29] 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. "Fully understanding the key mechanism behind successful interventions is often

[29] The direct cost of the intervention includes: (1) $12,500 spent on incentive payments to the faculty participants; and (2) $48,000 for the four research assistants who aided the faculty in implementing the experiment. To calculate the benefit of the experiment, using our estimating model for graduation, we predicted the probability of graduation with and without the treatment effect and show that the treatment induced 24 additional college graduates in the treatment group. Hence, the cost per graduate is approximately $2,521.

likely to take more than one experiment," (Banerjee et al. 2017, 96), and interventions with faculty in higher education contexts is no exception. Future work can and should offer additional experimentation with pedagogical approaches to feedback, alternate forms of faculty-student engagement in the college classroom, and should be cognizant of how such efforts may be received differently by different types of students (e.g., demographic background, preparation levels, etc.), by different messengers, and in different contexts (e.g., institutions, disciplines, course format, faculty incentives, etc.).

REFERENCES

Anderson, Drew M., Katharine M. Broton, Sara Goldrick-Rab, and Robert Kelchen. 2018. "Experimental Evidence on the Impacts of Need-Based Financial Aid: Longitudinal Assessment of the Wisconsin Scholars Grant." *Journal of Policy Analysis and Management* 39 (3): 720–39.

Angrist, Joshua, Daniel Lang, and Philip Oreopoulos. 2009. "Incentives and Services for College Achievement: Evidence from a Randomized Trial." *American Economic Journal: Applied Economics* 1 (1): 136–63.

Angrist, Joshua, Philip Oreopoulos, and Tyler Williams. 2014. "When Opportunity Knocks, Who Answers? New Evidence on College Achievement Awards." *Journal of Human Resources* 49 (3): 572–610.

Astin, Alexander W. 1993. What Matters in College? Four Critical Years Revisited. San Francisco, CA: Jossey-Bass.

Astin, Alexander W. 1999. Student Involvement: A Developmental Theory for Higher Education. *Journal of College Student Development* 40 (5): 518–29.

Athey, Susan, and Guido W. Imbens. 2017. "The Econometrics of Randomized Experiments Handbook of Economic Field Experiment." In *Handbook of Economic Field Experiments, Vol. 1*, edited by Abhijit Vinayak Banerjee and Esther Duflo, 73–140. Amsterdam: Elsevier.

Avery, Christopher, and Thomas J. Kane. 2004. "Student Perceptions of College Opportunities. The Boston COACH Program." In *College Choices: The Economics of Where to Go, When to Go, and How to Pay For It*, edited by Caroline M. Hoxby, 355–94. Chicago: University of Chicago Press.

Azmat, Ghazala, Manuel Bagues, Antonio Cabrales, and Nagore Iriberri. 2019. "What You Don't Know…Can't Hurt You? A Natural Field Experiment on Relative Performance Feedback in Higher Education." *Management Science* 65 (8): 3714–36.

Bailey, Martha J., and Susan M. Dynarski. 2011. "Inequality in Postsecondary Education." In *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, edited by Greg J. Duncan and Richard J. Murnane, 117–32. New York, NY: Russell Sage Foundation.

Ball, Sheryl, Catherine Eckel, and Christian Rojas. 2006. "Technology Improves Learning in Large Principles of Economics Classes: Using Our WITS." *American Economic Review* 96 (2): 442–46.

Balzer, William K., Michael E. Doherty, and Raymond O'Connor Jr. 1989. "Effects of Cognitive Feedback on Performance." *Psychological Bulletin* 106 (3): 410–33.

Bandura, Albert. 1993. "Perceived Self-Efficacy in Cognitive Development and Functioning." *Educational Psychologist* 28 (2): 117–48.

Bandura, Albert, and Dale H. Schunk. 1981. "Cultivating Competence, Self-Efficacy, and Intrinsic Interest Through Proximal Self-Motivation." *Journal of Personality and Social Psychology* 41 (3): 586–98.

Banerjee, Abhijit, Rukmini Banerji, James Berry, Esther Duflo, Harini Kannan, Shobhini Mukerji, Marc Shotland, and Michael Walton. 2017. "From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application." *Journal of Economic Perspectives* 31 (4): 73–102.

Barr, Andrew, and Benjamin Castleman. 2017. "The Bottom Line on College Counseling." Unpublished.

Baum, Sandy, Jennifer Ma, and Kathleen Payea. 2013. *Education Pays: The Benefits of Higher Education for Individuals and Society*. New York, NY: The College Board.

Beleche, Trinidad, David Fairris, and Mindy Marks. 2012. "Do Course Evaluations Truly Reflect Student Learning? Evidence from an Objectively Graded Post-Test." *Economics of Education Review* 31 (5): 709–19.

Bettinger, Eric P., and Rachel B. Baker. 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, Eric P. and Bridget Long. 2006. "The Increasing Use of Adjunct Instructors at Public Institutions: Are we Hurting Students?" In A*ssessing Public Higher Education at the Start of the 21st Century*, edited by R. Ehrenbergm.

Bettinger, Eric P., and Bridget Long. 2010. "Does Cheaper Mean Better? The Impact of Using Adjunct Instructors on Student Outcomes." *Review of Economics and Statistics* 92 (3): 598–613.

Bettinger, Eric P., Bridget Long, Philip Oreopoulos, and Lisa Sanbonmatsu. 2012. "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block Fafsa Experiment." *Quarterly Journal of Economics* 127 (3): 1205–42.

Braga, Michela, Marco Paccagnella, and Michelle Pellizzari. 2014. "Evaluating Students' Evaluations of Professors." *Economics of Education Review* 41: 71–88.

Broda, Michael, John Yun, Barbara Schneider, David Yeager, Gregory Walton, and Matthew Diemer. 2018. "Reducing Inequality in Academic Success for Incoming College Students: A Randomized Trial of Growth Mindset and Belonging Interventions." *Journal of Research on Educational Effectiveness* 11 (3): 317–38.

Brownback, Andy, and Sally Sadoff. 2019. "Improving College Instruction through Incentives." *Journal of Political Economy* 128 (8): 2925–72.

Carlson, Deven E., Felix Elwert, Nicholas Hillman, Alex Schmidt, and Barbara L. Wolfe. 2019. "The Effects of Financial Aid Grant Offers on Postsecondary Educational Outcomes: New Experimental Evidence from the Fund for Wisconsin Scholars." NBER Working Paper 26419.

Carrell, Scott E., and Michal Kurlaender. 2020a. "My Professor Cares: Experimental Evidence on the Role of Faculty Engagement: AEA RCT Registry." *American Economic Journal: Economic Policy*. https://www.socialscienceregistry.org/trials/5875.

Carrell, Scott E., and Michal Kurlaender. 2020b. "My Professor Cares: Experimental Evidence on the Role of Faculty Engagement." NBER Working Paper 27312.

Carrell, Scott E., and Michal Kurlaender. 2023. "Replication data for: My Professor Cares: Experimental Evidence on the Role of Faculty Engagement." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.38886/E169341V1.

Carrell, Scott E., Marianne E. Page, and James E. West. 2010. "Sex and Science: How Professor Gender Perpetuates the Gender Gap." *Quarterly Journal of Economics* 125 (3): 1101–44.

Carrell, Scott E., and Bruce Sacerdote. 2017. "Why Do College-Going Interventions Work?" *American Economic Journal: Applied Economics* 9 (3): 124–51.

Carrell, Scott E., and James E. West. 2010. "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors." *Journal of Political Economy* 118 (3): 409–32.

Castleman, Benjamin L., Lindsay C. Page, and Korynn Schooley. 2014. "The Forgotten Summer: Does the Offer of College Counseling After High School Mitigate Summer Melt Among College-Intending, Low-Income High School Graduates." *Policy Analysis and Management* 33 (2): 320–44.

Dobronyi, Christopher R., Philip Oreopoulos, and Uros Petronijevic. 2019. "Goal Setting, Academic Reminders, and College Success: A Large-Scale Field Experiment." *Journal of Research on Educational Effectiveness* 12 (1): 38–66.

Ehrenberg, Ronald G., and Liang Zhang. 2005. "Do Tenured and Tenure-Track Faculty Matter?" *Journal of Human Resources* 40 (3): 647–59.

Fairlie, Robert W., Florian Hoffmann, and Philip Oreopoulos. 2014. "A Community College Instructor Like Me: Race and Ethnicity Interactions in the Classroom." *American Economic Review* 104 (8): 2567–91.

Figlio, David N., Morton O. Schapiro, and Kevin B. Soter. 2015. "Are Tenure Track Professors Better Teachers?" *Review of Economics and Statistics* 97 (4): 715–24.

Hattie, John, and Helen Timperley. 2007. "The Power of Feedback." *Review of Educational Research* 77 (1): 81–112.

Hoffmann, Florian, and Philip Oreopoulos. 2009. "A Professor Like Me: The Influence of Instructor Gender on College Achievement." *Journal of Human Resources* 44 (2): 479–94.

Holzer, Harry J., and Sandy Baum. 2017. *Making College Work: Pathways to Success for Disadvantaged Students*. Washington, D.C.: Brookings Institution Press.

Hoxby, Caroline, and Christopher Avery. 2013. "The Missing "One-Offs": The Hidden Supply of High-Achieving, Low-Income Students." *Brookings Papers on Economic Activity* 44 (1): 1–65.

Hoxby, Caroline, and Sarah Turner. 2013. "Expanding College Opportunities." *Education Next* 13 (4).

Kuh, George D., and Shouping Hu. 2001. "The Effects of Student-Faculty Interaction in the 1990s." *Review of Higher Education* 24 (3): 309–32.

Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics* 112 (2): 443–77.

Leonhardt, David. 2014. "Is College Worth It? Clearly, New Data Say." New York Times, May 27. https://www.nytimes.com/2014/05/27/upshot/is-college-worth-it-clearly-new-data-say.html.

Karabenick, Stuart A., and John R. Knapp. 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 (2): 221–30.

Kluger, Avraham N., and Angelo DeNisi. 1998. "Feedback Interventions: Towards the Understanding of a Double-Edged Sword." *Current Directions in Psychological Science* 7 (3): 67–72.

List, John A., Azeem M. Shaikh, and Yang Xu. 2019. "Multiple Hypothesis Testing in Experimental Economics." *Experimental Economics* 22 (4): 773–93.

Murphy, M. C., Maithreyi Gopalan., Evelyn R. Carter, Katherine Emerson, Bette Bottoms, and Gregory Walton. 2020. "A Customized Belonging Intervention Improves Retention of Socially Disadvantaged Students at a Broad-Access University." *Science Advances* 6 (29).

National Center for Education Statistics. 2020. "Table 326.10." Digest of Education Statistics. https://nces.ed.gov/programs/digest/d20/tables/dt20_326.10.asp.

Oreopoulos, Philip, and Uros Petronijevic. 2019. "The Remarkable Unresponsiveness of College Students to Nudging and What We Can Learn from It." NBER Working Paper 26059.

Oreopoulos, Philip, Uros Petronijevic, Christine Logel, and Graham Beattie. 2020. "Improving Non-Academic Student Outcomes Using Online and Text-Message Coaching." *Journal of Economic Behavior and Organization* 171: 342–60.

Pew Research Center. 2014. *The Rising Cost of Not Going to College*. Washington DC: Pew Research Center.

Phillips, Meredith. and Reber, Sarah. 2019. "Does Virtual Advising Increase College Enrollment? Evidence from a Random Assignment College Access Field Experiment." NBER Working Paper 26509.

Pozo, Susan, and Charles A. Stull. 2006. "Requiring a Math Skills Unit: Results of a Randomized Experiment." *American Economic Review* 96 (2): 437–41.

Price, Joshua. 2010. "The Effect of Instructor Race and Gender on Student Persistence in STEM Fields." *Economics of Education Review* 29 (6): 901–10.

Ran, Florence Xiaotao, and Di Xu. 2017. "How and Why do Adjunct Instructors Affect Students' Academic Outcomes? Evidence from Two-Year and Four-Year Colleges." Center for Analysis of Postsecondary Education and Employment Working Paper.

Ray, Julie, and Stephanie Marken. 2014. *Life in College Matters for Life After College*. Washington, DC: Gallup.

Rendón, Laura I. 1994. "Validating Culturally Diverse Students: Toward a New Model of Learning and Student Development." *Innovative Higher Education* 19 (1): 33–51.

Rendón, Laura I. 2002. "Community College Puente: A Validating Model of Education. *Educational Policy* 16 (4): 642–67.

Snyder, Thomas D., and Sally A. Dillow. 2013. "Digest of Education Statistics 2012 (NCES 2014-015)." National Center for Education Statistics, Institute of Education Sciences. https://nces.ed.gov/pubs2014/2014015.pdf.

Steele, Claude M. 1988. "The Psychology of Self-Affirmation: Sustaining the Integrity of the Self." In *Advances in Experimental Social Psychology*, Vol. 21, edited by Leonard Berkowitz, 261–99. New York: Academic Press.

Stevens, Nicole M., MarYam Hamedani, and Mesmin Destin. 2014. Closing the Social-Class Achievement Gap: A Difference-Education Intervention Improves First-Generation Students' Academic Performance and All Students' College Transition." *Psychological Science* 25 (4): 943–53.

Strayhorn, Terrell. 2012. *College Students' Sense of Belonging: A Key to Educational Success for all Students*. New York: Routledge.

Thaler, Richard H., and Cass R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New London, CN: Yale University Press.

Tinto, Vincent. 1993. *Leaving College: Rethinking the Causes and Curses of Student Attrition. 2nd ed.* Chicago, IL: University of Chicago Press.

Turner, Sarah. 2018. *The Policy Imperative: Policy Tools Should Create Incentives for College Completion*. Washington, DC: American Enterprise Institute.

Umbach, Paul D., and Matthew R. Wawrzynski. 2005. "Faculty do Matter: The Role of College Faculty in Student Learning and Engagement." *Research in Higher Education* 46 (2): 153–84.

Walton, Gregory M., and Shannon Brady. 2017. "The Many Questions of Belonging." In *Handbook of Competence and Motivation 2nd Edition*, edited by Andrew Elliot, Carol Dweck, and David Yeager, 272–93). New York: Guilford Press.

Walton, Gregory M., and Geoffrey L. Cohen. 2011. "A Brief Social-Belonging Intervention Improves Academic and Health Outcomes of Minority Students." *Science* 331 (6023): 1447 –51.

Yeager, David S., and Gregory M. Walton. 2011. "Social-Psychological Interventions in Education: They're Not Magic." *Review of Educational Research* 81 (2): 267–301.

Yeager, David S., Gregory M. Walton, Shannon T. Brady, Ezgi N. Akcinar, David Paunesku, Laura Keane, Donald Kamentz, et al. 2016. "Teaching a Lay Theory before College Narrows Achievement Gaps at Scale." *Proceedings of the National Academy of Sciences* 113 (24): E3341–E48.