The Remarkable Unresponsiveness of College Students to Nudging And What We Can Learn from It

Philip Oreopoulos
University of Toronto

Uros Petronijevic
York University

December 2020

Abstract: We present results from a five-year effort to design promising virtual coaching interventions to improve college student achievement. Across nearly 20,000 students at three campuses, we find some improvement on mental health and study time but no effect on academic outcomes. We interpret the results with unique survey data and a model of student effort. Treated students learn more effort is needed to attain good grades and develop stronger preferences for high grades, but these effects are too small to translate into academic benefits. More comprehensive, social, and better-timed interventions are needed for helping students outside the classroom.

Acknowledgements: This project would not have been possible without the enthusiastic support of the Department of Economics at the University of Toronto, especially the instructors of first-year economics courses, Lee Bailey, Tenzin Yindok, Bob Gazzale, Bill Wolfson, Kieran Furlong, Avi Cohen, Jack Parkinson, Michael Ho, Zachary Mahone, Gordon Cleveland, Paola Salardi, Loren Brandt, Eric Kam, Erica Rzepecki, and Peter Foltin. We are also indebted to Aaron DeMello and James Lahey for their tireless commitment to designing the project's well-run website and online exercises, as well as for their help with organizing and extracting the experimental data. Matthew Hendrickson helped extract the administrative data. We are very grateful to all the coaches who provided experienced assistance to participants, and to the Social Science Research Council of Canada (Insight Grant #435-2018-0268 and Insight Development Grant # 430-2017-00779), the Jamal Poverty Action Lab, and the University of Toronto's Learning and Education Advancement Fund for financial support. The paper was completed while Oreopoulos was visiting the Russell Sage Foundation as a Visiting Scholar, almost entirely without procrastination. We are grateful for seminar participant feedback from numerous institutions. All remaining errors are our own. The experiments in this paper were registered with the AEA RCT Registry (AEARCTR-0000810).

I. Introduction

In nations like the United States and Canada, higher education continues to be promoted as a key tool for improving skills and wages (Deming, 2019; Acemoglu and Autor, 2012; Gurria, 2009; Psacharopoulos and Patrinos, 2018). Although individuals with more education realize better average outcomes than those with less, simply enrolling in college does not guarantee that students will be better off for two key reasons. First, a substantial fraction of current enrollees fails to graduate. In the United States, the six-year completion rate among students beginning a four-year postsecondary program is only 54.8 percent (Shapiro et al., 2019). Students who fail to complete college incur large up-front costs but can expect to earn similar incomes as individuals with only a high school degree, especially among enrollees in the bottom of their entry class distribution (Oreopoulos and Petronijevic, 2013). Second, many students who do earn a college degree do so with weak grades and questionable human capital gains. Arum and Roksa's (2011) seminal research, for example, finds little evidence of improved skills and learning among many attending college, as measured by tests designed to reflect critical thinking, complex reasoning, and writing.

A leading explanation for low learning gains is that students invest little time into their studies. Most college students spend fewer than 15 hours a week preparing outside of lecture for all of their courses, much less than the 25 to 40 hours per week usually recommended by university administrators (Brint and Cantwell, 2010; Babcock and Marks, 2011; Farkas et al., 2014). Working for pay or commuting long distances are not binding constraints for many of these students; rather, time-use surveys reveal that many students spend their time socializing or taking part in recreational activities instead of studying (Arum and Roksa, 2014, Oreopoulos et al., forthcoming).

While estimates of the *average* return to college remain significantly positive, including for students at the margin of admission (Oreopoulos and Petronijevic, 2013; Zimmerman, 2014; Ost et al., 2018), heterogeneous returns depend crucially on the role that students are willing and able to take in the development of their own human capital.[1] Of particular importance are students' cognitive and non-cognitive skills when entering college, the information they have about how to study effectively, and their willingness to devote time to studying at the expense of other activities (e.g., Nyblom, 2015). The key question motivating this paper is whether low-touch interventions can affect these kinds of inputs and, in turn, cause improvement in academic outcomes and overall college experiences.

We explore this question with a five-year sample of nearly 20,000 representative college students across all three campuses at the University of Toronto (UofT). Teaming up with instructors of first-year economics courses—who collectively teach about 5,000 students per year, including a quarter of all first-year students—we created the *Student Achievement Lab*, in which students needed to complete a one- to two-hour online 'warm-up' exercise within the first two weeks of the fall semester for a small grade requirement. Students registered an account, took a short introductory survey, and were then randomly assigned to treatment or control groups. We then linked our survey data to the university's administrative records to track academic outcomes and, in some cases, conducted follow-up surveys to collect non-academic outcomes such as study habits, aspirations, mental health, and perceptions of overall university experience.

The *Student Achievement Lab's* interventions are motivated by the notion that behavioral or psychological barriers may prevent students from realizing their preferred long-run outcomes

[1] Other determinants include financial cost (Lochner and Monge-Naranjo, 2012), incoming ability (Beattie et al., 2018), college and teaching quality (Chetty et al., 2017; Hoxby and Stange, 2018), and field of study (Kirkeboen et al. 2016).

(Lavecchia et al. 2016). Putting off studying for 'later', forgetting to take advantage of free tutoring services, or consistently getting distracted by social media are examples of how students' best intentions can go awry. Youth are particularly prone to these kinds of present or inattention biases (Giedd et al., 2012). Some nudges have proven helpful in getting students to complete a one-time action or a series of well-defined steps, such as completing a college application (Bettinger et al. 2012; Oreopoulos and Ford, 2019; Castleman et al. 2016), renewing financial aid (Castleman and Page, 2016), choosing selective colleges (Dynarski et al., 2018; Castleman and Sullivan, 2019), and choosing courses on time (Castleman and Page 2015). In contrast, nudging students toward improving study habits and attitudes has proven more challenging because it requires a sustained change in behavior over a prolonged period.

Prior studies find that offering structured, intensive, and personalized support can help. One of the most successful experimentally-tested programs is the Accelerated Study in Associate Program (ASAP), which requires that college students enroll full-time, attend mandatory tutoring, receive regular counseling and career advising services, and awards students free public transportation passes and funding for textbooks. ASAP doubled graduation rates at the City University of New York and had similarly large impacts on persistence in a replication attempt in Ohio (Scrivener et al., 2015; Sommo et al. 2018). Stay the Course (STC) and the Carolina Covenant aid program, two other comprehensive college-based support programs in Texas and North Carolina, respectively, also increased completion rates and credit accumulation (Evans et al., 2017; Clotfelter et al. 2018). While encouraging, these programs cost thousands of dollars per student and are difficult to scale. We also know little about how they improve academic outcomes, and why they do not help even more students, as one might expect given their intensity.

We use the experimental setting in the *Student Achievement Lab* to explore whether offering 'lighter-touch' and less costly interventions might also benefit students, and to learn more about the mechanisms by which students can be assisted during college. Over a span of five years, we designed and tested several promising interventions based on past research and consultations with college administrators. Most of the interventions were 'coaching' interventions, designed to better inform, motivate, advise, and remind students about effective study strategies for improving academic achievement and student experience. We group these interventions into four categories:[2] (1) *Online Coaching*, in which students were provided detailed advice about how to be a successful student; (2) online coaching with intensive follow-up communication through *One-Way Text Messages*; (3) online coaching with follow-up *Two-Way Text Messages* between students and experienced upper-year student coaches; and (4) online coaching with follow-up *in-person* regular meetings with coaches. Across all five years and interventions, our total sample consists of approximately 20,000 students.

We find that none of the interventions we test generate a significant improvement in student grades or persistence. We can rule out treatment effects larger than 8 percent of a standard deviation and find precise null impacts even when focusing on students more at risk of performing poorly and those attending the two satellite campuses that are more representative of less-selective commuter colleges.

These results, however, belie impacts on more intermediate and subjective outcomes. We find that our interventions improve study habits, such as weekly study hours and the likelihood of

[2] The *Student Achievement Lab* has been active for six years, from 2014-15 to 2019-20. In this paper, we only discuss the experiments and data relevant to the coaching interventions we evaluated in lab, which range from the 2015-16 to 2019-20 academic years. In addition to these coaching interventions, we have used the lab to test promising goal setting and mindset interventions from social psychology, which are the subjects of separate standalone papers.

meeting with a tutor or instructor. Study time increases, on average, by approximately two hours per week, but the estimated association between studying and grades (and causal estimates of these relationships from the literature) suggests that these improvements are not large enough to generate a significant change in aggregate academic outcomes. The interventions also improve subjective well-being, reduce stress, and make students feel more supported. Such impacts may be important independently of academic achievement, given the increase in attention by administrators to student experience and mental health.

While our previous studies (cited below) present findings separately for some of these interventions (and include additional estimated effects from more subtle treatment variations and sub-analyses), presenting our five-year effort collectively in this paper facilitates making more important contributions. First, our combined results emphasize the difficulty in utilizing low-cost efforts to change habits and influence student achievement, at least at the college level. Many researchers—us included—have been enticed by the prospect of applying behavioral economics to education for creating new cost-effective approaches for improving student outcomes. However, our collective results—given their consistency, fidelity, and precision—provide a clear demonstration of the limitations in using low-cost behavioral interventions to generate long-term benefits to students. Our paper underscores an emerging theme in the literature that, although nudging time-sensitive and specific actions may be feasible, a healthy dose of caution is warranted when seeking to use low-cost interventions to generate sustained long-term student gains (Oreopoulos, 2020; Page, Lee, and Gelbach, 2020).[3]

Our second contribution is to use newly gathered data from *weekly* surveys along with a model of student effort setting to better understand why our interventions generated some

[3] Recent failed replication attempts have also called into question the efficacy of interventions that previously showed promise in nudging relatively simple, time-sensitive actions (Bird et al, 2019; Gurantz et al, 2019).

intermediate effects on study behavior while not impacting academic achievement. Our model allows students to choose study intensity based on their preferences, abilities, expected effort-to-grade relationships, and psychological barriers that lead actual effort to differ from target effort. We then use our weekly data to measure changes in these factors over time and the role our interventions played in affecting each channel. We find clear evidence that students study four to five fewer hours per week than they intend and that our interventions had no impact on these gaps. Coaching did lower students' expectations about the efficacy of cramming for tests and increased their motivation to attain higher grades, although the magnitude of these effects is relatively small.

We also find that students adjust study effort and grade expectations in response to learning new information in an asymmetric way. Upon learning that it is easier to reach performance goals than originally believed, students decrease study time and leave grade expectations relatively constant. Students who learn it is harder to do well, however, do change study time and instead significantly reduce grade expectations, essentially coming to see poor performance as inevitable. We find that many of these students come to expect less of themselves shortly after midterm season in their first college semester—approximately only their *seventh* week in college.

These results shed new light on the nature and timing of college student decision-making, with important implications for the design of future interventions. While much of the existing literature explores how college students' beliefs about their academic abilities affect major (Arcidiacono, 2004; Stinebrickner and Stinebrickner, 2014a; Wiswall and Zafar, 2015) or dropout decisions (Stinebrickner and Stinebrickner, 2012, 2014b) after one or several semesters in college, we show that students also respond to new information about their abilities by adjusting the effort they invest in the accumulation of their human capital and their performance expectations. These revisions occur quite early into students' college careers—a little over halfway through their first college semester—suggesting that the time when students' learn about their performance on their first major evaluations in college may represent a critical period for targeted and personalized intervention (Carrel and Kurlaender, 2020).

As a final contribution, we also expand the list of outcomes on which we estimate treatment effects. By tracking students from our earlier interventions through their college careers, we estimate the medium-run effects of our prior interventions on students' academic outcomes (in contrast to our previous studies, which only report effects on contemporaneous outcomes). Collecting richer data in our most recent (unpublished) interventions also allows us to estimate the effects new measures of student well-being and detailed measures of study behavior. The lack of any medium-run effects on academic outcomes we report here further underscores the great challenge in using low-cost behavioral interventions to generate sustained benefits for students. Our findings suggest that more intensive and perhaps better-timed interventions are needed to meaningfully change students' performance and behavior.

The remainder of the paper proceeds as follows: in Section II, we describe our *Student Achievement Lab* setup and review each field experiment. We also describe our data and methodology, and present descriptive statistics. Section III presents the overall results. In Section IV, we present a model of student effort to interpret our results and discuss rational and time-inconsistent explanations of poor performance, as well as opportunities for policies to help. We corroborate our model and quantify the impact our experiments had on individuals' study-to-grade expectations, preferences, and procrastination in Section V and offer concluding remarks in Section VI.

II. The Student Achievement Lab: Setup, Interventions, and Data

A. Setup

The *Student Achievement Lab* (SAL) began in the fall of the 2014-15 academic year at the University of Toronto (UofT). In that first year, we conducted experiments only at the university's west-end satellite campus located in the city of Mississauga (we refer to this campus as UTM). UTM is primarily a commuter campus with approximately 12,500 undergraduate students. Roughly 80 percent of students live at home with their parent(s), slightly less than a quarter say that the campus was their first choice, and the majority say they work at least part-time while attending. Many of the students are immigrants or children of immigrants. Among undergraduates who entered in 2001, only 38 percent completed a degree in four years, while the six-year graduation rate was about 70 percent. SAL expanded in the following year (fall of 2015) to include UofT's two other campuses. The campus located in the east end of Toronto, the University of Toronto at Scarborough (UTSC), is similar to UTM, as it is primarily a commuter campus with completion rates of about 73 percent. UofT's St. George campus, UTSG, is located downtown and is more representative of a top four-year public college in the United States.[4] Students apply to each campus separately. Not surprisingly, UTSG is more selective and six-year completion rates are higher, at about 77 percent.[5]

We evaluated several promising interventions in SAL from 2014-15 to 2019-2020. In this paper, we discuss the coaching interventions we evaluated between 2015-16 and 2018-19, and

[4] The St. George campus is ranked as one of the top universities in the world: https://cwur.org/2018-19.php.
[5] The St. George Arts & Science program is about twice as big as UTM and UTSC. In 2016-17, the full-time headcount at St. George, UTM, and UTSC was 25,056, 12,967, and 11,902 respectively (University of Toronto, 2018).

later present supplemental weekly survey data collected during the 2019-20 academic year. During the fall semesters between 2014 and 2019, instructors of first-year economics courses at UofT incorporated into their course curriculum a small participation grade (usually 2 percent) for the completion of an online warm-up exercise lasting, on average, about an hour, with a deadline generally within the first two weeks of class. The grade requirement was highly effective in making almost all students participate (95 percent of all registered students at the start of the course, which at approximately 5,000 students per year constitutes 10 percent of the entire undergraduate student population).[6]

Students taking introductory economics courses are representative of the school's undergraduate student body. About a quarter of all first-year students at UofT enroll in a first-year economics course, half of which take the course as a requirement for their planned program of study. Students wanting to continue afterwards into one of the schools' competitive commerce or management programs must obtain a minimum grade (usually 67 percent) as part of that program's admissions requirements. Each year, about 30 percent of students drop their economics course before receiving an official grade. Of those who do complete, the 25th, 50th, and 75th percentiles in economics grades distribution are 58, 69.5, and 78 percent, respectively (using our baseline sample). Figures 1 and 2 depict students' academic performance overall. Figure 1 displays the distribution of grades averaged across courses completed by the end of the first fall semester. The distribution is similar to that for economics alone, with the median grade being 70.5 percent and the 25th percentile being 62.0 percent. Figure 2 shows the histogram of credits completed at the end of the first school year for our sample. Many students initially enroll in five credits to try to

[6] We restrict our sample to full-time students, defined as those paying full-time tuition, which permits them to enroll in at least 3.5 course credits over the school year.

complete their program in four years,[7] but by the end of the year, many drop some credits or fail to complete their courses. Only 30 percent of our sample received 5 or more credits by the end of the school year.

Students enrolled in first-year economics classes participated in SAL by logging in using their personal UofT account, or creating and verifying a new account, proceeding by first taking a short initial survey to collect data not available administratively (such as parents' education, grade and study expectations, education aspirations, and subjective tendencies to cram for exams), and then being randomized into different groups, which we categorize and describe below. During the last three years of the experiments, at the end of the fall semester or at the beginning of the winter (i.e., next) semester, we conducted a short follow-up survey also for a participation grade (usually worth 1 percent of the course's final grade for completion). We gathered information not available in administrative data, including questions about study habits, perceived learning outcomes, subjective well-being, attitudes towards grades, challenges with procrastination, and open-ended questions about first semester experiences, advice to other students, and feedback from treated students about the interventions.

B. Interventions[8]

1. Personality Test (Control Group)

[7] Students require 20 completed credits to earn a degree.

[8] All surveys and interventions in their original form are available to peruse online at https://studentachievementlab.org. For additional operational details not all covered in this paper, readers may also refer to appendices provided in Beattie et al. (2018) for the Personality Test and Oreopoulos and Petronijevic (2018), Oreopoulos et al. (2020), and Oreopoulos et al. (forthcoming) for many of the coaching interventions and follow-up surveys.

Students assigned to the control group were given a set of questions about time preferences, non-cognitive abilities, and interests. In order to make the exercise last as long as the treatment interventions, Control Group students were given two Big Five[9] personality tests: one based on an absolute score (e.g., Donnellan et al., 2006), making it possible for a student to score high in all five traits, and another based on a relative score (e.g., Hirsh and Peterson, 2008), indicating the extent to which one trait dominates a student's personality profile relative to other traits. The control group was also asked questions about risk tolerance (e.g., Dohmen et al., 2011), time preferences (e.g., Andersen et al., 2008), and grit (e.g., Duckworth and Quinn, 2009). The test took approximately 45 to 60 minutes to complete. Students were emailed a short report describing their relative Big Five scores and told that they might be interested in knowing which of their traits are most and least dominant.[10]

2. Online Coaching Only

Our interventions offer direct coaching advice about how to perform well in university and have a successful experience. Several of the beneficial comprehensive college support programs mentioned in the introduction offer coaching and mandatory workshops about studying and

[9] The five traits are agreeableness, conscientiousness, extraversion, openness to experience, and emotional stability.
[10] The Personality Test was not intended to affect subsequent academic performance or behavior, but data from respondents was used to explore which background and non-cognitive trait variables best predict the wide variance in first-year college performance. Beattie et al (2018) find that students who perform far below expectations also self-report greater tendency to procrastinate and being less conscientious ('gritty') than their peers. Those who perform unexpectedly and exceptionally well express purpose-driven goals and an intent to study more hours per week to obtain a high GPA. In a separate paper that uses follow-up survey data from SAL, Beattie et al (2019) examine the association between intermediate study inputs during a college semester and find that poor time management and lack of study hours are most associated with poor academic performance while large amounts of study time and regular use of student services are most associated with good academic performance. Worth noting as a prelude to this paper's discussion, both of these papers find that a student's high school grade used for admission is, by far, the most predictive variable for first-year performance, and that the additional non-cognitive variables examined do not improve predicted performance by much. A large variance remains even after accounting for observed differences in student background and study behavior.

performing well. Programs tested by Evans et al. (2018) and Bettinger and Baker (2014), in particular, have coaching as the main or key component. The personalized, ongoing, and proactive nature of these services, which we examine more below, may be important, but we intended to test as a baseline whether an inexpensive, one-time online exercise providing similar advice to what a coach would offer could generate even a small impact on academic achievement.

We tested two online-only coaching programs at SAL. The first, implemented in the 2015 fall semester and evaluated and discussed in length in Oreopoulos and Petronijevic (2018), asked students to think about the future they envision and the steps they could take in the upcoming year at UofT to help make that future a reality. The online module lasted approximately 60 to 90 minutes and led students through a series of writing tasks about their ideal futures, both at work and at home, what they would like to accomplish in the current school year, how they intend to follow certain study strategies to meet their goals, and whether they wanted to get involved with extracurricular activities at the university. The exercise aimed to make students' distant goals salient in the present and to provide information on effective study strategies and how to deal with the inevitable setbacks that arise during an academic year.

Together with Christine Logel, a social psychologist from the University of Waterloo, we designed a second online-only coaching treatment the following school year, 2016-17. This intervention is the focus of Oreopoulos, Petronijevic, Logel, and Beattie (2020), and it incorporates elements of goal setting and mindset interventions from psychology and our previous coaching treatment. Split into two parts, the intervention allows each student to focus on the challenges they think are particularly important. Part One presents students with six broad factors critical to academic success,[11] with subsequent sections elaborating on each factor and taking students through tasks that draw on psychology research on attitude and behavior change. Part Two presents students with eight institutional barriers to success and invites them to choose the two barriers most relevant to future students like them, identify and write about a reason why students might struggle with this problem, and identify and write about a potential solution.[12]

Both exercises mentioned in this section offer detailed and specific online coaching advice for performing well in university and having a good experience. We therefore group them both into the same category, which we call 'Online Coaching Only'. Table A5 in Appendix A includes baseline results with separate treatment effects (which are similar).

3. Online Plus One-Way Text Coaching

To help students stay motivated and remember study advice, a random subset of students finishing the online coaching exercise were also offered the opportunity to receive follow-up communication during the school year by text message or email. Students were told that they were selected by lottery to participate in a pilot project designed to help with their goals and provide extra support outside the classroom. About 85 percent of those invited provided a cell phone number. The remainder received emails with similar content. The initiative was called *You@UofT*—a name we chose to associate the program directly with the university and its effort to support students' individual goals.

[11] These include studying enough, studying effectively, seeking help, attending class, staying motivated, being patient and taking a long-term perspective.
[12] Figures A1 and A2 in Appendix A include sample screenshots of the initial instructions and video, and one of Part One's modules about the importance of staying motivated while studying.

In one treatment arm during the 2015-16 school year, students received mostly one-way messages, designed deliberately not to solicit a response—a design feature that allowed us to avoid having to hire, train, employ, and manage real coaches, making the marginal cost of the program almost zero (sending one text message costs about US$0.0075).[13] Text messages were typically three to four lines in length, while emails were longer and provided more detail (students received both text and email messages). The messages typically focused on three themes: academic and study preparation advice, information on the resources available at the university, and motivation and encouragement. Figure A3 in Appendix A shows a screen shot of the coaching manager we used to view outgoing (and incoming) messages for the one-way text message coaching treatment. Messages were signed from the 'You@UofT Support Team' rather than any individual person. Students were free to opt out of receiving email messages, text messages, or both at any time after the exercise, although few chose to do so.

In the 2016-17 school year at the UTM campus, we partnered with an existing for-profit company in the business of sending one-way text messages to college students to improve academic achievement and persistence.[14] With this treatment arm, we aimed to investigate whether experienced commercial organizations can design more effective text message coaching programs. Randomly selected students still completed the online coaching exercise, but then were offered the outsourced one-way text message coaching (they did not know that the messages sent to them were from the outside organization). Students who did not provide cell phone numbers received our regular email messages instead. The text message program remained labelled

[13] This intervention is described in more detail and evaluated in Oreopoulos and Petronijevic (2018).
[14] A more detailed description and evaluation of this intervention appears in Oreopoulos, Petronijevic, Logel, Beattie (2020).

You@UofT and references were made to UTM student services. Some messages invited students to text back yes/no or numbered responses to receive automated replies.

Again, for the purpose of reporting results from a generalizable set of interventions from SAL, we estimate our main treatment effects below by grouping together the one-way-text coaching program administered by us and the one administered by the for-profit company. Similar but less precise results are presented separately in Table A5 in Appendix A.

4. Online Plus Two-Way Text Coaching

We investigated the impact of more intensive and personalized coaching by introducing two-way text message coaching, in which students were assigned to experienced upper-year undergraduate coaches whom they could message back with questions or simply check in about how their week was progressing and whether any challenges had arisen. Coaches were recruited based on their academic transcript and existing experience with mentoring, tutoring, and coaching students through other student services. They also received training from the university's Academic Skills Centre and from a one-day workshop discussing the You@UofT program and how to best communicate with students via text.

In the 2016-17 school year, a random subset of first-year economics students who completed the online coaching exercise were also offered an individual coach who would send them messages throughout the year and with whom they could communicate back.[15]

[15] More information about this intervention and the recruitment and qualifications of the coaches is provided is in Oreopoulos, Petronijevic, Logel, Beattie (2020).

Approximately 90 percent of students chose to opt in by providing their phone numbers, and less than 3 percent later chose to opt out. Those who did not provide a number received weekly email messages of study advice and motivation instead.

Coaches initiated communication with each of their students at least once a week (often twice a week), typically using pre-programed (but auto-populated with student-specific information) batch messages designed to stimulate conversation. Once contact was established, conversations evolved organically, with coaches usually trying to determine how students were progressing, both academically and emotionally. We encouraged coaches to follow up with individual students on recently discussed issues. Figure A4 in Appendix A provides a sample conversation using two-way coaching and our platform during the 2016-17 year.

We designed an online and coaching intervention the following year to emphasize the importance of sufficient study time. Many college administrators and faculty recommend two or three hours of study each week for each hour a student spends in class, implying at least 25 to 35 hours of effort outside of class for someone enrolled full-time. In contrast, many of our participants in SAL report studying fewer than 15 hours per week for all their courses, with more than a quarter of our sample studying fewer than 10 hours per week. Poorly performing students who study only a few hours per week are unlikely to benefit from any intervention that does not increase engagement outside the classroom.

SAL participants assigned to coaching at UTSG and UTM in 2017-18 and at all campuses in 2018-19 were guided through a planning exercise.[16] We first told students about UofT's

[16] The intervention in the 2017-18 academic year is described and evaluated in full in Oreopoulos, Patterson, Petronijevic, Pope (forthcoming). The 2018-19 treatment has not been evaluated in prior work, apart from a limited evaluation of treatment effects on student study time (not academic outcomes) serving as supplemental evidence in Oreopoulos, Patterson, Petronijevic, Pope (forthcoming).

recommendation for weekly study time, showed them several stories from past students about the importance of sufficient study time, and invited them to write about how they could motivate themselves to stay committed to a regular study routine. We then guided students through making a regular study schedule by building a weekly calendar. Figure A5, in Appendix A, displays a screen shot of this part. Most students could upload their weekly schedules to their electronic phone or computer calendars.[17] We invited all treated students to receive follow-up communication with a virtual coach, who would send them a study tip and check in with them each week about their weekly study progress.[18] As with the earlier coaching interventions, the minority of students who did not provide a cell phone number message received similar email messages instead.

The virtual coaching programs were well received. Figure 3 charts text-back response rates from students who provided cell phone numbers. Combining samples over the three years that two-way text coaching was offered, we see that more than 65 percent replied at least once to their coach during the first semester. Weekly response rates were relatively high, especially during the first month, when a third to a half of eligible students replied every week. Students not responding still may have benefited from the advice and reminders we sent. As a quality check, we contacted some students who were not responding to any text messages. They mentioned that they felt too busy to reply but wanted to keep receiving the messages because they found them helpful. Figure 4 reinforces this conclusion, indicating feedback about the text-message coaching program from our follow-up surveys. Most students enjoyed the program and felt that they were doing better in

[17] For the 2018-19 school year, students treated with this planning intervention could also indicate deadlines for particular tests, exams, and writing assignments. Based on these deadlines, we uploaded to their calendars suggested study strategies prior to these deadlines to prepare.

[18] Students who provided their cell phone numbers were assigned to a specific coach, and each coach was assigned a few time slots during the week to be the coach who was ‘on call’. During each on-call time for a given coach, we sent a batch message to all students who were assigned to that coach to spur productive conversation. If students replied while their coach was still on call, that coach would continue the conversation. If students replied after their coach’s shift ended, the coach who was currently on call or the team manager was responsible for closing the conversation.

university at least in part because of their coach. Seventy percent of respondents preferred that the coaching program continue into the following semester (should resources be available), and 87 percent said that it should be offered to the cohort of students next year. We also received several personal text and email messages at the end of the program expressing gratitude and appreciation from having participated.

5. Online and Face-to-Face Coaching

To compare the lower-cost, lower-touch interventions above with more intensive coaching efforts for helping students, we randomly offered a small sample of students during the 2015-16 and 2016-17 school years a coach to meet in person rather than communicating only through text. The interventions were provided only at the UTM campus. Oreopoulos & Petronijevic (2018) provides a detailed description of the face-to-face coaching program. Briefly, after completing the online coaching exercise, 24 randomly selected students were offered one of four personal coaches in 2015-16 and 66 students were offered one of nine personal coaches the following year. Coaches arranged weekly 30-60 minutes meetings with their assigned students, and to reached out to reschedule when meetings did not occur. Coaches were also available in between meetings via Skype, email, or text. Students were sent messages of advice and motivation from their coaches, much like the other coaching programs described above.

C. Data and Methodology

Our baseline sample for estimating the effects of interventions described above includes all full-time UofT students between the 2015-16 and 2018-19 school years taking a fall semester introductory economics course and who at least started a SAL online warm-up exercise before October 1.[19] In Section V below, we incorporate detailed weekly survey data from the 2019-20 year, which allow us to investigate treatment mechanisms and student decision making. Students received a grade worth 1 to 2 percent for completion of the exercise, and about 95 percent of all initially registered students completed the exercise within the first few weeks of September.

Table 1 shows the number of students assigned to each intervention within each campus-year cluster. The table also compares the actual percentage of students assigned to each intervention relative to the percentage we should have expected had the assignment process been truly random. The second column of Table 1 indicates, for example, that in the 2015-16 academic year, 19.6 percent of first year students participating in SAL were assigned to the online coaching only, 30.7 percent were assigned to online and one-way text coaching, 0.4 percent to face-to-face coaching, and the remaining 29.6 percent to the control group.[20] These proportions are all very close to those we expected to obtain according to the randomization rule. The total sample size over all five years was 19,864 students.

We estimate treatment effects by regressing outcome variables on treatment dummy variables plus fixed effects for each of the eight campus-year clusters listed in Table 1. The fixed

[19] Full-time students are those registered to take at least 3.5 course credits over the school year. Typically, 5.0 credits each year are needed to complete a program in four years.

[20] Statistics in 2015 are reported separately for first-year and non-first-year students because we also evaluated a mindset intervention (inspired by the social psychology literature) in that year, and only first-year students were randomly assigned to that intervention. As mentioned above, we do not explore these mindset interventions in this paper. Because approximately 20 percent of first-year students were assigned to the mindset condition, the percentages in column (2) do not sum to 100. We also conducted mindset interventions at UTSG in 2017, randomly assigning approximately 36 percent of students to those interventions, which is again why the percentages in column (7) do not sum to 100.

effects are necessary because the interventions we designed changed over time, as did the sample populations. In addition, it was sometimes the case that some demographic groups of students (e.g., international or first-year students) or students at particular campuses were disproportionately assigned to certain interventions relative to other groups of students or students attending other campuses. We therefore include the cluster fixed effects to account for the mechanical correlation between treatment status and cohort, campus location, or background variables introduced by our assignment rules. The treatment effects may be interpreted as average outcome differences between those from treatment groups and those from the control group within a given campus-year cluster. We do not condition these regressions on any additional background variables for ease of interpretation and because of missing high school admissions grades for some students (Table A1 in Appendix A shows that baseline results do not change when we do).

In addition to data collected through the warm-up exercise itself and follow-up surveys, we linked students to administrative admissions records and academic performance (e.g., credit accumulation, GPA) data. Column 2 in Table 2 displays descriptive mean characteristics for the control group from our full sample. Column 3 indicates the corresponding standard deviation. A few observations from the table are particularly noteworthy. Most students self-report aspiring to pursue graduate studies after completing their undergraduate degree (66 percent). This widespread ambition suggests that good grade performance should matter to many. Indeed, the expected fall grade average is 80.7 percent.[21] Thirty-two percent of the students are international students, implying they pay larger tuition fees and have not lived in Canada until very recently. An even larger fraction does not speak English at home. Most students are admitted with very high grades—the average admissions grade (typically the top 6 high school courses) is 85.4 percent.

[21] Percentage grades that are 80 percent or above correspond to an A- or higher at UofT.

Almost a third of students are first-generation (with both parents having less than a university education). Students expect to study in the fall semester an average of 18 hours a week, with a large standard deviation of 12 hours.

Columns 4, 6, 8, and 10 show estimated differences in mean characteristics between treatment groups and the control group (along with respective standard errors listed in columns 5, 7, 9, and 11). Estimates include fixed effects for the sample clusters mentioned above and listed in Table 1. Out of the 72 estimates, two are significant at the 1 percent level, three at the 5 percent level, and eight at the 10 percent level, close to what would be expected by chance. Even these statistically significant differences are generally small, due to the large sample sizes. Together with Table 1 that shows each of the intervention groups appear to be in proportional size to what would be expected from random assignment, we take these results to suggest students were credibly randomly assigned in each experiment.

UofT administrative data allow us to track academic performance within the university until the start of the 2019 term. Table 3 shows means of the outcome variables for the control group that we can measure depending on the school year the experiment began. Credits earned and course grades for the fall semester during which the experiment began are observed for each of the four cohorts. Many students do not complete the credits needed each year to graduate from their programs on time. For our earliest experiment conducted at the beginning of the 2015-16 school year, we observe that only 46 percent of the first-year students in the control group are recorded as graduating by the end of their fourth year. About 10 percent of first-year students fail to persist into second year. By fourth year, 20 to 25 percent are no longer registered with the university.

III. Results

A. First Term Academic Performance

We first present the effects of the four interventions on academic outcomes during the fall semester in which each experiment began. Outcomes are regressed on intervention indicators plus fixed effects for the cluster groups listed in Table 1.[22] The second column of Table 4 shows estimated effects on missing grade data at the end of the first fall semester. Not having any grade data may indicate that a student dropped out of the program entirely or that they enrolled in only full-year courses and grades are not yet available.[23] The findings suggest that the interventions generally had no impact on the likelihood of missing grade data compared to the fraction of students missing data in the control group (13 percent). The exception is for students receiving Online Plus One-Way Text Coaching, with the estimated impact being positive, implying a counterintuitive increase in the likelihood of missing recorded grades. We discount the importance of this result given the fact that we do not find effects for the other, more intensive interventions, and that we do not find an impact for the same treatment on credits earned over the entire school year (results shown in Table 5).

Column 3 in Table 4 shows estimated treatment effects on non-missing average fall semester grades, measured in percentage points. The control group's mean average grade in the

[22] The same table in which the regressions also condition on background variables (showing similar results) is shown as Table A1 in Appendix A.
[23] Most courses at UofT are one-semester courses. Even courses that tend to last a school year, like economics, have been split across two semesters (e.g., micro and macro) to make course selection easier. An exception is at the UTM campus, where several large first-year courses remain defined as full-year.

fall semester is 69.2 percent with a wide standard deviation of 13.4 percentage points. None of the four estimated effects are statistically or economically significant, with the largest effect being only 1.9 percent of a standard deviation. Columns 4 through 8 in Table 4 display null distributional treatment effects for each treatment as well—that is, none of the interventions affect the likelihood of receiving a fall grade average greater than 50, 60, 70, 80, or 90 percent respectively. Appendix Tables A2, A3, and A4 show similarly small and insignificant effects on average course grades across each campus separately, on first-semester average math grades, and on first-semester average economics grades, respectively. We therefore conclude that none of the coaching interventions impacted fall semester grades.

B. Persistent Academic Outcomes

Table 5 shows treatment effect estimates beyond the first term, with column 2 again showing the null effects on fall semester grades from Table 4 as a reference point. Impacts on winter semester grades are also mainly null, except for the online-only intervention. Column 4 indicates that, on average, students from the control group earn 3.6 credits by the end of the first school year of the experiment, and that this average is no different for the other intervention groups. The estimates are very close to zero with small 95 percent confidence regions. We can rule out effects larger than 6 percent of a standard deviation.

Eighty percent and 75 percent of students in our control group enroll in courses at UofT the second and third year, respectively, after taking the warm-up exercise. We find no significant differences between these persistence rates and those for students in any of our intervention groups.

Only in the third year since taking a warm-up exercise do we find significant impacts on credits earned. These results are driven by the online and follow-up coaching programs given in the second year of SAL. Students receiving the online coaching intervention (with or without one-way coaching) earn about 9 percent of a standard deviation more credits than those in the control group. The impact on the 24 students selected to receive proactive face-to-face coaching is particularly large—63 percent of a standard deviation, though this result is only one out of thirty-eight estimates in the table that is significant at the 1 percent level. Table A5 shows similarly small and insignificant results of each treatment separately on both contemporaneous and persistent outcomes.

C. Academic Outcomes for Students at Greater Risk of Poor Performance

In previous work examining some of these interventions separately, we estimated treatment effects for dozens of different sub-groups after conducting a pre-analysis to focus on students thought to be more at risk of poor performance than others (Oreopoulos and Petronijevic, 2018; Oreopoulos et al., 2020; Oreopoulos et al., forthcoming). We found no convincing evidence that the interventions improved first-year academic outcomes for any of the sub-groups examined.[24]

In this paper, we summarize heterogeneous effects by focusing on students at risk of performing poorly academically in their first term. Specifically, we first use the control group

[24] Some subgroup examples include students who are male, first year, first generation, international students, live with their parents, working at least 8 hours per week, not sure about their program of study, self-report they tend to procrastinate, intend to complete their education with no more than an undergraduate degree, and expect to earn less than an A- grade average.

sample to estimate a students' propensity (probability) for receiving a grade less than 60 percent (a C minus), conditioning on a cubic function for high school admissions grade, mother's education, father's education, age, days since warm-up exercise introduced before registering, indicator variables for English as a second language and gender, and fixed effects for clustered sample group used for randomization. Each treated and control student for which we had such background information was then assigned a propensity score and ranked in order from highest at risk for predicted poor performance to lowest.[25] Table 6 shows estimated intervention effects on fall semester grades by the end of the same term that the warm-up exercise was introduced, for students with a non-missing high school admissions grade. This sample tends to omit some students who completed high school outside the province of Ontario, including outside of Canada. As indicated in column 1, however, all average treatment effect estimates are again insignificant in this sample. Likewise, there are no clear positive effects of any treatment on students who are at risk of earning low grades, consistent with there being no impact on academic outcomes across a wide range of student subgroups.

D. Mental Health and Student Experience

We conducted follow-up surveys at the end of the fall semester as part of the three experiments that occurred in 2016, 2017, and 2018. As with the initial warm-up exercise, students received a

[25] We estimate the propensity score using a leave-one-out procedure for students in the control group to avoid introducing bias in the subsequent analysis of treatment effects.

small grade for completion to encourage participation.[26] We use the follow-up surveys to investigate whether our interventions affected non-academic outcomes and intermediate outcomes that we cannot observe with administrative data. We asked a standard question about subjective well-being: "All things considered, how satisfied are you with your life as a whole these days?" Students responded on a 1 (Not at all satisfied) to 7 (Absolutely satisfied) scale. We also asked how satisfied they were with their university experience and whether they have felt stressed, sad, or depressed since the beginning of the academic year (0 (rarely or none of the time), 1 (some or a little of the time), 2 (occasionally or a moderate amount of the time), or 3 (most or all of the time)).

Table 7 shows estimated treatment effects on standardized measures of these variables (all converted to have mean zero and standard deviation one). The table reveals large impacts for students assigned to receive face-to-face personal coaching. Self-reported university and life satisfaction are 20 and 23 percent of a standard deviation higher compared to the control group, respectively. The impacts on feeling stressed or depressed are also large but imprecisely estimated. If we create an overall mental health measure by averaging across these standardized variables, we estimate a 25.8 percent increase. There is also suggestive evidence that the interventions with text-message coaching improved overall mental health, though the impacts are smaller, with point estimates ranging from 3.7 to 8.5 percent of a standard deviation and significant only at the 10 percent level. Combining all online coaching treatments to increase statistical power, the treatments are estimated to raise overall mental health by 4.4. percent of a standard deviation, significant at the 5 percent level.

[26] There are no significant treatment effects on starting or completing the follow-up surveys (the first row of Table 7 shows estimated impacts), though participation rates were lower in general (about 76 percent). This was largely due to students having already dropped the economics course and no longer being invited or required to take the survey.

During one of the follow-up surveys, in 2016-17, we asked students about their university experience so far. Specifically, using a 1 to 6 scale, we asked whether students agree they feel like they belong at their university, whether being a university student is an important part of how they see themselves, whether they think their university wants them to be successful, and how confident they feel that they have the ability to succeed at their university. The bottom panel of Table 7 shows estimated coaching effects from our interventions on standardized versions of these measures. Again, for students offered face-to-face coaching, students indicate feeling much more supported and confident. The program seems to generate a clear sense that the university is trying to support their education. Students' sense of belonging and university support is 27.8 percent of a standard deviation higher with face-to-face coaching. They feel significantly more confident they will succeed as well. Those assigned to Two-Way Text Message Coaching also feel more supported, but less so than those with Face-to-Face Coaching. Overall feelings of university support are about 6 percent of a standard deviation higher than in the control group.

E. Study Behavior and Attitudes

We asked students at the end of the term about how much they studied during a typical week outside of midterms and finals. Table 8 shows estimated effects from our coaching interventions on this standardized outcome. Students assigned to Online and Two-Way Text Coaching studied 11.3 percent of standard deviation more, on average, than students in the control group, or about 1.3 hours. We find no significant effects for those assigned to receive face-to-face follow-up with a personal coach, though these estimates have wide confidence intervals (we cannot reject zero at the 95 percent significance level, but we cannot reject an effect size of 27.8 percent either). Those assigned to Online and Two-Way Text Coaching are also significantly less likely to report they cram for exams, less likely to miss class, and more likely to feel they manage their time well. We find some marginal significant effects on these outcomes for those assigned to only the online treatment, without follow-up coaching. Finally, we find some less precise but notably larger estimated effects on positive study strategies from face-to-face coaching, including rewriting course material in one's own words, seeking feedback, and managing time well. If we average over these standardized measures to create a summary measure of overall positive study behavior, we find a marginally significant effect from online coaching only (6 percent of a standard deviation), a larger impact from Online and Two-Way Text Coaching (13 percent) and an even larger effect from face-to-face coaching (19 percent). Overall, similar to the pattern of results found in Table 7 for the estimated treatment effects on mental health and student experience outcomes, we find small significant effects on study behaviors from the virtual coaching treatments, and large effects from Face-to-Face coaching.

One concern with these results is that treated students may feel more obliged to self-report more study hours than the control group, even though actual hours are the same. To address this, we asked multiple questions about study time during the last follow-up survey in the 2018-19 academic year (in which we only tested one intervention). As indicated in the bottom half of Table 8, we find significant effects from assigning students to receive Two-Way Text Coaching for all of our study time measures: self-reported weekly study time across all courses (measured in hours), for only their economics course, and the amount they plan to study each week the following (winter) semester. We also asked each student to create a brief time diary documenting what they did 'yesterday' (i.e., the day before the took the follow-up survey). Added up, students studied, on average, 3.3 hours per day with a standard deviation of 2.7 hours. We estimate students assigned to receive Online and Two-Way Text Coaching report studying an average of 0.3 hours more in the previous day, which averages to 2.1 hours over a week, similar to the estimated effects using the subjective weekly study time variables. The magnitudes of these impacts on study time are also similar, ranging from 10 to 20 percent of a standard deviation.[27]

IV. A Model of Student Effort

Why were the interventions we evaluated ineffective at improving overall academic achievement? Perhaps students already optimize when choosing how much they want to study and how to study efficiently relative to their abilities and preferences. Or perhaps these low-cost interventions are not intensive or personal enough to meaningfully change habits or goals. To explore these issues further, we describe a simple model of study effort to better understand the mechanisms by which our interventions affected study behavior but not academic performance. We then map the model to our survey data gathered during the fall semesters of our fifth and sixth experimental years (2018-19 and 2019-20), track how students' beliefs about their academic abilities, study choices,

[27] In Oreopoulos et al. (forthcoming), we argue that our treatment effects on study time are real and that study time generally affects achievement positively. We then show that the association between study time and grades is positive, but weak. At most, a one-standard deviation (13 hour) increase in weekly study time is associated with a 5.72 percentage-point increase in mean math grades. If we assume this as a ball-park estimate for the causal return to study time, our treatment-driven increase of 2.28 hours of studying per week is predicted to cause an increase in mean grades of 1.01 percentage points or 6.1 percent of a standard deviation. This is a small effect on achievement and one which we often cannot reject as being the treatment effect of our coaching interventions. Taken together, our results therefore suggest that our coaching interventions improved study behavior but not enough in magnitude to observe a significant and meaningful improvement in academic performance.

and grade expectations change over the semester, and measure the impact our interventions had on these objects.

Four main takeaways arise from the analysis. First, students' beliefs about their academic abilities and their preferences for attaining high grades rationally determine their study effort and grade expectations. Second, students systematically study four to five fewer hours per week than intended, however, suggesting procrastination or other behavioral barriers are also at play. Third, our interventions increased study time not by reducing procrastination but by increasing academic ambition (which we measure as a willingness to spend time studying to obtain higher grades) and by causing students to believe that greater weekly study time (relative to their initial beliefs) is needed to achieve high grades. Fourth, better timed and more personalized interventions may be needed: upon learning it is harder to do well than initially expected, many students substantially revise down their grade expectations but do not change study time. Grade expectations decline particularly sharply around midterm season (and do not recover subsequently), suggesting there may exist critical periods in which to intervene with personalized support to help students avoid simply accepting poor performance as inevitable.

A. The Education Production Function and Student Expectations

In the model, students take their expected abilities and preferences at the beginning of the semester as given and decide how much study effort to exert. They then learn more about their abilities and preferences, revise their initial expectations, and update their study decisions and grade expectations accordingly. The difference between the time they report studying at the end of the semester and the time they expect to study at the beginning is a function of both rational information updating and a behavioral deviation from that update, which we refer to as procrastination.[28]

Let y_i denote the grade earned by student i at the end of the fall semester. We assume that the weekly study effort of each student, s_i, is mapped into grades according to the following linear production technology:

$$y_i = \alpha_i + \beta_i s_i + \epsilon_i, \tag{1}$$

where α_i is the academic ability of student i—i.e., the grade she would expect to earn without any study effort—β_i is the return to each unit of additional studying for student i, and ϵ_i is an error term with mean zero.[29]

Students are uncertain about their academic abilities and returns to study effort at the beginning of the fall semester. We let $\hat{\alpha}_{i0}$ and $\hat{\beta}_{i0}$ denote, respectively, student i's expected ability and return to study effort at the start of the semester. Similarly, we let $\hat{\alpha}_{i1}$ and $\hat{\beta}_{i1}$ represent the updated values for these objects at the end of the semester. For a given amount of study intensity at time t (s_{it}), student i therefore expects to earn the following grade

$$\mathbb{E}_t(y_i|s_{it}) = \hat{\alpha}_{it} + \hat{\beta}_{it} s_{it}, \tag{2}$$

where $t = 0$ and 1 denote the beginning and end of the semester, respectively.[30] With their grade expectations in mind, students then make study decisions according to their preferences over grades and the cost of study effort.

[28] The difference could also arise from mistakes in time management, over-confidence with initial expectations, or lack of salience.
[29] We assume the simple linear specification for the production technology to keep the analysis tractable and to allow for an intuitive mapping of the theory to the survey data, where we ask students about their expected abilities and returns to studying during initial and follow-up surveys.
[30] More precisely, in our data, the beginning and end of the semester are the times when students take the initial and follow-up surveys.

B. Student Preferences

We assume that students perceive the benefits of higher grades in discrete categories, defined by the grade cutoffs that correspond to the letter grades A, B, and all other letter grades that are up to and below a C. Specifically, we let θ_{it}^{j} denote the utility benefit obtained by student i when she earns letter grade j and assume that $\theta_{it}^{A} > \theta_{it}^{B} > \theta_{it}^{C}$. Student i earns an A when $y_i > y^A$, earns a B when $y_i > y^B$, and earns a C when $y_i > y^C$, where $y^A > y^B > y^C$. At both the beginning and end of the semester, each student exerts of level of student intensity s_{it} to increase her expected grade, given by equation (2). The cost of study effort is given by the strictly increasing and convex function $c(s_{it})$.

We assume the benefit students derive from higher grades only changes (discretely) when they earn a grade that crosses a threshold for a higher letter grade: continuous changes in percentage grades within a given letter grade category do not give rise to any change in the benefit students derive from their study effort. We make this assumption because the patterns in our data suggest that students do indeed place much importance on attaining grades that correspond to certain thresholds. Some of this behavior is due to explicit thresholds determining whether students are admitted to specialized or honors programs. In Appendix B, we show that student percentage grade expectations bunch at multiples of ten, which indicate transitions between letter grades at UofT, and that only 30 percent of students report preparing for a test until they completely understand the material, with the remaining 70 percent preparing only enough to earn various letter grades. [31,32]

C. Student Decision-Making and the Interpretation of Observed Study Outcomes

The time students report studying each week at the end of the semester represents a combination of rational revisions to their initial expected study times, which reflect updated information about their academic abilities and preferences, and behavioral deviations from these rational revisions, which we conceptualize as procrastination.

Information-Driven Choice

Using her beliefs in each time period and equation (2), student i determines the minimum amount of study effort that is required for her to expect to earn letter grade j in time period t as

$$s_{it}^{j} = \frac{y^{j} - \hat{\alpha}_{it}}{\hat{\beta}_{it}}, \tag{3}$$

[31] A model in which the benefit of higher grades is continuous (and increasing and concave) in the grade earned delivers similar predictions about student behavior but with some important differences. In particular, a model with a continuous benefit implies that students revise their expected grades upward when receiving a positive information update about their abilities and that they revise study time upward when receiving a negative information update. We do not find these patterns in the data. Instead, we see an asymmetric response, with students who receive a positive update leaving grade expectations relatively unchanged but decreasing study time choices and students who receive a negative update by downgrading grade expectations substantially but leaving study time relatively unchanged. These patterns can be more easily generated by the threshold-based model we present here.

[32] Also note that, for the purpose of our model, we deliberately group all grades up to and below a C into one category. We do so because (i) allowing for more grade categories does not add to the model's main insights and (ii) the data are consistent with students not differentiating much between letter grades that are a C or below. In Appendix B, we show that less than 2 percent of students expect to earn a grade that is a C or below, both across all courses and economics specifically, while only 9 percent of students report preparing for tests by studying enough to only earn a C or less.

where $j = A, B,$ or C and $t = 0$ or 1. We assume that the distribution of student ability is such that the study time required for the lowest letter grade of C is non-negative for all students, implying that $y^C \geq \hat{\alpha}_{it}$ $\forall$ i and t.

When choosing between whether to exert enough study effort to expect an A or only enough to expect a B, student i compares the additional benefit of earning an A to the cost of additional studying, opting to aim for an A when

$$\theta_{it}^A - \theta_{it}^B \geq c(s_{it}^A) - c(s_{it}^B), \tag{4}$$

where $s_i^{A,t}$ and $s_i^{B,t}$ are defined according to equation (3). Likewise, when choosing between aiming for a B or a C, student i studies enough to expect a B when

$$\theta_{it}^B - \theta_{it}^C \geq c(s_{it}^B) - c(s_{it}^C). \tag{5}$$

As discussed, the descriptive evidence suggests that few students approach their studies by aiming for a C or below. For ease of exposition, we assume that no student prefers to aim for a C over a B.[33] Formally, we normalize the benefit of obtaining a letter grade of C to zero for all students ($\theta_{it}^C = 0$ $\forall$ i, t) and assume that the following condition holds

$$\underline{\theta}_t^B > c\left(\frac{y^B - \underline{\hat{\alpha}}_t}{\underline{\hat{\beta}}_t}\right) - c\left(\frac{y^C - \underline{\hat{\alpha}}_t}{\underline{\hat{\beta}}_t}\right) \tag{6}$$

for $t = 0$ and 1.[34]

[33] As mentioned, we have only three grade thresholds in the model for ease of exposition, making the letter grade C our lower bound. The same intuition can be obtained from a model with more grade thresholds, in which no student prefers to attain a failing grading (F) over a D. However, the more nuanced model would add very little useful content at the expense of expositional clarity.

[34] Here, the underlined objects represent the minimum benefit of obtaining a B grade across all students, and the minimum values of perceived academic ability and the return to studying in each period across all students. Because $c(\cdot)$ is strictly increasing and convex, the right-hand side of equation (6) is decreasing in both α and β, and is therefore maximized at the minimum values of both objects. Likewise, the left-hand side is lowest at $\underline{\theta}_t^B$, implying that condition (6) guarantees no student prefers to study only enough to expect a letter grade of C. In the model, all students will therefore study enough to expect to earn either an A or a B, which is consistent with the descriptive evidence in the data.

With this framework in hand, the optimal study choice of student i in period t is written as

$$s_{it}^{*} = \begin{cases} s_{it}^{A} \text{ if } \theta_{it}^{A} - \theta_{it}^{B} \geq c(s_{it}^{A}) - c(s_{it}^{B}) \\ s_{it}^{B} \text{ if } \theta_{it}^{A} - \theta_{it}^{B} < c(s_{it}^{A}) - c(s_{it}^{B}) \end{cases}. \tag{7}$$

Behavioral Barriers

We assume that observed study time at the end of the semester is given by the rational quantity implied by equation (7) and a behavioral deviation caused by procrastination. Specifically, we write observed study time at the end of the semester, $\tilde{s}_{i1}$, as

$$\tilde{s}_{i1} = \lambda_{i}^{p} + s_{i1}^{*} + \nu_{i}, \tag{8}$$

where λ_{i}^{p} is a student-specific procrastination term and ν_{i} is mean-zero noise. Seen this way, observed study outcomes are a function of rational behavior—based on preferences and expectations about academic ability—and behavioral challenges, such as procrastination tendencies and distractions. In Section V below, we demonstrate how multiple measures of study time from our survey data allow us to identify average procrastination behavior ($\bar{\lambda}^{p}$), while holding constant changes in study time that are driven by rational information updating. We also use our weekly survey data to show how procrastination fluctuates on a weekly basis over the semester.

D. Analyzing Changes in Study Time Over the Semester

We now describe how rational study choices and grade expectations change as students update their beliefs about their abilities and their preferences. We then decompose the difference between actual and initially expected study time into a component driven by rational updating and a component driven by behavioral barriers like procrastination.

Information Updating. We first consider how students who initially plan to study enough to earn a letter grade of A update their grade expectations and study decisions upon learning new information about their abilities. We discuss the intuition here, while Proposition 1 in Appendix B establishes the formal results. When students originally believe they are putting forth enough effort to earn top grades and effort exertion is costly, they reduce their effort upon learning that they are of higher academic ability or that each unit of effort is more productive. In contrast, when students learn it is more difficult to earn top grades than originally expected, and it no longer pays off to continue aiming for an A, students revise their grade expectations down and either marginally reduce or do not change study effort.[35] Proposition 2 in Appendix B establishes analogous predictions for students who are originally aiming for a B. When these students learn it is harder to do well, they increase study time and continue aiming for a B; in contrast, when they learn it is slightly easier to do well, they decrease study time and either do not change or revise up their expected grade.

Taking the propositions together, the model implies students modulate their study effort in response to new information and in a potentially asymmetric way with respect to positive and negative information updates. That is, depending on their initial grade expectations and the size of the information update, the model outlines cases where students who realize it is *harder* to do well respond by not changing (or marginally decreasing) study time choices and decreasing grade expectations. Students are likely to make revisions of this nature when they are initially aiming for an A and receive an intermediate negative shock to their beliefs about their academic abilities.[36] The model also outlines cases where students who realize it is *easier* to earn an A respond by

[35] This is the case when the downward revision to beliefs is not too large. For large negative updates, students increase study effort to ensure that even a lower expected grade is attainable.
[36] When they are initially aiming for an A and receive a large negative shock, they increase study time but still revise grade expectations down.

decreasing study time choices and not changing grade expectations. Revisions of this type are likely to occur when students are originally aiming for an A or aiming for a B and receive a relatively small information update. In the next section, we show the data are consistent with these predictions, as we find that students who receive negative updates to their beliefs about their academic abilities respond by revising grade expectations down but not increasing study time, while students who receive positive updates respond by significantly reducing study time but revising grade expectations far less.

Preferences (Academic Ambition). Holding beliefs about academic ability constant, students' preferences may also change over the course of the semester, thus affecting their study time choices. We interpret a change in preferences that makes students value higher grades more as in increase in $\theta_{it}^{A} - \theta_{it}^{B}$. Proposition 3 in Appendix B formally establishes the intuitive idea that, for a given academic ability and return to studying, students are willing to work harder to earn higher grades when the value they place on attaining higher grades increases.[37]

Procrastination. Rational revisions to study time and grade expectations in our model are driven by changes to information about academic ability and changes in preferences for earning high grades. We also emphasize—both theoretically and empirically—that these rational updates to study choices occur separately from procrastination behavior. That is, using equation (8), the difference between the actual number of hours per week a student reports studying at the end of the semester ($\tilde{s}_{i1}$), and her original expected study time, is

$$\tilde{s}_{i1} - s_{i0}^{*} = \underbrace{\lambda_i^p}_{\text{Procrastination}} + \underbrace{(s_{i1}^{*} - s_{i0}^{*})}_{\text{Rational Update}} + \nu_i. \tag{9}$$

[37] We frame the proposition in terms of the maximum amount of time students are willing to study to earn *A* because we present evidence in Section V that our interventions cause treated students to report being willing to study more hours to earn higher grades than control students. We interpret this as suggestive evidence that our coaching interventions changed students' perceived benefits of higher grades.

Equation (9) makes clear that both students with high and low initial study expectations (s_{i0}^*) can procrastinate, as even students with low initial study goals may optimally desire to revise those goals up throughout the semester but fail to do so because they procrastinate. Indeed, in Section V below, we show that our average measure of procrastination does not differ between students with low and high initial study goals.

In summary, we consider three mechanisms through which our coaching interventions could affect study behavior: (i) changing the information students have about their academic abilities or returns to studying ($\hat{\alpha}_{it}$ and $\hat{\beta}_{it}$), (ii) changing the value students place on earning high grades ($\theta_{it}^A - \theta_{it}^B$), and (iii) helping students reduce procrastination. All are plausible channels through which the interventions could have caused students to increase study time. The interventions emphasized the importance of adequate study time and cautioned students against studying at the last minute, potentially causing them to revise down their expected grades without study effort ($\hat{\alpha}_i$). We also provided students with effective study strategies and tips to help make their study time more effective ($\hat{\beta}_i$), and we asked students to reflect on their long-run goals and the long-term benefits of doing well in college to increase the value they place on earning high grades ($\theta_{it}^A - \theta_{it}^B$). In addition, our programs attempted to mitigate students' tendencies to procrastinate (λ_i^p) by keeping their goals salient throughout the semester and by offering advice on time management and reminders to study.

V. Supporting Evidence for the Model and Decomposing Treatment Effects

We now turn to the data to corroborate our model. We do so with two data sources: (i) the initial and follow-up surveys from the experiment in the 2018-19 academic year and (ii) the data we gathered from the 2019-20 intervention. The intervention in the 2019-20 academic year was a more comprehensive data-gathering effort than in prior years: in addition to students completing an initial survey at the start of the semester and a follow-up survey at the end, students also completed *weekly* surveys, in which they reported their weekly study times and grade expectations over the course of the semester.[38] We use this data to assess how study time and grade expectations evolve on a weekly basis over the semester. As mentioned above, information updating, changes to preferences, and procrastination determine study time and grade expectations in our model and our SAL interventions could potentially influence all of them. We now describe how we construct empirical analogues to each of these objects, before testing the extent to which our coaching interventions affected each factor below.

Information Updating. We concisely summarize changes in both $\hat{\alpha}_{it}$ and $\hat{\beta}_{it}$ with changes in the study time that students expect is needed to earn an A, $s_{it}^{A} = \frac{80-\hat{\alpha}_{it}}{\hat{\beta}_{it}}$, which we label as 'academic savvy'.[39] We construct a measure of academic savvy for each student using the answers they provided in the initial and follow-up surveys during the 2018-19 and 2019-20 experiments. In both years, the initial and follow-up surveys asked students to report the percentage grade they thought they would earn if they studied varying numbers of hours per week.[40] Using the reported

[38] Half of the students in the 2019-20 intervention completed these weekly surveys in which we invited them to plan their study schedule for the week ahead and report their grade expectations. The other half completed different weekly exercises in which we provided tips for doing well in college and invited them to identify common barriers to success in university and potential solutions to those barriers. Both groups needed to complete at least eight weekly exercises (out of 13 possible weeks over the course of the semester). There is no natural control group in the 2019-20 academic year, as both groups received weekly information that could have potentially caused improvement in academic performance and overall well-being. We therefore do not report treatment effects from this wave of data, instead only using it to assess how study time and grade expectations evolve on a weekly basis over the course of the semester.
[39] Recall that a percentage grade of 80 percent or higher is considered an A at UofT.
[40] In 2018-19, students were asked about the grade they believed they would earn in *only their economics course* if they studied 0, 1, 3, 7, 12, and 20 additional hours per week for the course, on top of any cramming two days before the midterm and final exam. In 2019-20, students were asked about the overall grade they believed they would earn *averaged across all courses* if they studied 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 hours per week for all their courses in addition to planned cramming before midterms and exams.

expected percentage grades as the dependent variable and the hours of study options as the independent variable, we estimate two (student-specific) regressions for each student—one with data from the initial survey and one with data from the follow-up—which allow us to construct estimates of each student's expected ability ($\hat{\alpha}_{it}$) and return to studying ($\hat{\beta}_{it}$) at both the beginning and end of the semester. We measure expected ability at each time as the estimated intercept from the relevant regression, while taking the estimated slope as the excepted return to studying. [41]

Preferences (Academic Ambition). During both the 2018-19 and 2019-20 experiments, students were also asked how many hours per week they would be willing to study if doing so guaranteed an overall average grade across all courses of 70, 75, 80, or 85 percent.[42] Because students were told to assume they would receive the grade with certainty if they studied the stated time (regardless of the opportunity cost of that time), the difference between the hours they are willing to study per week to earn an 80 percent average and a 70 percent average reflects their preferences for earning an A relative to a B, independently of any disutility (or cost) of study effort. We therefore take this difference to reflect student preferences for good grades ($\theta_{it}^{A} - \theta_{it}^{B}$).

Procrastination. Students simply falling short of their initial target study hours is suggestive of procrastination problems, but it does not isolate changes in study time that are driven by procrastination. The key challenge is to identify procrastination-driven changes separately from changes that are driven by information updating, despite not directly observing the study time students would have rationally selected only because of information updating. We approach this challenge in two ways, each relying on the *timing* of when students answered survey questions.

[41] We use a simple linear model for tractability and ease of exposition. Regressing expected percentage grades on a quadratic function of hours and proceeding with estimates from that exercise does not meaningfully change the results.
[42] In 2019-20, students were only asked about grades of 75 and 85 percent.

First, among both the 2018-19 and 2019-20 cohorts, we measure average weekly procrastination by taking the difference between students' target hours in the winter (i.e., *next*) term and the hours they report having actually studied during an average week throughout the fall term. Both answers are recorded *at the same time*, during the follow-up survey at the end of the fall term, ensuring that the information students have about their academic abilities and preferences is the same when answering both questions. Procrastination does not determine how many hours they intend to study next semester, however, which instead reflects a choice based on present information. We therefore treat expected weekly study time next semester as an observable proxy to the time students would have rationally chosen to study this semester based on just their preferences and updated information. Denoting expected study next semester as s_{i2}^{*} but assuming it is in fact equal to s_{i1}^{*} in our model, allows us to use equation (9) and obtain the difference between expected study time next semester and actual study time this semester as

$$s_{i2}^{*} - \tilde{s}_{i1} = s_{i1}^{*} - \left(\lambda_i^p + s_{i1}^{*} + \nu_i\right) = -\lambda_i^p - \nu_i. \quad (10)$$

Equation (10) reflects the amount by which study time during the fall semester was affected by procrastination and mean-zero noise. Averaging equation (10) over all students therefore provides an average measure of procrastination, while holding constant (or removing) the effect of information updating.

Second, we also measure procrastination using information on planned and actual study time each week from the *weekly* surveys students completed during the 2019-20 intervention. Each week, students indicated how many hours they planned to study and then had to record how many hours they actually studied when they completed their survey for the following week. It is reasonable to assume that relatively little information updating occurs over the course of one week, implying that the difference between planned and actual study time, recorded on a weekly basis, is largely driven by procrastination or other behavioral barriers and not information updating. As we demonstrate below, both measures of procrastination deliver remarkably consistent results.

A. Characteristics that Relate to Study Effort and Grade Outcomes

Using the information students provided in the end-of-semester follow-up surveys, the top panel of Table 9 documents the associations between study effort and the information updating, preference, and procrastination variables described above, while the bottom panel documents how these variables associated with academic achievement. Three main points emerge.

First, recorded study time during the fall semester depends negatively on students' beliefs about their academic abilities and positively on their preferences for studying. Across both cohorts, study time is decreasing in students' expected grades without studying ($\hat{\alpha}_{it}$) and increasing in the number of hours they believe they need to study to earn an A, consistent with students studying more only when they believe it is required.[43] Also across both cohorts, fall semester study time is strongly correlated with academic ambition ($\theta_{it}^{A} - \theta_{it}^{B}$), with a 10-hour increase in a students' willingness to study for a guaranteed 'A' versus 'B' being associated with about a 2-hour increase in actual study time.[44] All regressions condition of high school grade average, suggesting that

[43] The different magnitudes between some relationships across the cohorts owes to the academic beliefs variables in 2019-20 being measured across all courses and not just economics.

[44] As further supporting evidence for the importance of attitudes toward studying, another proxy variable—how much a student agrees, on a 1 to 6 scale, that they 'like to study'—is also highly correlated with actual study time.

both students' beliefs about their own study effectiveness and their preferences for good grades predict observed study behavior independently of incoming ability.[45]

Second, the bottom panel of Table 9 shows that beliefs about academic ability and preferences for high grades are positively associated with end-of-term grades while procrastination tendencies are negatively associated. Among the 2018-19 cohort, for example, a standard-deviation increase in the grade that students expect without studying plus a standard-deviation increase in the expected return to studying are together associated with a 5.4 percentage-point increase in the actual fall term grade (40 percent of standard deviation). Academic ambition also positively associates with academic achievement in the fall term, with a willingness to study 10 extra hours for an A over B being associated with 1 to 1.5 percentage-point increase in the average fall grade. Measuring procrastination as the difference between students' target hours in the winter (i.e., *next*) term and the hours they report having studied during an average week throughout the fall term shows that procrastination habits are negatively associated with academic achievement in both cohorts.

Third, our data also reveal that students procrastinate between 4 to 5 hours per week, on average. The bottom panel of Table 9 shows that the mean procrastination is 4.9 and 4.5 hours per week among the 2018-19 and 2019-20 cohorts, respectively. We also construct a more direct measure of procrastination using the weekly surveys from the 2019-20 cohort. Students recorded the number of hours they planned to study each week and then came back to record the number of hours they actually studied in that week. We take the difference between the two variables to reflect weekly procrastination and plot the mean of that variable over time in Figure 5. We normalize the

[45] We do not include the measure of procrastination we have for both cohorts in these analyses because it is constructed as the difference between expected study time in the winter semester and actual study time in the fall semester, making it mechanically correlated with the dependent variable (fall semester study time).

week in which students completed the exercise relative to the week of their first midterm in economics. Averaging across all student-weeks, mean procrastination is 4.2 hours per week (depicted by the solid red line)—a remarkably similar estimates to the one we obtain using our first measure of procrastination. The weekly surveys, however, reveal interesting dynamics in procrastination over time, showing that procrastination is minimized during the week of midterms at below three hours per week and maximized at nearly five and a half hours per week when students are four weeks removed from midterm season.

B. Information and Preference Updating and Changes in Study Time and Grade Expectations

Having shown how the key variables of interest from the model are correlated with student study time and performance, we now show how *changes* in students' beliefs about their academic abilities and their preferences are associated with *changes* in study time and grade expectations over the semester. We do so to track how student behavior changes over time and responds to new information.

Four main findings emerge. First, changes to students' beliefs about their academic abilities are negatively associated with changes in study effort and positively associated with changes in grade expectations. Second, students respond to new information about their academic abilities in an asymmetric way: those who realize it is easier to do well reduce study time and do not change grade expectations much, while those who realize it is harder do not change study time but revise grade expectations down substantially. Third, these adjustments occur approximately halfway through the first semester in college, suggesting the need for early and personalized intervention.

And fourth, students who procrastinate more during the semester report lower study time relative to initial expectations but changes in preferences for high grades appear to be a weak predictor of changes in study effort and grade expectations.

Measuring Changes in Beliefs about Academic Ability and Changes in Preferences

To concisely use all the available information when tracking changes in students' beliefs about their academic abilities, we measure information updating for each student as the change in 'academic savvy'—that is, the change in the study hours that are required for students to expect to earn at least an A—over the course of the semester:

$$\Delta s_i^A = \frac{80 - \hat{\alpha}_{i1}}{\hat{\beta}_{i1}} - \frac{80 - \hat{\alpha}_{i0}}{\hat{\beta}_{i0}}. \tag{11}$$

When Δs_i^A is positive, students receive a negative information update during the semester, learning that it is more difficult to earn an A than initially expected and that more study time is required to do so. The opposite is true when Δs_i^A is negative.[46] This is our preferred measure of information updating because it has an intuitive connection to our model.

Among students in the 2019-20 cohort, we also construct the change in preferences for earning an A over a B average from the initial to follow-up survey by taking the difference between our academic ambition measures recorded at follow-up and at baseline: [47]

$$\Delta\theta_i = (\theta_{i1}^A - \theta_{i1}^B) - (\theta_{i0}^A - \theta_{i0}^B)$$

[46] Again, for the 2018-19 cohort, this change reflects beliefs specific to students' economics course; for the 2019-20 cohort, it reflects beliefs across all courses.
[47] We cannot consider changes in preferences among the 2018-19 cohort because we did not record the preferences variable in the initial survey during that intervention.

The difference captures the change in students' willingness to study extra time to guarantee an A average, with positive values indicating a stronger preference for higher grades at the end of the semester than at the start and negative values indicating a weaker preference.

Explaining Changes in Study Effort and Grade Expectations Between the Start and End of the Semester

Table 10 shows how changes in academic savvy relate to changes in study times and grade expectations among the 2018-19 cohort, while Table 11 also considers changes in students' preferences and the average amount of time per week that students procrastinate during the semester (as measured by the mean difference between planned and actual study time across all weekly exercises that a student completed) among the 2019-20 cohort. In both tables, we also estimate specifications with additional control variables, including various demographic and background variables, and flexible (cubic) functions of students' initial expected study times and expected grades. The dependent variables in tables are often changes relative to these initial expectations, making it important to flexibly control for systematic changes throughout the semester that are potentially correlated with information updating.[48]

The estimated coefficients on changes academic savvy are similar across specifications with and without additional controls and are economically significant. Columns 1 and 2 in both Tables 10 and 11 show the association between changes in students' reported weekly study times (from the initial to follow-up surveys) and changes in students' academic savvy (i.e., changes in

[48] For example, suppose some students initially submit very high and unrealistic expectations for study time and grades. We would expect that these students mechanically revise down both study time and grade expectations, and, if such students are also more likely to be overly optimistic about their academic abilities, we would expect to find a correlation between these mechanical revisions and our measure of information updating. Flexibly controlling for the relationship between changes in expectations and initial expectations allows us to identify the effect of information updating conditional on this relationship.

the weekly study hours students believe are required to earn an A). The results indicate that students report studying more hours per week (relative to initial expectations) when they learn that more study time is required to earn good grades. Columns 1 and 2 in Table 10, for example, show that when students expect to have to study 6.5 hours more per week to earn an A in economics—a one standard-deviation change in the independent (academic savvy) variable—they actually study 0.8 hours (16 percent of a standard deviation) more per week for economics than they originally report expecting to study on the baseline survey. Columns 1 and 2 in Table 11 show similar results when considering changes in academic savvy and study time across all courses (not just economics). Specifically, when students learn they have to study approximately 11.4 hours more per week to earn an A average over all courses (a one standard deviation change in the academic savvy variable), they report actually studying 3.5 hours (0.3 standard deviations) more per week than they reported expecting to study at baseline.

The specifications in Table 11 additionally show that changes in grade preferences have a (small) influence on changes in study time independently of students learning about their academic savvy: an 8.7-hour change in willingness to study—a one standard deviation change—associates with a half-an-hour increase in study time (or 4 percent of a standard deviation), although this effect is statistically insignificant. As expected, students' procrastination habits are negatively associated with study time, with a one standard deviation increase in weekly procrastination (6 hours) being associated with a 1.2-hour greater reduction in actual study time relative to expectations.

Columns 3 and 4 in Tables 10 and 11 show that students revise their grade expectations down during the semester when they learn it is harder to earn an A than they originally expected. The point estimate in column 4 of Table 10, for example, indicates that a one standard-deviation increase in required study time is associated with students expecting to earn economics grades that are approximately 6 percentage points (0.32 standard deviations) lower than they originally believed. The point estimate on the change in academic savvy in column 4 of Table 11 shows a similar relationship among the 2019-20 cohort: when a student's academic savvy changes such that they expect to need to study one standard deviation more (11.4 hours per week) to earn an A across all courses, they revise their grade expectations down by 2.2 percentage points (0.27 standard deviations). Conditional on the change in study time students believe is required to earn high grades (as measured by changes in academic savvy), changes in student preferences for good grades are not associated with grade expectation revisions, but a one-standard deviation increase in weekly procrastination is associated with a 0.6-percentage point reduction in the expected grade.

In columns 5 and 6 of Table 10, we show that students accurately revise their grade expectations upon learning new information. Specifically, the dependent variable in these specifications is the difference between students' *realized* economics grades and their expected grades at the start of the semester. The point estimates imply that a one standard-deviation higher change in academic savvy (i.e., the change in the weekly study hours students believe is required to earn an A) is associated with students scoring 4.5 percentage points lower than originally expected. Similar patterns prevail among the 2019-20 cohort in Table 11. Columns 7 and 8 show that a one standard-deviation increase in information updating is associated with an overall grade average that is 1.5 percentage points (or 11 percent of a standard deviation) lower than initially expected.[49] Again, it appears that changes in preferences for high grades have little association with grade revisions and realizations, conditional on changes in academic savvy and other

[49] Columns 5 and 6 in Table 11 also show that students additionally revise down the minimum overall grade average that is acceptable to them when they learn it is harder to well than they originally thought. Column 6, for example, indicates that when academic savvy increases by one standard deviation, students lower their minimum acceptable grade by 1.32 percentage points (0.13 standard deviations).

background variables; however, students who procrastinate more during the semester do go on earn significantly lower than expected grades at the end of the semester.

To further explore the dependence of study time and grade expectation revisions on information updating as measured by changes in academic savvy, we produce several binned-scatter plots between these variables, along with the underlying regression functions, estimated using simple regressions (without additional control variables) with the student-level data. The panels of Figure 6 plot the simple relationships between each of the dependent variables in Table 10 and changes in academic savvy, while the panels of Figure 7 do the same for the dependent variables in Table 11.[50]

Both figures show two main facts. First, while there is a clear positive relationship between changes in academic savvy and study time revisions, students respond to information updates in an asymmetric way: they revise their study time down when learning they need to study fewer hours than initially expected to earn an *A* but they do not revise study time up by much when learning they need to study more hours. In Figure 6(a), for example, the average change in study hours among students who learn they need to study more (those to the right of zero on the horizontal axis) is not statistically different from zero while those who learn they need to study less (those to the left of zero) revise their study time in economics down by 1.34 hours per week. Similar asymmetric responses are seen across all courses in the 2019-20 cohort in Figure 7(a). Second, there is a clear negative association between revisions to grade expectations and changes in beliefs about academic ability, but this response is also asymmetric. In Figures 6(b) and 7(b),

[50] The panels in Figures 6 and 7 are binned scatter plots. Each binned scatter plot is created by first grouping students into 20 equal-width bins (vigintiles) in the distribution of the variable on the x-axis and calculating the mean of both the y- and x-axis variables within each bin. The circles represent these means, while the lines represent the associated linear fit from the underlying student-level data.

students who receive a positive information update revise their expected grade down only a little (or not at all), while students who receive large negative updates revise their grade expectations down substantially.[51]

The patterns presented so far between information updating and changes to study times and grade expectations are consistent with our model's predictions of asymmetric responses to new information. That is, students who learn it is harder to do well (those to the right of zero on the x-axes in Figures 6 and 7) respond with small changes to study time but large downward revisions to grade expectations. In contrast, students who learn it is easier (those to the left of zero on the x-axes in Figures 6 and 7) respond with large reductions in study time but relatively small revisions to grade expectations. It appears that students revise their grade expectations correctly, on average, as realized grades follow a similar profile to the profile of revised expectations (see Figures 6(c) and 7(d)). These documented patterns are all robust, appearing in both the 2018-19 and 2019-20 cohorts.

Weekly Changes in Grade Expectations and Study Times Throughout the Semester

We further substantiate the information content in measured changes to students' academic savvy by using the weekly survey data among the 2019-20 cohort. Specifically, we categorize students according to whether the information update reflected by the change in academic savvy is positive or negative. Students receive a positive update about their academic abilities when the change in academic savvy is negative (i.e., fewer hours are required to earn an A average than originally expected) and they receive a negative update when the change is positive (i.e., more hours are required to earn an A average). We therefore label students as having received a positive

[51] It is also worth noting that nearly all students revise their percentage grade expectations down, on average.

ability update when they are to the left of zero on the x-axis in Figure 7 and as having received a negative update when they are to the right of zero. In Figure 8, we then plot the evolution of grade and weekly study time goals over the course of the semester, where we normalize the week in which the exercise was completed relative to the week of the first economics midterm.

The time series plots in Figure 8 show that, after midterm season in first semester, students substantially revise both their goals for their expected grades and for study time during a typical week. Figure 8(a) shows the evolution of the mean planned overall grade (across all students who completed an exercise in each week) separately among students who received positive and negative ability updates over the semester. Figure 8(b) shows the regression-adjusted means (relative to the week before the first economics midterm) from specifications that include student fixed effects. The figures, which can be conceptualized as the timeseries analogues to Figure 7(b), show that students who received a negative ability update over the semester begin to substantially revise their grade expectations down two weeks after the first economics midterm (when most students receive their grades back), while the trend in grade expectations is relatively flat prior to midterm season. Students who received a positive update also revise expectations down after midterms, but by approximately 70 percent less than those with negative updates.

Figure 8(c) shows the evolution of mean planned weekly study time during a typical week without midterms while Figure 8(d) shows the evolution of the regression-adjusted means. These figures represent the time series analogues to Figure 7(a) and show that students who received a positive update start to plan to study fewer hours per week after midterm season, while those who received a negative update do not meaningfully adjust study goals over the course of the semester.

Overall, the patterns in Figure 8 are consistent with both the direction and the asymmetric nature of the responses documented in Figure 7. Specifically, students who receive negative updates to their beliefs about their academic abilities revise down their grade expectations but do not change study intensity by much (or at all). In contrast, students who revise their beliefs about their study effectiveness upwards over the semester make smaller revisions to the grade expectations but reduce study time more substantially. Further, Figure 8 shows clear evidence that these changes to grade and study goals happen rather abruptly, right after midterm season in first semester. For many students in our sample, the reflects the seventh week of a thirteen-week fall semester, suggesting that many students come to accept that they will not perform as well as they expected halfway through the semester.

C. Treatment Effects on Academic Expectations, Ambition, and Procrastination

Having presented evidence for the key determinants of student study time and grade expectations implied by our model, we now explore which of these variables were impacted by our coaching treatment. We do so by using treatment and control group data from 2018-19 intervention. Table 12 presents treatment effects on variables capturing students' beliefs about their academic abilities, students' preferences for higher grades, and our measure of procrastination in this sample (the difference between planned study time *next semester* and actual study time in the fall semester).

Our coaching intervention did have a small impact on academic savvy, mostly by reducing the grade students believed they would get from minimal studying. In particular, the online and coaching intervention reduced the grade expected from only cramming by 1.27 percentage points, which our model and descriptive evidence above show is associated with students increasing study time. Treated students also report believing the need to study 0.87 hours more per week to earn an A in economics than control students. Our coaching intervention had the largest impact on preferences for high grades or academic ambition, raising the number of hours students are willing to study for a grade of 80 versus 70 by one, and causing a similar increase to the willingness to study for a grade of 85 versus 75, representing an increase of 15 percent of a standard deviation. We also find supporting evidence for an increase in preferences for high grades from estimating significant treatment effects on a variable measuring students' agreement that earning good grades matters more than just ensuring program completion—a variable which likely reflects, in part, students' perceived benefits from higher grades.

Students procrastinate or significantly deviate from their study intentions, as shown above, but our interventions do not reduce this gap. Table 12 shows this lack of effect by reporting estimated treatment effects on multiple measures of procrastination and distraction. We estimate insignificant effects when measuring procrastination as the difference between winter target hours and actual fall term hours and when considering either the population of students with initially high target study hours (above the median) or initially low target hours. We also find no impact on students' self-reported tendencies to feel distracted by social media and video screens (TV, Netflix, etc.).

In summary, our coaching interventions led students to increase study time by causing them to believe it is harder to earn good grades and by increasing the value they place on earning higher grades. The interventions did not successfully assist students with realizing their study time goals, however, as treated and control students do not differ in both measured procrastination and self-reported tendencies to get distracted.

VI. Conclusion

This paper summarizes a five-year effort to improve college performance through an inexpensive and scalable setup in which thousands of students completed a one- to two-hour online exercise for a small grade at the start of the academic year. After registering and completing a brief survey, students were randomly assigned to interventions that we group here into four categories: 1) Online Coaching, in which students were given helpful advice for academic success; 2) Online and One-Way Text Coaching and 3) Online and Two-Way Text Coaching, in which students also received follow-up text messages of tips, reminders and, in the latter case, an opportunity to communicate regularly with a personal coach; and 4) Online and Face-to-Face Coaching in which real coaches were assigned to students and proactively tried to meet regularly with them.

The fidelity of the experiments was very high. The grade requirement ensured a large representative sample of students from a large first-year economics course participated in the experiments at low cost. About 95 percent of those asked to complete the exercise did so. Feedback and open-ended responses suggested that students took the tasks seriously, thought carefully about the information provided, and were overall quite positive about the experience. Most of those who received follow-up virtual and face-to-face coaching wished for the program to continue for them and to be offered to future students.[52]

Despite the positive experiences, we found that our coaching interventions had no discernable impact on contemporaneous or persistent academic outcomes. The interventions did, however, marginally increase student study time and make students feel a greater sense of support and belonging. We explained these results by paring newly gathered weekly survey data with a

[52]The platform provides a unique way to collect a large set of quantitative and qualitative data over time. Other colleges and institutions can administer our exercises at their own institutions or modify them to ask other questions and try other interventions Details of the interventions and assistance for designing similar experiments are available on this paper's online appendix and through the website studentachievementlab.org.

simple model of student effort. The model highlighted four main reasons for poor student performance: low academic ability, low expected return from studying, low preferences for good grades, and procrastination or other behavioral barriers—all of which may change during the semester.

Our data suggest that many students realize over the semester that they need to work harder to attain their grade goals. They react, however, by committing to study a small amount more than before or not changing study behavior at all, and instead begin to expect lower academic performance as inevitable. We find that these rational revisions to study behavior and grade expectations occur for many students approximately halfway through their first semester of college. We also find evidence of considerable procrastination, estimating that, on average, students in our sample study four to five fewer hours a week than they would prefer based on rational choices. Our coaching interventions increased study time by making students realize they needed to study more to attain good grades and increasing the value students place on earning higher grades. The interventions did not, however, affect students' tendencies to procrastinate.

Taken together, the evidence collectively points to the need for more comprehensive, better timed, and more personalized interventions for supporting students while in college. Our results cast serious doubt about the ability of low-touch programs to create the changes in student attitudes and behavior that are needed to sustain long-run benefits.

References

Acemoglu, Daron, and David Autor. 2012. "What does human capital do? A review of Goldin and Katz's The Race between Education and Technology." *Journal of Economic Literature*, 50(2): 426-463.

Andersen, Steffen, Glenn W Harrison, Morten I Lau, and E Elisabet Rutström. 2008. "Eliciting risk and time preferences." *Econometrica*, 76(3): 583-618.

Arcidiacono, Peter. 2004. "Ability sorting and the returns to college major." *Journal of Econometrics*. 121: 343-375.

Arum, Richard, and Jospia Roksa. 2011. "Academically Adrift: Limited Learning on College Campuses." Chicago, IL: University of Chicago Press.

Arum, Richard, and Josipa Roksa. 2014. "Aspiring Adults Adrift. Chicago, IL: University of Chicago Press.

Babcock, Philip, and Mindy Marks. 2011. "The falling time cost of college: Evidence from half a century of time use data." *Review of Economics and Statistics*, 93(2): 468-478.

Beattie, Graham, Jean-William P. Laliberte, and Philip Oreopoulos. 2018. "Thrivers and divers: Using non-academic measures to predict college success and failure." *Economics of Education Review*, 62: 170-182.

Beattie, Graham, Jean-William Laliberte, Catherine Michaud-Leclerc, and Philip Oreopoulos. 2019. "What sets college thrivers and divers apart? A contrast in study habits, attitudes, and mental health." *Economic Letters*, 178: 50-53.

Bettinger, Eric, and Rachel Baker. 2014. "The Effects of Student Coaching: An Evaluation of a Randomized Experiment in Student Advising." *Educational Evaluation and Policy Analysis,* 35(1): 3-19.

Bettinger, E. P., Long, T. B., Oreopoulos, P., & Sanbonmatsu, L. 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): 1205-1242.

Bird, Kelli A.m Benjamin L. Castleman, Jeffrey T. Denning, Joshua Goodman, Cait Lamberton & Kelly Ochs Rosinger. 2019. "Nudging at Scale: Experimental Evidence from FAFSA Completion Campaigns." NBER Working Paper Number 26158. Cambridge, MA: National Bureau of Economic Research.

Brint, Steven, and Allison M. Cantwell. 2010. "Undergraduate time use and academic outcomes: Results from the University of California Undergraduate Experience Survey 2006." *Teachers College Record*, 112: 2441-2470.

Carrell, Scott E. and Michal Kurlaender. 2020. "My Professor Cares: Experimental Evidence on the Role of Faculty Engagement." NBER Working Paper Number 27312. Cambridge, MA: National Bureau of Economic Research

Castleman, B. L., & Page, L. C. 2015. "Summer Nudging: Can Personalized Text Messages and Peer Mentor Outreach Increase College Going Among Low-Income High School Graduates?" *Journal of Economic Behavior & Organization*, 115: 144-160.

Castleman, B. L., & Page, L. C. 2016. "Freshman Year Financial Aid Nudges: An Experiment to Increase FAFSA Renewal and College Persistence." *The Journal of Human Resources*, 51(2): 389-415.

Castleman, B. L., Page, L. C., & Meyer, K. 2016. "Customized Nudging to Improve FAFSA Completion and Income Verification." SSRN Working Paper http://dx.doi.org/10.2139/ssrn.2854345.

Castleman, Benjamin L. and Zachary Sullivan. 2019. "Cash for college apps: The effects of conditional cash transfers on selective college enrollment." University of Virginia Working Paper.

Chetty, Raj, John N. Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2017. "Mobility Report Cards: The Role of Colleges in Intergenerational Mobility." NBER Working Paper 23618. Cambridge, MA: National Bureau of Economic Research.

Clotfelter, Charles T., Steven W. Hemelt and Helen F. Ladd. 2018. "Multifaceted Aid for Low-Income Students and College Outcomes: Evidence from North Carolina." *Economic Inquiry*, 56(1): 278-303.

Deming, David. 2019. "What does education do?" Kershaw Award Lecture, APPAM Conference Fall 2018. https://www.dropbox.com/s/30zgwghqn88tdnq/David%20N.%20Kershaw%20Award%20What%20Does%20Education%20Do.mp4

Donnellan, M Brent, Frederick L Oswald, Brendan M Baird, and Richard E Lucas. 2006. "The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality." *Psychological Assessment*, 18(2): 192.

Dohmen, Thomas, Armin Falk David Huffman Uwe Sunde Jürgen Schupp, and Gert G. Wagner. 2011. "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences." *Journal of the European Economic Association,* 9(3): 522-550.

Duckworth, Angela Lee, and Patrick D Quinn. 2009. "Development and validation of the Short Grit Scale (GRIT–S)." *Journal of Personality Assessment*, 91(2): 166–174.

Dynarski, Susan, C.J. Libassi, Katherine Michelmore, Stephanie Owen. 2018. "Closing the Gap:

The Effect of a Targeted, Tuition-Free Promise on College Choices of High-Achieving, Low-Income Students." NBER Working Paper No. 25349. Cambridge, MA: National Bureau of Economic Research.

Evans, William N., Melissa S. Kearney, Brendan C. Perry, James X. Sullivan. 2017. "Increasing community college completion rates among low-income students: Evidence from a randomized controlled trial evaluation of a case management intervention" NBER Working Paper 24150. Cambridge, MA: National Bureau of Economic Research.

Farkas, Gary J., Ewa Mazurek, and Jane R. Marone. 2014. "Learning style versus time spent studying and career choice: Which is associated with success in a combined undergraduate anatomy and physiology course?" *Anatomical Sciences Education*, 9: 121-131.

Giedd, Jay N., Michael Stockman, Catherine Weddle, Maria Liverpool, Gregory L. Wallace, Nancy Reitano Lee, Francois Lalonde, and Rhloshel K. Lenroot. 2012. "Anatomic Magnetic Resonance Imaging of the Developing Child and Adolescent Brain." In *The Adolescent Brain: Learning, Reasoning, and Decision Making* edited by Valerie F. Reyna, Sandra B. Chapman, Michael R. Dougherty and Jere Confrey. Washington, D.C.: American Psychological Association.

Gurantz, O., Howell, J., Hurwitz, M., Larson, C., Pender, M., & White, B. 2019. "Realizing Your College Potential? Impacts of College Board's RYCP Campaign on Postsecondary Enrollment" EdWorkingPaper No.19-40.

Gurría, Angel. 2009. Remarks for the OECD launch of Education at a Glance 2009. OECD Conference Centre, Paris, 8 December 2009.

Hirsh, Jacob B, and Jordan B Peterson. 2008. "Predicting creativity and academic success with a "fake-proof" measure of the Big Five." *Journal of Research in Personality*, 42(5): 1323–1333.

Hoxby, Caroline M., and Kevin Stange. 2018. "Introduction to 'Productivity in Higher Education," Chapter in forthcoming NBER book *Productivity in Higher Education*, Caroline M. Hoxby and Kevin Stange, editors. University of Chicago Press.

Kirkeboen, Lars. J., Edwin Leuven, and Magne Mogstad. 2016.. "Field of Study, Earnings, and Self-Selection." *Quarterly Journal of Economics,* 131(3): 1057-1111.

Lavecchia, Adam, Heidi Liu, Philip Oreopoulos. 2016. "Behavioral Economics of Education: Progress and Possibilities." In *Handbook of Economics of Education Volume 5* edited by Eric A. Hanushek, Stephen J. Machin, Ludger Woessmann, 1-74. North Holland Press, Amsterdam.

Lochner, Lance, and Alexander Monge-Naranjo. 2012. "Credit Constraints in Education." *Annual Review of Economics*, 4: 225-256.

Martin Nybom. 2017. “The Distribution of Lifetime Earnings Returns to College.” *Journal of Labor Economics,* 35(4): 903-952.

Oreopoulos, Philip. 2020. “Promises and Limitations of Nudging in Education.” IZA Institute of Labor Economics Discussion Paper Number 13718.

Oreopoulos, Philip, and Reuben Ford. 2019. “Keeping College Options Open: A Field Experiment to Help All High School Seniors Through the College Application Process.” *Journal of Policy Analysis and Management*, 38(2): 299-329.

Oreopoulos, Philip, Richard W. Patterson, Uros Petronijevic, and Nolan G. Pope. “Low-Touch Attempts to Improve Time Management among Traditional and Online College Students.” *Journal of Human Resources*, forthcoming.

Oreopoulos, Philip, and Uros Petronijevic. 2013. “Making college worth it: A review of the returns to higher education.” *Future of Children*, 23(1): 41-65.

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: 342-360.

Oreopoulos, Philip, and Uros Petronijevic, 2018. “Student Coaching: How Far Can Technology Go?” *Journal of Human Resources*, 53(2): 299-329.

Ost, Ben, Weixiang Pan, and Douglas Webber. 2018 “The Returns to College Persistence for Marginal Students: Regression Discontinuity Evidence from University Dismissal Policies." *Journal of Labor Economics,* 36 (3): 779-805.

Page, Lindsay C., Jeonghyun Lee, and Hunter Gehlbach (2020). “Conditions under which college students can be responsive to nudging,” Annenberg Institute at Brown University, Working Paper: 20-242.

Psacharopoulos, George, and Harry Anthony Patrinos. 2018. “*Returns to investment in education: a decennial review of the global literature.*” *Education Economics,* 26(5): 445-458.

Scrivener, Susan, Michael J. Weiss, Alyssa Ratledge, Timothy Rudd, Colleen Sommo, and Hannah Fresques. 2015. “Doubling Graduation Rates: Three-Year Effects of CUNY’s Accelerated Study in Associate Programs (ASAP) for Developmental Education Students.” Report, MDRC, New York, NY.

Shapiro, D., Dundar, A., Huie, F., Wakhungu, P., Bhimdiwala, A., and Wilson, S. E. 2019. Completing College: Eight Year Completion Outcomes for the Fall 2010 Cohort (Signature Report No. 12c). Herndon, VA: National Student Clearinghouse Research Center.

Sommo, Colleen, Dan Cullinan, and Michelle S. Manno. 2018. “Doubling graduation rates in a new state: Two-year findings from the ASAP Ohio Demonstration," Policy Brief, MDRC, New York, NY. https://www.mdrc.org/publication/doubling-graduation-rates-new-state

Stinebrickner, Ralph and Todd R. Stinebrickner. 2012. "Learning about Academic Ability and the College Dropout Decision." *Journal of Labor Economics*, 30(4): 707-748.

Stinebrickner, Ralph and Todd R. Stinebrickner. 2014a. "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout." *Review of Economic Studies*, 81: 426-472.

Stinebrickner, Ralph and Todd R. Stinebrickner. 2014b. "Academic Performance and College Dropout: Using Longitudinal Expectations Data to Estimate a Learning Model." *Journal of Labor Economics*, 32(3): 601-644.

University of Toronto. 2018. "Enrolment Report 2017-18, February 16, 2018, Planning and Budget Office.

Wiswall, Mathew and Basit Zafar. 2015. "Determinants of College Major Choice: Identification using an Information Experiment." *Review of Economic Studies*, 82: 791-824.

Zimmerman, Seth D. 2014. "The Returns to College Admission for Academically Marginal Students." *Journal of Labor Economics*, 32(4): 711-754.

Table 1
Random Assignment to Different Treatment and Control Groups Across Year and Campus

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Year	**2015**	**2015**	**2016**	**2016**	**2016**	**2017**	**2017**	**2018**
Campus	**All**	**All**	**UTSG**	**UTM**	**UTSC**	**UTSG**	**UTM**	**All**
Sample	**1st Years**	**Upper Years**	**All**	**All**	**All**	**All**	**All**	**All**
Group								
Online Coaching Only	**962**	**334**	**1152**		**600**			
	(19.6) [19.6]	(28.4) [29.4]	(34.1) [33.3]		(50.2) [50.0]			
Online and One-Way Text Coaching	**1506**	**493**		**710**				
	(30.7) [30.0]	(41.9) [40.0]		(48.7) [47.5]				
Online and Two-Way Text Coaching			**1118**			**787**	**670**	**2723**
			(33.1) [33.3]			(26.8) [33.3]	(47.4) [50.0]	(49.7) [50.0]
Online and Face-to-Face Coaching	**17**	**7**		**66**				
	(0.4) [0.4]	(0.6) [0.6]		(4.5) [5.0]				
Control (Personality Test)	**1451**	**342**	**1112**	**681**	**595**	**1106**	**743**	**2689**
	(29.6) [30.0]	(29.1) [30.0]	(32.9) [33.3]	(46.7) [47.5]	(49.8) [50.0]	(37.7) [33.3]	(52.6) [50.0]	(50.3) [50.0]
Total Sample Size **19864**	**3936**	**1176**	**3382**	**1457**	**1195**	**1893**	**1413**	**5412**

Notes: The table displays the number of University of Toronto students enrolled in a first-year economics course assigned to each experiment category by year, campus and sample. Values in round brackets show the percent assigned to a group relative to each randomized sample. Values in square brackets show the expected percent assigned to each group based on the assignment rule. UTM = University of Toronto at Mississauga campus, UTSG = University of Toronto at St. George (downtown) campus, UTSC = University of Toronto at Scarborough campus.

Table 2
Descriptive Statistics and Balance Tests

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Control Mean [standard dev.]		**Online Coaching**		**Online and One-Way Text Coaching**		**Online and Two-Way Text Coaching**		**Online and Personal Coaching**	
Want Grad. Degree	0.66	[0.472]	-0.01	[0.011]	0.02	[0.012]	-0.01	[0.009]	-0.01	[0.051]
Father's Education [0-8]	5.1	[2.26]	-0.05	[0.053]	-0.03	[0.060]	-0.03	[0.043]	-0.22	[0.245]
Mother's Education [0-8]	4.73	[2.21]	0.00	[0.052]	0.04	[0.058]	-0.06	[0.041]	-0.19	[0.237]
First Generation Student	0.29	[0.456]	0.01	[0.011]	0.00	[0.012]	0.01	[0.009]	0.02	[0.049]
Parent has Grad. Degree	0.34	[0.473]	0.00	[0.011]	0.00	[0.012]	-0.02	[0.009]*	-0.01	[0.051]
First-Year Student	0.787	[0.410]	0.00	[0.008]	0.00	[0.009]	0.01	[0.007]*	0.06	[0.038]
International Student	0.32	[0.466]	-0.01	[0.010]	0.00	[0.011]	0.00	[0.008]	0.05	[0.046]
Tendency to Not Cram [1-7]	3.9	[1.50]	0.05	[0.035]	0.00	[0.040]	0.03	[0.028]	-0.01	[0.162]
Exp. Avg. Weekly Study Hrs.	18	[12.02]	0.06	[0.279]	0.27	[0.314]	-0.11	[0.223]	-0.89	[1.283]
Exp. Avg. Weekly Work Hrs.	6.89	[9.44]	-0.13	[0.218]	-0.23	[0.245]	0.02	[0.174]	0.09	[1.001]
Exp. Fall Grade [0-100]	80.7	[6.76]	0.07	[0.150]	0.24	[0.169]	-0.01	[0.120]	0.63	[0.691]
# Days Since Sept 1 Began Exercise	10.7	[4.69]	-0.01	[0.098]	0.21	[0.110]*	-0.05	[0.078]	-1.91	[0.451]***
Grit Score: Finish What I Begin [1-5	3.8	[0.825]	-0.02	[0.026]	-0.07	[0.044]	0.03	[0.016]**	0.17	[0.106]
English Mother Tongue	0.42	[0.493]	0.00	[0.012]	0.00	[0.013]	0.00	[0.010]	-0.09	[0.054]*
Male	0.48	[0.500]	0.00	[0.012]	0.00	[0.013]	0.00	[0.010]	-0.01	[0.055]
Age	20	[2]	0.02	[0.040]	0.08	[0.045]*	-0.02	[0.032]	0.17	[0.185]
No High School Grade Data	0.27	[0.446]	0.00	[0.011]	0.04	[0.012]***	0.00	[0.008]	0.05	[0.048]
HS Grade Admissions Avg [0-100]	85.4	[7.2]	0.03	[0.189]	-0.05	[0.209]	0.09	[0.153]	-0.91	[0.846]

Notes: Column 1 lists each background variable (recorded prior to random assignment). Want Grad. Degree = highest expected education attainment is more than a Bachelor degree. Father and mother education categories range from none (0) to Doctorate degree (8). Exp. = Expected. Avg. = Average. HS = High School. Grad. = Graduate. Column 2 displays the mean of these variables among the control group, while column 3 shows the standard deviation. Columns 4, 6, 8, and 10 show the difference between the variable mean for the indicated treatment and control groups. Columns 5, 7, 9, and 11 show the estimated standard errors for these differences. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels respectively.

Table 3

Control Group Outcome Means and Standard Deviations by Cohort and Year Since Experimen

(1) Relative Year Since Exp. Began	(2) 2015	(3) 2016	(4) 2017	(5) 2018
Fall Grade Avg., Year 1	69.1	68.6	68.5	70.2
	[14.1]	[13.6]	[13.0]	[13.0]
Winter Grade Avg., Year 1	68.4	68.2	68.2	69.8
	[15.6]	[15.7]	[15.4]	[14.7]
Year 1 Grade Average	67.6	66.5	67.1	69.1
	[13.6]	[14.3]	[13.9]	[13.3]
Year 2 Grade Average	68.8	69.1	70.1	71
	[12.8]	[13.1]	[12.0]	[12.9]
Year 3 Grade Average	70.7	71.5	72.6	
	[12.8]	[11.8]	[11.8]	
Year 4 Grade Average	72.4	72.8		
	[12.1]	[13.0]		
Year 1 Total Credits Earned	3.5	3.4	3.3	3.5
	[1.5]	[1.6]	[1.6]	[1.6]
Year 2 Total Credits Earned	3.2	3.1	3.2	
	[1.9]	[1.9]	[2.0]	
Year 3 Total Credits Earned	3	2.9		
	[2.0]	[2.1]		
Year 4 Total Credits Earned	2.7			
	2.1			
Persistence Year 2	0.911	0.892	0.897	0.856
Persistence Year 3	0.838	0.811	0.8	
Persistence Year 4	0.804	0.747		
Graduated by End of Year 4	0.461			

Notes: The table shows outcome means and, in square brackets, standard deviations for the control groups from each year of the experiment. Grade averages (Avg.) are listed as a percernt. Persistence variables show the fraction of first-year students in the first year of the experiment with any grade data in the following second, third, and fourth years. The graduation variable indicates the fraction officially gradating with any degree by the Fall Term of 2019 (after four years for first-years in the 2015 experiment).

Table 4
Estimated Treatment Effects on Initial Fall Term Grades [0-100]

	Outcome Variable						
(1)	(2) **Missing Fall Grade**	(3) **Fall Term Grade**	(4) **Grade>50**	(5) **Grade>60**	(6) **Grade>70**	(7) **Grade>80**	(8) **Grade>90**
Online Coaching Only	0.002	0.125	-0.007	0.015	0.015	-0.001	0.002
	[0.008]	[0.340]	[0.007]	[0.010]	[0.013]	[0.010]	[0.004]
Online and One-Way Text Coaching	0.023	0.255	0.007	0.014	0.009	0.001	0.002
	[0.009]**	[0.385]	[0.007]	[0.011]	[0.014]	[0.012]	[0.005]
Online and Two-Way Text Coaching	0.002	-0.057	0.001	0.001	0.001	0.004	-0.001
	[0.006]	[0.270]	[0.005]	[0.008]	[0.010]	[0.008]	[0.003]
Online and Face-to-Face Coaching	-0.007	-0.423	-0.006	-0.05	-0.039	-0.018	-0.013
	[0.037]	[1.532]	[0.029]	[0.045]	[0.056]	[0.047]	[0.019]
Control Mean [& st.dev.]	0.129 [0.336]	69.17 [13.4]	0.928 [0.258]	0.801 [0.4]	0.531 [0.5]	0.212 [0.409]	0.027 [0.162]
Sample Size	19,864	17,102	17,102	17,102	17,102	17,102	17,102

Notes: The table shows coefficient estimates from regressing the indicated outcome variable on the different treatment categories plus fixed effects for each randomized group listed in Table 1. Grades are measured as a percent at the end of the fall term averaged over all courses completed in the first year of each experiment. Grade>X is an indicator variable for whether the Fall Term Grade exceeds X. Control means, standard deviations and sample sizes are also shown at the bottom. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level respectively.

Table 5
Estimated Treatment Effects on Academic Performance and Persistence

	Outcome									
(1)	(2) **Fall Grade Year 1**	(3) **Winter Grade Year 1**	(4) **Credits Earned Year 1**	(5) **Final Grade Year 1**	(6) **Persisted Year 2**	(7) **Credits Earned Year 2**	(8) **Final Grade Year 2**	(9) **Persisted Year 3**	(10) **Credits Earned Year 3**	(11) **Final Grade Year 3**
Online Coaching Only	0.125	0.802	0.003	0.416	0.01	0.022	0.457	0.016	0.155	0.337
	[0.340]	[0.397]**	[0.034]	[0.341]	[0.009]	[0.038]	[0.350]	[0.010]	[0.060]**	[0.507]
Online and One-Way Text Coaching	0.255	0.282	0.006	0.085	-0.013	0.064	-0.029	-0.011	0.135	0.372
	[0.385]	[0.447]	[0.037]	[0.378]	[0.010]	[0.042]	[0.383]	[0.011]	[0.054]**	[0.455]
Online and Two-Way Text Coaching	-0.057	-0.362	-0.02	-0.167	-0.002	-0.023	-0.137	-0.004	Not available yet	
	[0.270]	[0.432]	[0.037]	[0.375]	[0.010]	[0.061]	[0.561]	[0.016]		
Online and Face-to-Face Coaching	-0.423	0.459	0.004	0.089	0.042	0.208	1.092	0.049	1.068	4.265
	[1.532]	[1.783]	[0.152]	[1.541]	[0.041]	[0.168]	[1.537]	[0.046]	[0.328]***	[2.740]
Control Mean [& st.dev.]	69.2 [13.4]	68.7 [15.3]	3.6 [1.4]	67.6 [13.8]	0.8 [0.4]	2.9 [1.8]	69.8 [12.8]	0.745 [0.436]	3.1 [1.7]	71.6 [12.1]

Notes: The table shows coefficient estimates from regressing the indicated outcome variable on the different treatment categories plus fixed effects for each randomized group listed in Table 1. The year indicates the year since the experiment began. Control means and standard deviations are also shown at the bottom. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level respectively.

Table 6
Estimated Treatment Effect on Fall Term Grade by Propensity Score Groups for Low Fall Grade

	Sample						
(1)	(2) **Full Sample [with HS grade]**	(3) **pscore> 90 pctile**	(4) **pscore> 80 pctile**	(5) **pscore> 70 pctile**	(6) **pscore> 60 pctile**	(7) **pscore> 50 pctile**	(8) **pscore<= 50 pctile**
Online Coaching Only	-0.146	-1.493	-1.277	-1.221	-1.403	-1.042	0.957
	[0.397]	[1.248]	[0.869]	[0.711]*	[0.609]**	[0.540]*	[0.533]*
Online and One-Way Text Coaching	0.007	-0.983	0.213	-0.053	-0.071	0.082	-0.182
	[0.443]	[1.200]	[0.851]	[0.699]	[0.613]	[0.556]	[0.678]
Online and Two-Way Text Coaching	-0.096	-1.949	-0.707	-0.934	-1.271	-0.856	0.292
	[0.317]	[1.379]	[0.946]	[0.741]	[0.619]**	[0.534]	[0.347]
Online and Face-to-Face Coaching	-0.201	-6.785	-0.059	0.458	0.719	1.822	-5.798
	[1.720]	[4.337]	[3.086]	[2.519]	[2.216]	[2.019]	[3.169]*
Control Mean [& st.dev.]	68.2 [13.6]	60.7 [14.6]	61.6 [14.4]	62.2 [14.3]	63.1 [14.2]	63.8 [14.1]	72.5 [12.0]
Sample Size	12,907	1,237	2,469	3,734	5,014	6,297	6,610

Notes: The table shows coefficient estimates from regressing Fall Grades (in percent) from the experimental year on the different treatment categories plus fixed effects for each randomized group listed in Table 1. Except for Column 2, the samples include only those with non-missing high school grade data. Regression results are shown using different samples, restricted by the indicated percentile cut-offs of a propensity score for the likelihood of receiving a low grade (less than 60) based on background characteristics. See text for more details on the calculation of this score. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level respectively.

Table 7
Estimated Treatment Effect Outcomes on Reported Mental Health and College Experience

	Outcome Coefficients [standard errors in brackets]					
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome	**Online Coaching Only**	**Online and One-Way Text Coaching**	**Online and Two-Way Text Coaching**	**Online and/or Text Coaching**	**Online and Face-to-Face Coaching**	**Sample Size**
Completed Follow-up Survey (Cont.Mn=0.76)	-0.015 [0.010]	0.01 [0.016]	-0.01 [0.007]		-0.034 [0.039]	11,446
Subjective Mental Health Outcomes (standardized)						
Life Satisfaction	0.08 [0.038]**	0.078 [0.062]	0.042 [0.026]		0.317 [0.153]**	8,140
Univ. Satisfaction	0.021 [0.030]	0.066 [0.049]	0.023 [0.021]		0.208 [0.121]*	8,140
Feeling Less Stressed	0.013 [0.024]	0.029 [0.039]	0.017 [0.016]		0.136 [0.096]	8,140
Feeling Less Depressed	0.011 [0.041]	0.081 [0.062]	0.034 [0.044]		0.166 [0.153]	4,342
Overall Mental Health	0.041 [0.031]	0.085 [0.050]*	0.037 [0.021]*		0.279 [0.125]**	8,140
Overall Mental Health				0.044 [0.018]**	0.258 [0.122]**	8,140
Subjective Feelings of Support (standardized)						
Sense of Belonging	-0.061 [0.041]	0.103 [0.062]*	0.016 [0.045]		0.278 [0.153]*	4,276
University Wants Me to Succeed	0.056 [0.042]	0.06 [0.062]	0.09 [0.045]**		0.431 [0.154]***	4,276
University Supports Me	0.044 [0.042]	0.057 [0.062]	0.093 [0.045]**		0.278 [0.153]*	4,276
Confident I Can Succeed	0.047 [0.041]	0.048 [0.062]	0.009 [0.045]		0.288 [0.152]*	4,276
Overall Sense of Support	0.029 [0.041]	0.089 [0.062]	0.069 [0.045]		0.425 [0.152]***	4,276
Overall Sense of Support				0.057 [0.032]*	0.409 [0.150]***	4,276

Notes: The table shows coefficient estimates from regressing the indicated standardized outcome variable (with mean zero, standard deviation one) on the different treatment categories plus fixed effects for each randomized group listed in Table 1. Except for the first row, the sample is restricted to those responding to the follow-up surveys taken near or after the end of the first year fall term. See text for more details. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level respectively.

Table 8
Estimated Treatment Effect Outcomes on Reported Study Behavior

	Outcome Coefficients [standard errors in brackets]					
(1)	(2) **Online Coaching Only**	(3) **Online and One-Way Text Coaching**	(4) **Online and Two-Way Text Coaching**	(5) **Online and/or Text Coaching**	(6) **Online and Face-to-Face Coaching**	(7) Sample Size
Study Behavior (standardized)						
Average Weekly Study Hours	0.067 [0.035]*	0.031 [0.053]	0.113 [0.023]***		0.006 [0.142]	9,662
Tend Not to Cram for Exams	0.038 [0.023]	0.003 [0.026]	0.054 [0.020]***		0.004 [0.107]	9,662
Number of Missed Classes	-0.205 [0.307]	-0.137 [0.319]	-0.135 [0.032]***		NA	3,995
Review Past Mistakes to Learn	0.021 [0.040]	-0.031 [0.056]	0.044 [0.042]		0.224 [0.152]	4,830
Rewrite Material in Own Words	0.013 [0.040]	-0.113 [0.056]**	-0.009 [0.042]		0.264 [0.152]*	4,830
Get Writing Feedback	0.002 [0.040]	0.029 [0.056]	0.054 [0.042]		0.287 [0.152]*	4,830
Meet with Tutor	-0.006 [0.040]	-0.149 [0.056]***	0.067 [0.042]		0.161 [0.152]	4,830
Manage Time Well	0.064 [0.037]*	-0.001 [0.056]	0.074 [0.025]***		0.332 [0.151]**	8,770
Overall Positive Study Behavior	0.06 [0.036]*	-0.051 [0.054]	0.127 [0.023]***		0.19 [0.145]	9,718
Overall Positive Study Behavior				0.091 [0.020]***	0.263 [0.143]*	9,718
Study Time (from 2018-19 data)	**Online Coaching Only**	**Online and One-Way Text Coaching**	**Online and Two-Way Text Coaching**	**Online and/or Text Coaching**	**Online and Face-to-Face Coaching**	Cont. Mn [Std Dev]
Weekly Study Time Fall Sem.			2.28 [0.409]***			14.4 [12.7]
Econ Weekly Study Time Fall Sem.			0.99 [0.172]***			5.0 [3.7]
Time Diary Daily Alone Study Time Fall Sem.			0.26 [0.085]***			2.6 [2.5]
Time Diary Daily Group Study Time Fall Sem.			0.06 [0.052]			0.7 [1.6]
Time Diary Daily Total Study Time Fall Sem.			0.32 [0.092]***			3.3 [2.7]
Current Weekly Study Time in Winter Sem.			2.18 [0.518]***			14.0 [10.6]

Notes: The table shows coefficient estimates from regressing the indicated standardized outcome variable (with mean zero, standard deviation one) on the different treatment categories plus fixed effects for each randomized group listed in Table 1. The sample is restricted to those responding to the follow-up surveys taken near or after the end of the first year fall term. See text for more details. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level respectively.

Table 9
Characteristics that Relate to Study Effort and Grade Outcomes in 2018-19 and 2019-20 Cohorts

(1)	(2)	(3)	(4)	(5)	(6)	(7)
		2018-19 Cohort			**2019-20 Cohort**	
Dep. Var.: Fall Semester Average Weekly Study Time						
	Mean [Std. Dev]			**Mean [Std. Dev]**		
Alpha (exp. grade attainable with minimal cramming)+	54.15	-0.132		44.27	-0.521	
	[12.28]	[0.055]**		[22.35]	[0.020]***	
Beta (slope between grade and weekly study time)+	1.72	0.628		1.1	-12.844	
	[0.69]	[0.939]		[0.51]	[0.874]***	
Expected Weekly Hrs Needed to Get an A+	14.94		0.171	30.56		0.255
	[7.46]		[0.080]**	[15.52]		[0.013]***
Extra Weekly Hrs Willing to Study to Guarantee A vs. B	8.32	0.217	0.205	9.78	0.177	0.2
	[7.25]	[0.073]***	[0.073]***	[7.39]	[0.025]***	[0.028]***
Like to Study (1 - 6 scale)	3.56	1.21	1.293	3.56	1.703	1.868
	[1.19]	[0.442]***	[0.444]***	[1.21]	[0.148]***	[0.166]***
High School Grade	87.08	0.31	0.298	88.35	0.255	0.238
	[6.67]	[0.086]***	[0.086]***	[5.89]	[0.031]***	[0.035]***
Dependent Variable Mean and Standard Deviation	18.92			16.42		
	[12.75]			[11.99]		
Observations		525	520		2844	2844
R-squared		0.077	0.064		0.353	0.19
Dep. Var.: End of Fall Term Average Grade						
	Mean [Std. Dev]			**Mean [Std. Dev]**		
Alpha (exp. grade attainable with minimal cramming)+		0.341			0.193	
		[0.044]***			[0.021]***	
Beta (slope between grade and weekly study time)+		1.698			6.779	
		[0.756]**			[0.933]***	
Expected Weekly Hrs Needed to Get an A+			-0.48			-0.115
			[0.064]***			[0.013]***
Extra Weekly Hrs Willing to Study to Guarantee A vs. B		0.15	0.151		0.108	0.113
		[0.058]**	[0.059]**		[0.026]***	[0.026]***
Hrs of Weekly Procrastination (Planned Winter - Actual Fall)	4.9	-0.024	-0.037	4.52	-0.118	-0.111
	[12.97]	[0.030]	[0.031]	[9.41]	[0.021]***	[0.021]***
Like to Study (1 - 6 scale)		0.142	0.062		0.75	0.744
		[0.351]	[0.355]		[0.158]***	[0.158]***
High School Grade		0.532	0.531		0.717	0.716
		[0.069]***	[0.071]***		[0.033]***	[0.033]***
Dependent Variable Mean and Standard Deviation	69.72			69.26		
	[13.65]			[13.49]		
Observations		493	488		2746	2,746
R-squared		0.251	0.234		0.2	0.197

Notes: Columns 3, 4, 6, and 7 show coefficient estimates from regressing the indicated dependent variable (Dep. Var.) on the independent variables listed in the table rows. The sample in columns 2 to 4 is restricted to those responding to the 2018-19 follow-up survey taken near or after the end of the first year fall term. The sample in columns 5 to 7 is restricted to those responding to the 2019-20 follow-up survey taken near the end of the first year fall term. exp = expected. Hrs = Hours. Wkly = Weekly. std. = standard deviation. + In the 2018-19 data, variables pertain to the realationship between weekly study time and grades in economics; in the 2019-20 data, they pertain to the relationhsip ebtween weekly study time and grades across all courses. See text for more details. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level respectively.

Table 10
Information Updating Revisions in Study Times and Grade Expectations in 2018-19 Cohort

	(1)	(2)	(5)	(6)	(7)	(8)
	Actual - Expected Study Time in Economics		Difference in Expected Econ Grades: At Follow-up – At Baseline		Actual Econ Grade - Expected Econ Grade at Baseline	
Change in Academic Savvy	0.122***	0.129***	-0.987***	-0.858***	-0.716***	-0.602***
	[0.019]	[0.019]	[0.072]	[0.083]	[0.072]	[0.073]
Observations	1,765	1,241	1,765	1,241	915	661
Controls?	N	Y	N	Y	N	Y

Notes : Each regression is estimated at the student level and the dependent variable indicated in the column headings. Control variables include high school admissions grade average, age, expected weekly study time across all courses reported during the baseline survey, expected weekly study time in economics reported at during the baseline survey, the number of days it took for the student to start the online warmup exercise, campus fixed effects, commute time to campus (in minutes), cubic functions of students' initially expected economics grade, initially expected weekly study time in economics, and initially expected study time across all courses, indicators for expected performance categories, English as a second language, gender, first-year status, first-generation status, international student status, intending to earn more than a BA, self-reported enjoyment of studying, frequent use of a calendar, believing the first midterm in a course determines subsequent outcomes, the belief that grades do not matter as long as one graduates, managing time well, and having a strong tendency to study at the last minute. Robust standard errors are reported in brackets. *** indicates significance at the 1 percent level; ** indicates significance at the 5 percent level; and * indicates significance at the 10 percent level.

Table 11
Information & Preference Updating Revisions in Study Times and Grade Expectations in 2019-20 Cohort

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Actual - Expected Study Time in All Courses		Difference in Expected Overall Grade: At Follow-up – At Baseline		Difference in Min. Acceptable Overall Grade: At Follow-up – At Baseline		Actual Overall Grade - Expected Overall Grade at Baseline	
Change in Academic Savvy	0.380***	0.320***	-0.256***	-0.230***	-0.149***	-0.129***	-0.197***	-0.132***
	[0.028]	[0.028]	[0.020]	[0.019]	[0.023]	[0.028]	[0.029]	[0.028]
Change in Preferences	0.090***	0.054	0.039	0.028	0.097***	0.062*	0.068**	0.013
	[0.035]	[0.034]	[0.025]	[0.027]	[0.029]	[0.034]	[0.033]	[0.033]
Mean Weekly Procrastination	-0.216***	-0.202***	-0.111***	-0.083**	-0.114**	-0.077	-0.216***	-0.140**
	[0.049]	[0.049]	[0.040]	[0.041]	[0.050]	[0.059]	[0.059]	[0.057]
Observations	1,741	1,354	1,741	1,354	1,709	1,329	1,621	1,309
Controls?	N	Y	N	Y	N	Y	N	Y

Notes: Each regression is estimated at the student level and the dependent variable indicated in the column headings. Control variables include high school admissions grade average, age, expected weekly study time across all courses reported during the baseline survey, expected weekly study time in economics reported at during the baseline survey, campus fixed effects, commute time to campus (in minutes), cubic functions of students' initially expected overall grade, initially expected weekly study time in economics, and initially expected study time across all courses, indicators for English as a second language, gender, first-year status, first-generation status, international student status, intending to earn more than a BA, self-reported enjoyment of studying, and having a strong tendency to study at the last minute. Robust standard errors are reported in brackets. *** indicates significance at the 1 percent level; ** indicates significance at the 5 percent level; and * indicates significance at the 10 percent level.

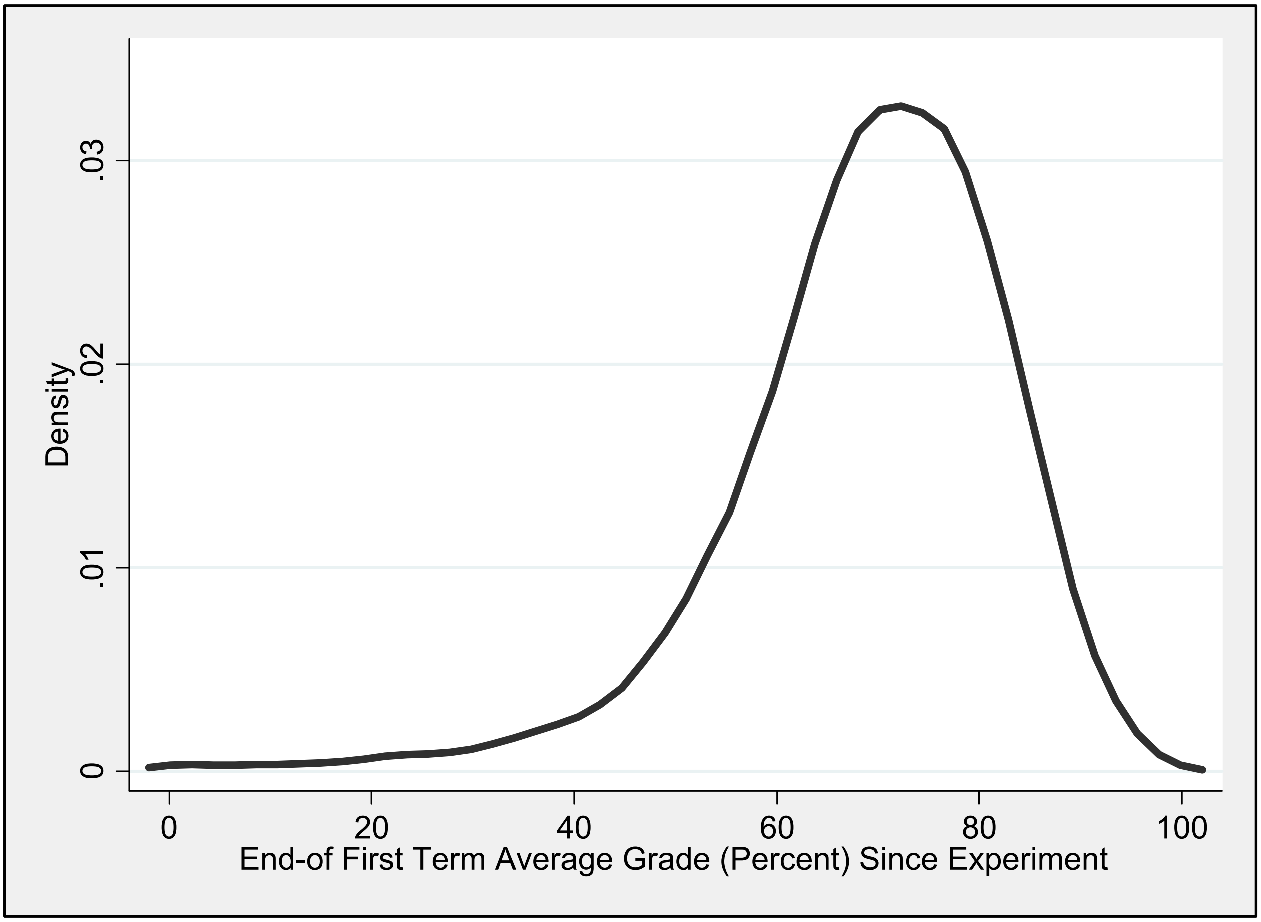
Table 12
Online Plus Two-Way Coaching Effects on Study Expectations, Ambition, and Procrastination

(1)	(2) Mean	(3) Std. Dev.	(4) Outcome Coef.	(5) Std. Error	(6) Sample Size
Beliefis about Academic Abilities					
Alpha (exp. grade attainable with minimal cramming)	54.75	[12.12]	-1.27	[0.567]**	1,833
Beta (slope between grade and weekly study time)	1.73	[0.69]	-0.02	[0.032]	1,833
Expected Weekly Hrs Needed to Get an A in Econ.	14.5	[7.37]	0.87	[0.347]**	1,826
Preferences for High Grades					
Extra Weekly Hrs Willing to Study to Guarantee A vs. B	7.82	[7.13]	1.03	[0.348]***	1,740
Extra Weekly Hrs Willing to Study to Guarantee A+ vs. B+	9.52	[8.62]	1.33	[0.409]***	1,740
Grades Don't Matter So Long As I Graduate (1-7 scale)	2.57	[1.29]	-0.11	[0.041]**	3,784
Procrastination					
Procrastination (Winteter Target Hrs - Fall Actual Hrs)	4.94	[12.94]	-0.05	[0.597]	1,892
Procrastination for Students with Low Initial Target Hrs	4.88	[12.17]	0.11	[0.836]	993
Procrastination for Students with High Initial Target Hrs	4.99	[13.64]	-0.23	[0.856]	899
Social Media, Screens Distract Me (standardized)	0	[0.84]	-0.04	[0.059]	808

Notes: The table reports estimated treatment effects from online and two-way coaching for the time-management program tested during the 2018-19 academic year. Sample sizes vary because some outcomes are collected from different surveys with different response rates (not correlated with treatment), and some variables were asked to a random subset only. exp = expected, hrs = hours. The social media variable is the average of standardized students' responses to their subjective agreement to the degree to which social media and video distract them. Students with low (high) initial target hours are those with stated target weekly study hours below (equal or above) the median (18 hours). *** indicates significance at the 1 percent level; ** indicates significance at the 5 percent level; * indicates significance at the 10 percent level.

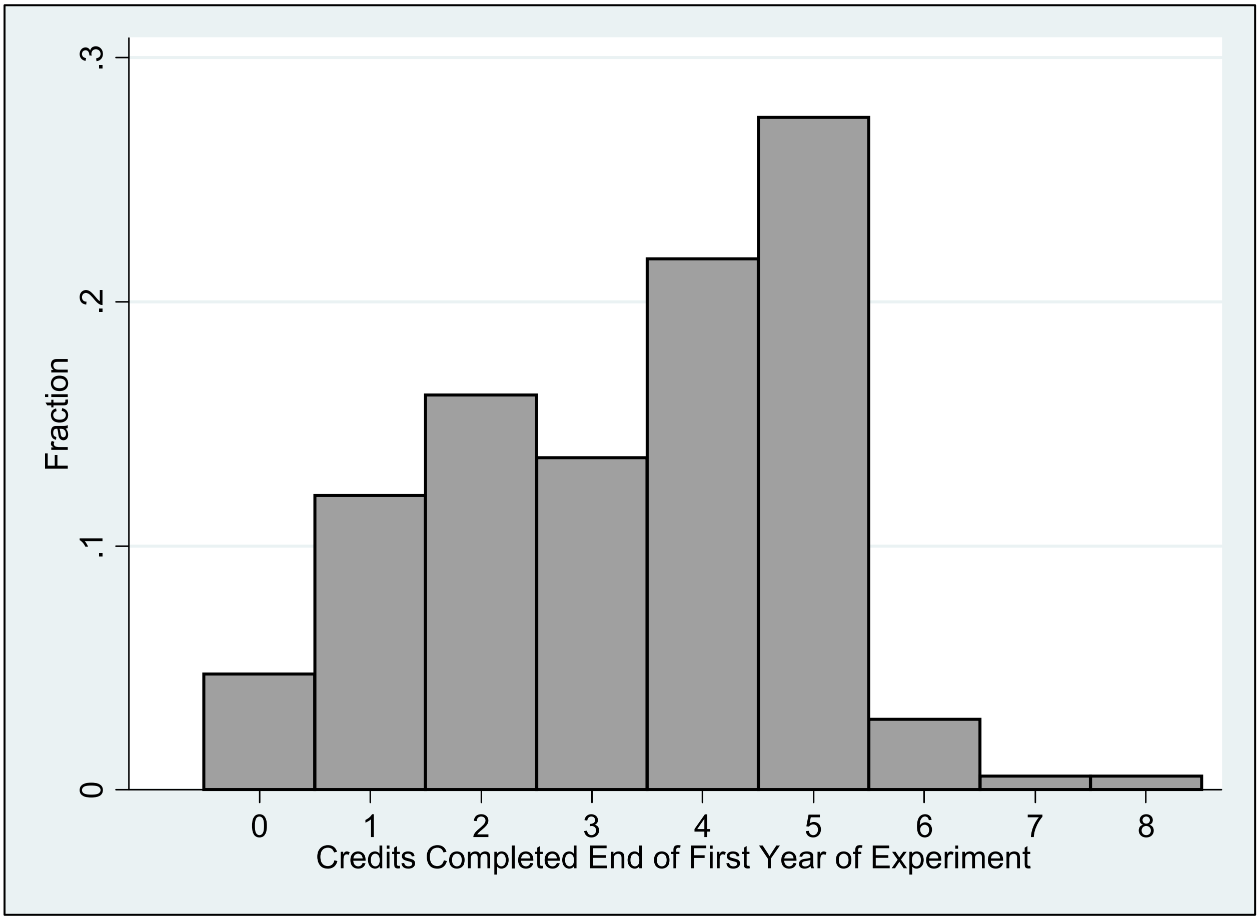
Figure 1

Fall Term Grade Distribution



Notes: Figure 1 graphs the kernel density estimate of all first year fall term grade averages for this paper's main sample of 2014-2018 first-year economics students. The density was calculated using a bandwidth of 2 and STATA's kdensity command. The median grade is 70.5, the 25th percentile is 62.0.

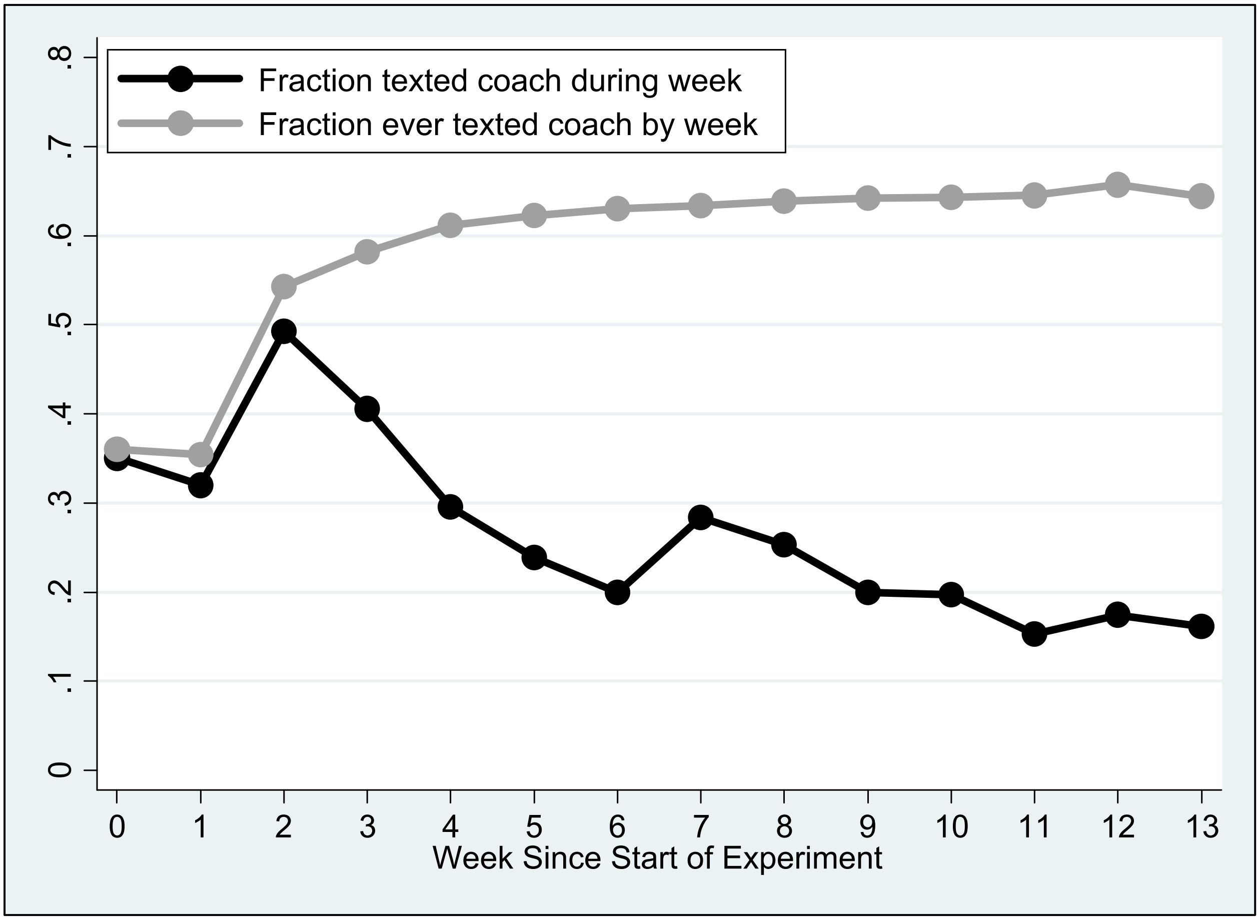
Figure 2
Total Credits Completed by End of First Year of Experiment (Sept-Aug)



Notes: Figure 1 displays the histogram of total credits completed by the end of the first year of the experiment. A full course load to graduate in four years with summers off would typically be 5 credits. The sample includes all first-year economics students in this paper's main sample (2014-2018).

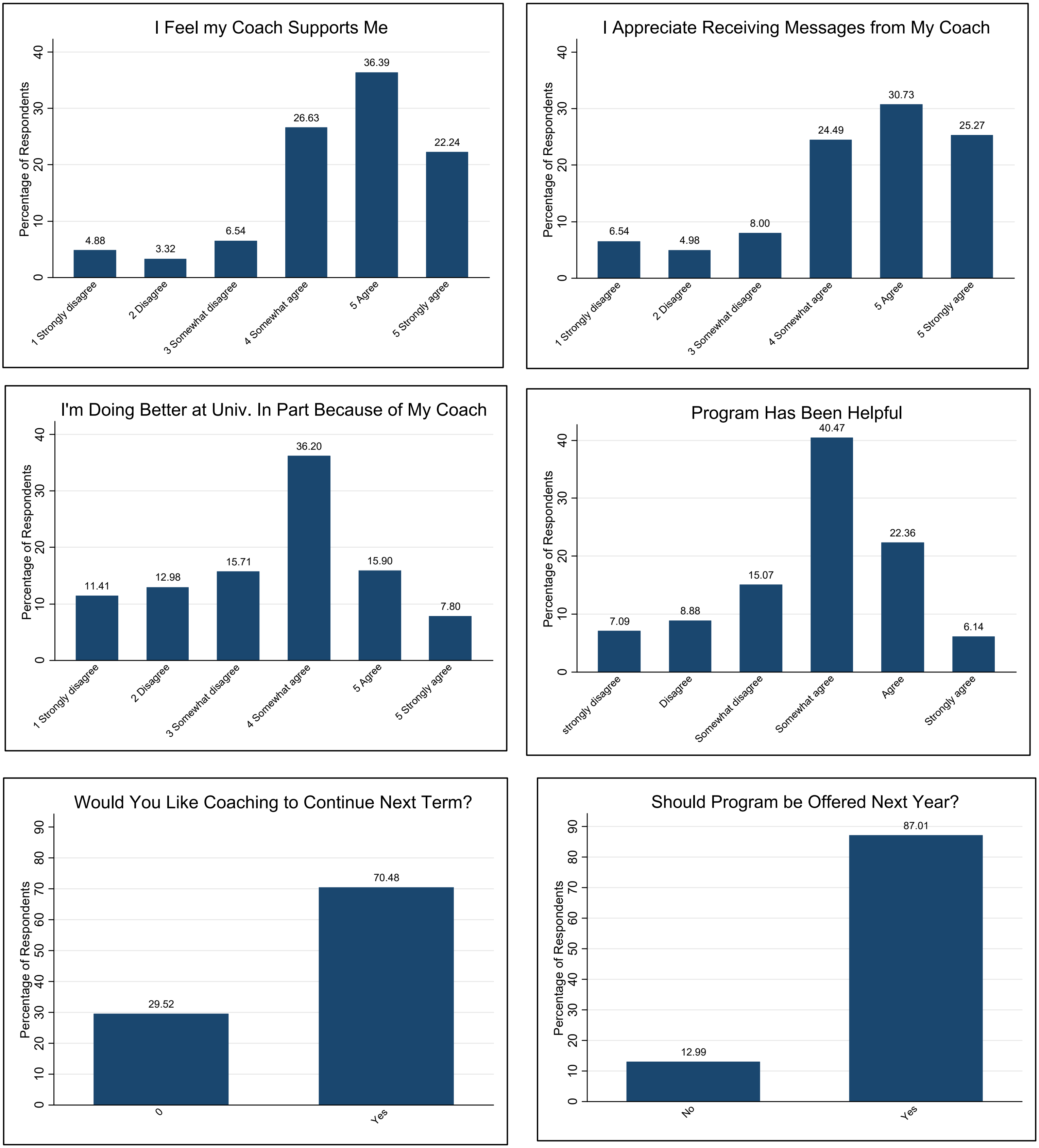
Figure 3

Fraction of Students Assigned to a Virtual Coach That Texted Back in a Given Week Since Start of Experiment And Fraction Ever Texted Back



Notes: The sample includes students agreeing to receive text-message coaching with Two-Way communication at the start of the 2016, 2017, and 2018 school years. The lighter line displays the fraction of this sample who ever texted back as of the indicated week during the first fall term of the experiment (with zero being the first Sunday after September 1). The darker line displays the fraction of this sample who texted anything back in a given week.

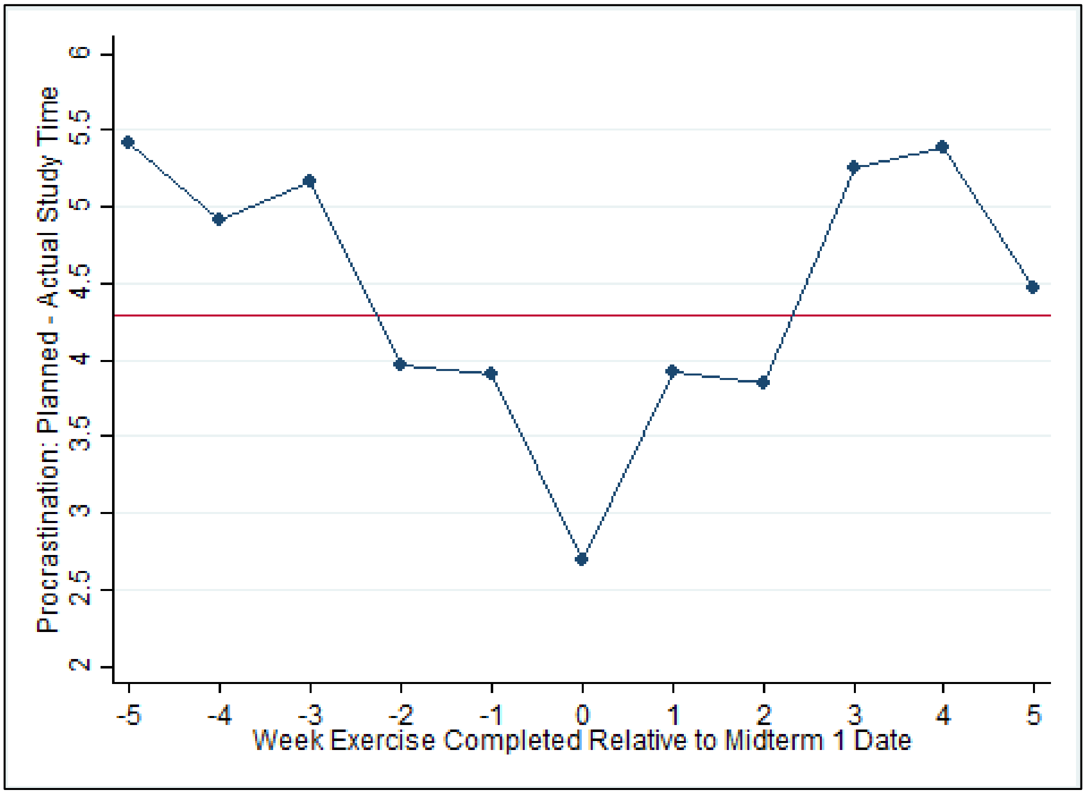
Figure 4: Student Feelings About the 2-Way Text-Message Coaching Program



Notes: The first three panels show the percentages of students in the text-message coaching program in 2016 who strongly disagree, disagree, somewhat disagree, somewhat agree, agree, and strongly agree with the statement that appears as the title of each panel. The last three panels show students responding about the 2018 coaching program.

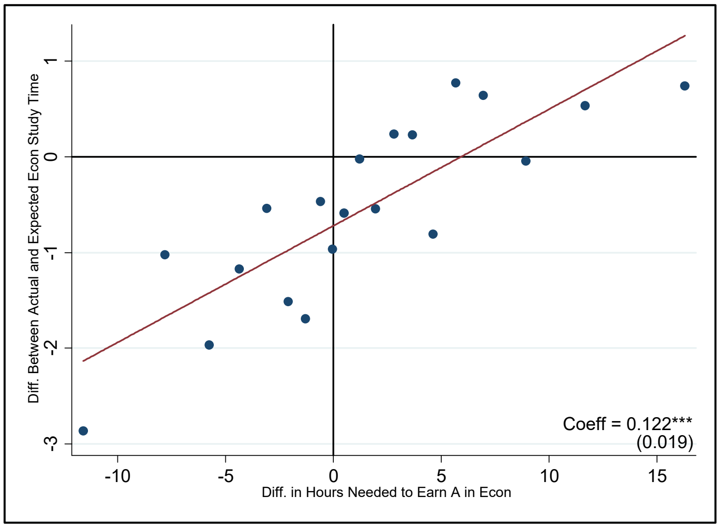
Figure 5

Planned minus Actual Reported Weekly Study Time

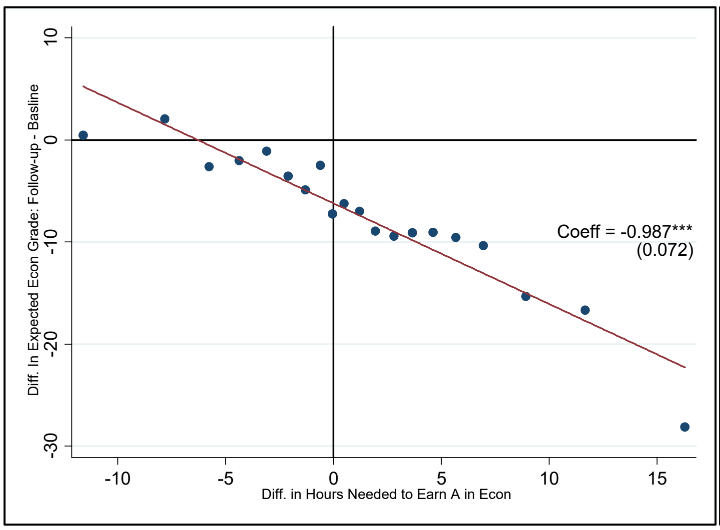


Notes: The sample is restricted to students responding to the 2019-20 weekly surveys about study times and grade expectations. The horizontal axis measures the week of the fall semester in which students completed each exercise relative to the week in which their first midterm in economics was held. The vertical axis measures the difference between the number of hours students planned to study in each week and the number of hours they actually studied in that week. Each point in the graph represents the average of this difference taken across all students who completed an exercise in that week. The horizontal solid red line represents the mean of y-axis variable taken across all student-week observations.

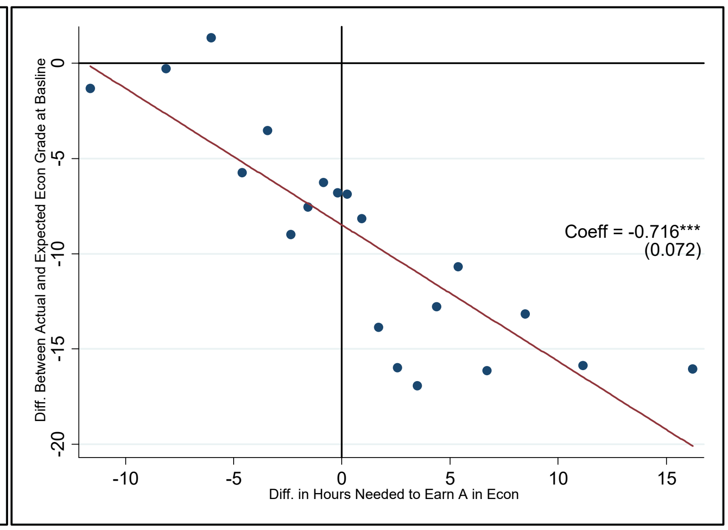
Figure 6: Study Time and Grade Expectation Revisions and Information Updating (2018-19 Cohort)



(a): Change in Econ Study Hours vs. Change in Hours for A



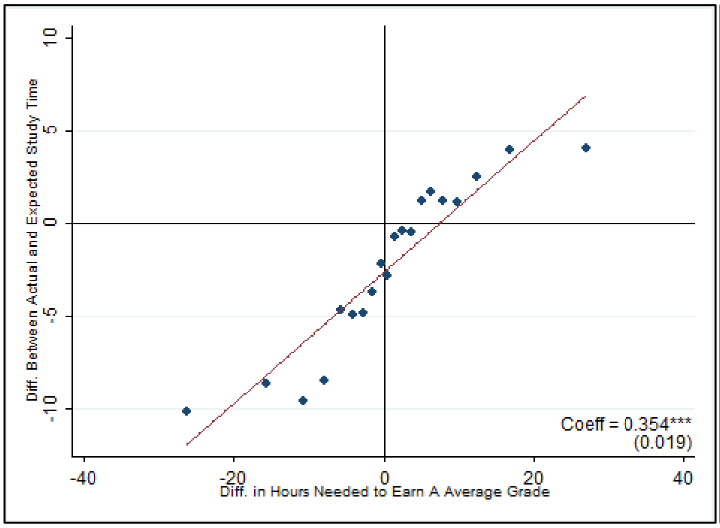
(b): Change in Econ Grade Expectation vs. Change in Hours for A



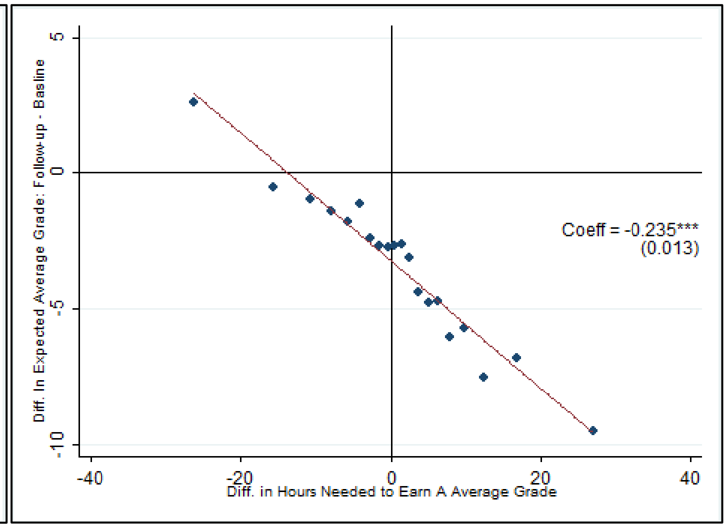
(c): Actual – Expected Econ Grade vs. Change in Hours for A

Notes: Panel (a) shows the relationship between changes in students' study times and measures of changes in students' beliefs about their academic abilities. Panels (b) and (c) show the relationships between changes in students' expected and realized economics grades and measures of changes in students' beliefs about their academic abilities. Each binned scatter plot is created by first grouping students into 20 equal-width bins (vigintiles) in the distribution of the variable on the x-axis and calculating the mean of both the y- and x-axis variables within each bin. The circles represent these means, while the lines represent the associated linear fits from the underlying student-level data.

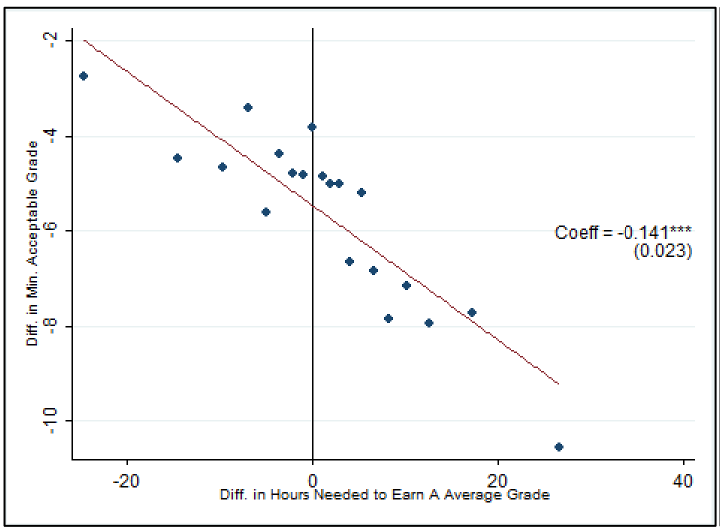
Figure 7: Study Time and Grade Expectation Revisions and Information Updating (2019-20)



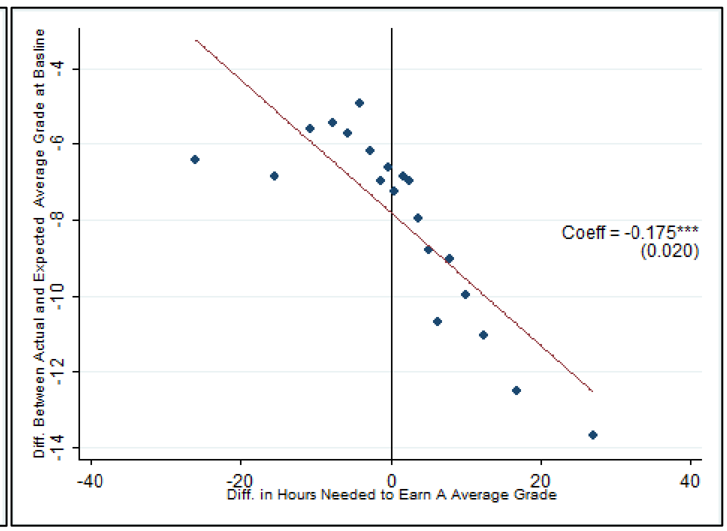
(a): Change in Study Hours vs. Change in Hours for A



(b): Change in Expected Grade vs. Change in Hours for A



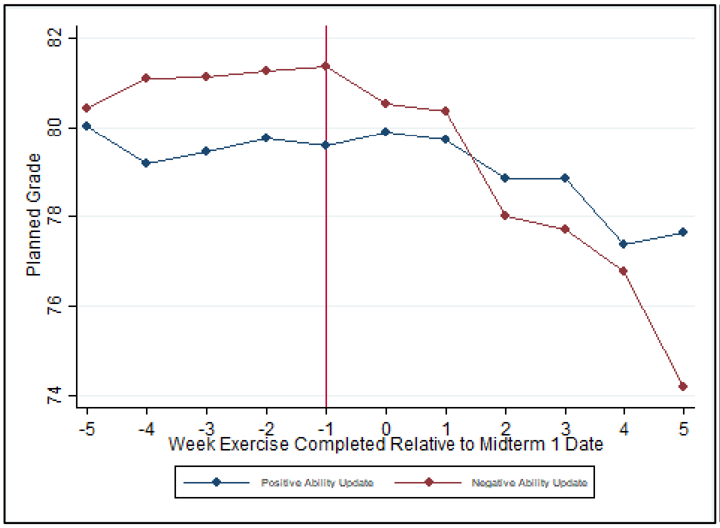
(c): Change in Min. Acceptable Grade vs. Change in Hours for A



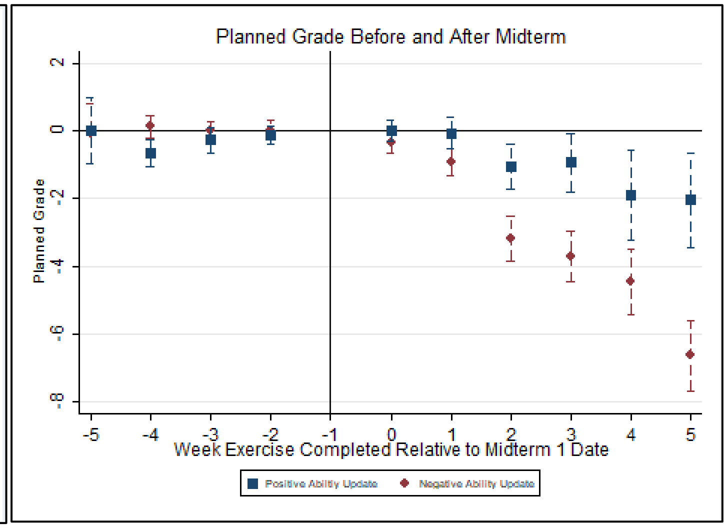
(d): Actual – Expected Average Grade vs. Change in Hours for A

Notes: Panel (a) shows the relationship between changes in students' study times and changes in students' beliefs about their academic abilities. Panel (b) shows the relationship between changes in students' grade expectations and changes in students' beliefs about their academic abilities. Panel (c) shows the relationship between changes in students' minimum acceptable grade average and changes in students' beliefs about their academic abilities. Panel (d) show the relationship between the difference in students' realized and expected average fall semester grade and changes in students' beliefs about their academic abilities. Each binned scatter plot is created by first grouping students into 20 equal-width bins (vigintiles) in the distribution of the variable on the x-axis and calculating the mean of both the y- and x-axis variables within each bin. The circles represent these means, while the lines represent the associated linear fits from the underlying student-level data.

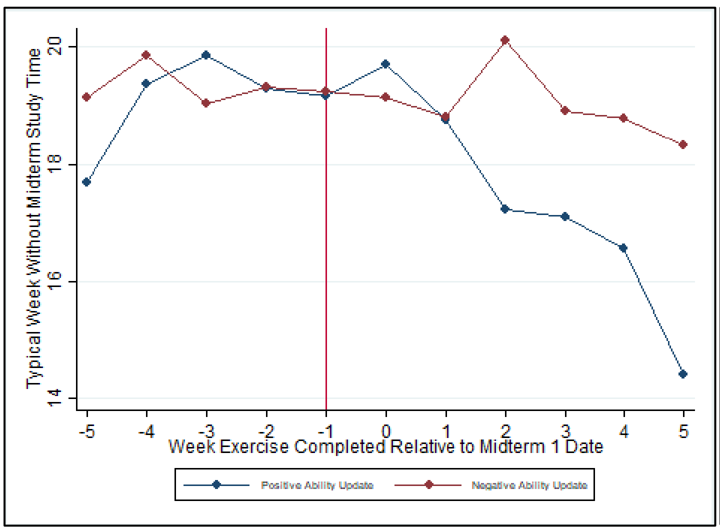
Figure 8: Grade and Weekly Study Time Expectations Over the Semester



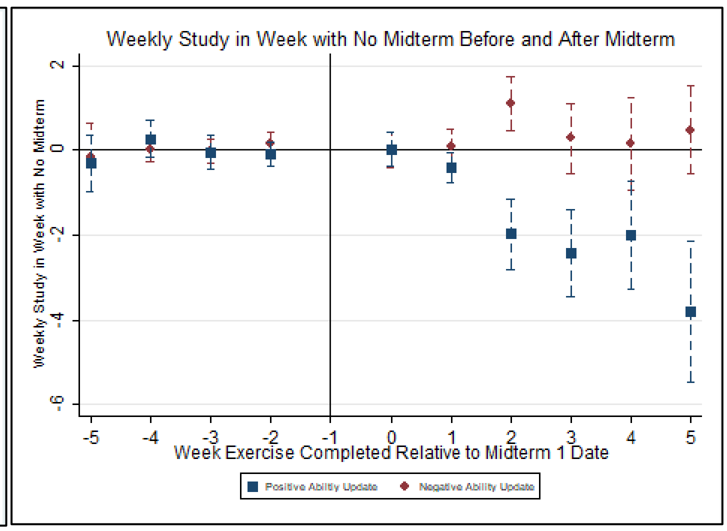
(a): Expected Overall Average Grade



(b): Expected Overall Average Grade—Reg Adjusted



(c): Planned Weekly Study Time Weeks without Midterms



(d): Planned Weekly Study Time Weeks without Midterms—Reg Adjusted

Notes: The sample is restricted to students responding to the 2019-20 weekly surveys about study times and grade expectations. In each panel, the horizontal axis measures the week of the fall semester in which students completed each exercise relative to the week in which their first midterm in economics was held. In panels (a) and (b), the vertical axis measures students' planned grade average over all their courses, while in panels (c) and (d), the vertical axis measures students planned weekly study time across all their courses. In panels (a) and (c), we plot the average of the y-axis variable across all students who completed an exercise in each week. In panels (b) and (d), we plot regression-adjusted means of the y-axis variable in each week from regressions that included student fixed effects, along with the 95-percent confidence interval for each mean.

Table A1
Estimated Treatment Effects on Initial Fall Term Grades [0-100], With Additional Control Variables

	Outcome Variable						
(1)	(2) **Missing Fall Grade**	(3) **Fall Term Grade**	(4) **Grade>50**	(5) **Grade>60**	(6) **Grade>70**	(7) **Grade>80**	(8) **Grade>90**
Online Coaching Only	0.003	0.060	-0.008	0.014	0.013	-0.002	0.002
	[0.007]	[0.319]	[0.006]	[0.010]	[0.012]	[0.010]	[0.004]
Online and One-Way Text	0.013	0.193	0.006	0.014	0.007	-0.002	0.002
Coaching	[0.008]	[0.362]	[0.007]	[0.011]	[0.014]	[0.011]	[0.005]
Online and Two-Way Text	-0.003	-0.096	0.000	-0.001	0.000	0.004	-0.001
Coaching	[0.006]	[0.253]	[0.005]	[0.008]	[0.009]	[0.008]	[0.003]
Online and Face-to-Face	-0.018	-0.399	-0.006	-0.051	-0.034	-0.010	-0.013
Coaching	[0.032]	[1.442]	[0.029]	[0.044]	[0.054]	[0.045]	[0.019]
Control Mean [& st.dev.]	0.129 [0.336]	69.2 [13.4]	0.928 [0.258]	0.801 [0.4]	0.531 [0.5]	0.212 [0.409]	0.027 [0.162]
Sample Size	18,885	17,102	17,102	17,102	17,102	17,102	17,102

Notes: The table shows coefficient estimates from regressing the indicated outcome variable on the different treatment categories plus fixed effects for each randomized group listed in Table 1. The regressions also include the following conditional variables: a set of cubic polynomial terms for father and mother's education and for age; indicator variables for: English as a second language, any parent with more than an undergraduate degree, high school admissions grade, interacted with whether high school grade is missing from the administrative data. Grades are measured as a percent at the end of the fall term averaged over all courses completed in the first year of each experiment. Grade>X is an indicator variable for whether the Fall Term Grade average exceeds X. Control means, standard deviations and sample sizes are also shown at the bottom. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level respectively.

Table A2
Estimated Treatment Effects on Initial Fall Term Grades [0-100]
by Campus

(1)	(2) **All UofT**	(3) **St. George Campus**	(4) **Mississauga Campus**	(5) **Scarborough Campus**
Online Coaching Only	0.125 [0.340]	-0.609 [0.432]	1.225 [0.990]	0.973 [0.636]
Online and One-Way Text Coaching	0.255 [0.385]	-0.053 [0.589]	0.815 [0.615]	-0.748 [0.946]
Online and Two-Way Text Coaching	-0.057 [0.270]	-0.315 [0.331]	0.007 [0.548]	0.423 [0.805]
Online and Face-to-Face Coaching	-0.423 [1.532]		0.719 [1.620]	
Control Mean [& st.dev.]	69.2 [13.4]	72.1 [12.8]	65.2 [13.4]	67.8 [13.0]
Sample Size	17,102	8,937	5,016	3,149

Notes: Same as in Table 4.

Table A3
Estimated Treatment Effects on Initial Full Year Math Grades [0-100]

	Outcome Variable						
(1)	(2) **Missing Yr1 Math Grade**	(3) **Year 1 Math Grade**	(4) **Grade>50**	(5) **Grade>60**	(6) **Grade>70**	(7) **Grade>80**	(8) **Grade>90**
Online Coaching Only	-0.011	1.281	0.004	0.029	0.043	0.028	0.009
	[0.011]	[0.518]**	[0.010]	[0.013]**	[0.014]***	[0.012]**	[0.008]
Online and One-Way Text Coaching	-0.004	-0.108	-0.011	-0.014	0.013	0.017	-0.003
	[0.013]	[0.604]	[0.012]	[0.015]	[0.017]	[0.015]	[0.009]
Online and Two-Way Text Coaching	-0.008	0.327	0.003	-0.003	0.014	0.007	-0.001
	[0.009]	[0.428]	[0.008]	[0.011]	[0.012]	[0.010]	[0.006]
Online and Face-to-Face Coaching	-0.021	2.708	-0.020	0.114	0.152	0.151	0.013
	[0.052]	[2.590]	[0.050]	[0.066]*	[0.072]**	[0.062]**	[0.039]
Control Mean [& st.dev.]	0.384 [0.486]	65.9 [18]	0.861 [0.346]	0.69 [0.463]	0.462 [0.5]	0.235 [0.424]	0.076 [0.265]
Sample Size	19,864	12,333	12,333	12,333	12,333	12,333	12,333

Notes: Same as in Table 4, but outcome is course average only for math courses taken over first year of experiment.

Table A4
Estimated Treatment Effects on Initial Full Year Economics Grades [0-100]

	Outcome Variable						
(1)	(2) **Missing Yr1 Econ Grade**	(3) **Year 1 Econ Grade**	(4) **Grade>50**	(5) **Grade>60**	(6) **Grade>70**	(7) **Grade>80**	(8) **Grade>90**
Online Coaching Only	0.010	0.680	0.004	0.014	0.021	0.021	0.023
	[0.010]	[0.429]	[0.008]	[0.012]	[0.013]	[0.011]*	[0.011]**
Online and One-Way Text Coaching	0.010	0.164	0.003	-0.003	0.009	0.004	0.016
	[0.011]	[0.494]	[0.009]	[0.013]	[0.015]	[0.013]	[0.013]
Online and Two-Way Text Coaching	0.015	0.027	0.004	0.000	-0.005	0.002	0.016
	[0.008]*	[0.348]	[0.007]	[0.009]	[0.011]	[0.009]	[0.009]*
Online and Face-to-Face Coaching	0.002	-0.560	-0.024	-0.006	-0.002	0.023	0.040
	[0.045]	[2.119]	[0.040]	[0.058]	[0.065]	[0.056]	[0.055]
Control Mean [& st.dev.]	0.224 [0.417]	67 [16.1]	0.892 [0.31]	0.734 [0.442]	0.5 [0.5]	0.233 [0.423]	0.162 [0.368]
Sample Size	19,864	15,241	15,241	15,241	15,241	15,241	15,241

Notes: Same as in Table 4, but outcome is course average only for economics courses taken over first year of experiment.

Table A5
Treatment Effects on Academic Performance and Persistence
Estimated Separately for All SAL Experiments

(1)	(2) **Fall Grade Year 1**	(3) **Winter Grade Year 1**	(4) **Final Grade Year 1**	(5) **Credits Earned Year 1**	(6) **Persisted Year 2**	(7) **Final Grade Year 2**	(8) **Credits Earned Year 2**
Online General Coaching without Follow-Up	-0.145	0.466	0.122	-0.059	0.013	0.318	0.041
	[0.534]	[0.611]	[0.518]	[0.051]	[0.014]	[0.514]	[0.050]
Online General Coaching with Text Follow-Up	0.203	0.057	-0.175	-0.053	-0.019	-0.116	0.066
	[0.478]	[0.546]	[0.463]	[0.046]	[0.013]	[0.462]	[0.045]
Online General Coaching with F2F Follow-Up	2.086	4.338	2.533	-0.03	0.009	3.197	0.649
	[3.006]	[3.416]	[2.895]	[0.286]	[0.081]	[2.908]	[0.285]**
Customized Coaching without Follow-Up	0.457	1.272	0.807	0.044	-0.001	0.577	0.006
	[0.486]	[0.555]**	[0.485]*	[0.048]	[0.013]	[0.482]	[0.047]
Customized Coaching with Text Follow-Up	0.251	0.374	0.492	0.017	-0.015	-0.076	-0.031
	[0.593]	[0.672]	[0.585]	[0.058]	[0.016]	[0.580]	[0.056]
Customized Coaching with F2F Follow-Up	-1.436	-0.87	-0.697	0.06	0.067	0.323	0.053
	[1.807]	[2.090]	[1.832]	[0.181]	[0.051]	[1.815]	[0.178]
Customized Coaching with For-Profit Text Follow-Up	0.04	0.536	0.531	0.12	0.013	0.057	0.08
	[0.760]	[0.884]	[0.755]	[0.074]	[0.021]	[0.781]	[0.076]
Time Management Coaching with Follow-Up	-0.101	-0.245	-0.315	-0.024	0.005	0.106	0.013
	[0.306]	[0.349]	[0.307]	[0.030]	[0.008]	[0.316]	[0.030]
Control Mean [& st.dev.]	69.2 [13.4]	68.7 [15.3]	67.6 [13.8]	3.6 [1.4]	0.8 [0.4]	69.8 [12.8]	2.9 [1.8]

Notes: Same as Table 5**.**

Figure A1
Screen Shot of 2016 Online Program

General Instructions

The University of Toronto and the Department of Economics want to better understand our students' thoughts about transitioning into university. We will use this information to evaluate the resources we plan to provide for future students.

This exercise involves 2 parts:

- In Part 1, **you will be asked to think about your own** education and future. This will help us understand how students think about various strategies for having a good year and working towards their goals.
- In Part 2, **you will be asked to tell us why you think other students** struggle and to suggest ways your peers might overcome challenges. This section is intended to help us understand how UofT can support future students to overcome barriers

The exercise should take about 45 to 90 minutes to complete. Please try your best to write for the amount of time specified and feel free to take longer if you need to. Please take your time and be thoughtful. If you need a few minutes to walk around and take a break, please feel free to do so.

You'll be asked to help us understand your thoughts and feelings about getting the most out of university.

At the end of the exercise, we will email a copy of your notes to your account address. Reflect on them at a later time, as you may have additional thoughts.

If you need to take a break or two to get up and walk around or help you think, please feel free to do so. Thank-you and Enjoy!

Proceed through the exercise by clicking the Next (Save) button. You can go back to previous pages by clicking Previous (Save). Each time you click Next or Previous, the data you have entered on that page will be saved.

Save Next (Save)

General Instructions

Part 1: How to Succeed at U of T
Study enough
Study effectively
Get help when you don't understand
Keep up and go to class
Stay motivated
Be patient and take a long-term perspective

Part II
Identifying the Barriers to Success
Digging Deeper Into the Top 2 Issues

One last thing...

Congratulations! You are finished

Figure A2
Screen Shot of 2016 Online Program

5) Staying Motivated

It's not always easy to study with so many other activities competing for students' time. Spending time with friends, watching videos, or even cleaning can seem preferable. Students can help stay committed to learning by frequently reminding themselves what motivates them.

- [ ] 1) For some, motivation comes from thinking about how their education can be used to help achieve their long-term career and family related goals.
- [ ] 2) For others, who may not have a clear sense of their long-term goals yet, it comes from wanting to keep their options open. Good grades often open doors to graduate school and help impress potential employers after graduation.
- [ ] 3) For others, it's about challenging themselves to do their best and focusing on learning as much as they can about how the world works.
- [ ] 4) Or, for others, motivation comes from the idea of using their education one day to help others and make a real difference in the world.

Click on the number above corresponding to what you think is the strongest source of motivation for doing well in school for most incoming UofT students.

Please tell us what motivates you to do well at UofT and why

0 word(s)

Previous (Save) | Save | Next (Save)

General Instructions

Part 1: How to Succeed at U of T
Study enough
Study effectively
Get help when you don't understand
Keep up and go to class
Stay motivated
Be patient and take a long-term perspective

Part II
Identifying the Barriers to Success
Digging Deeper Into the Top 2 Issues

One last thing...

Congratulations! You are finished

Figure A3:
Screen Shot of One-Way Coaching Manager

Figure A4
Screen Shot of Two-Way Coaching Manager

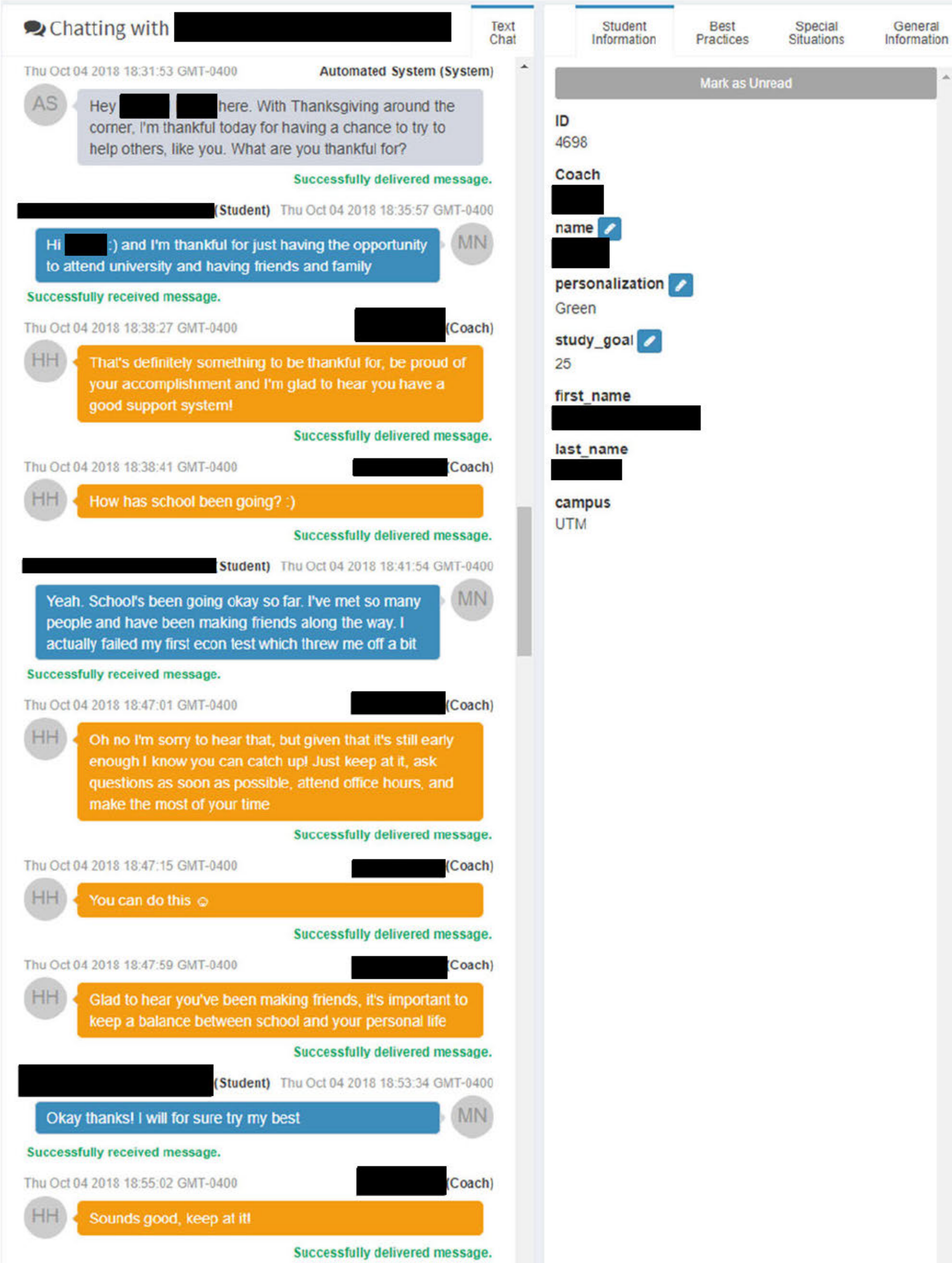


Figure A5
Screen Shot of Planning Treatment

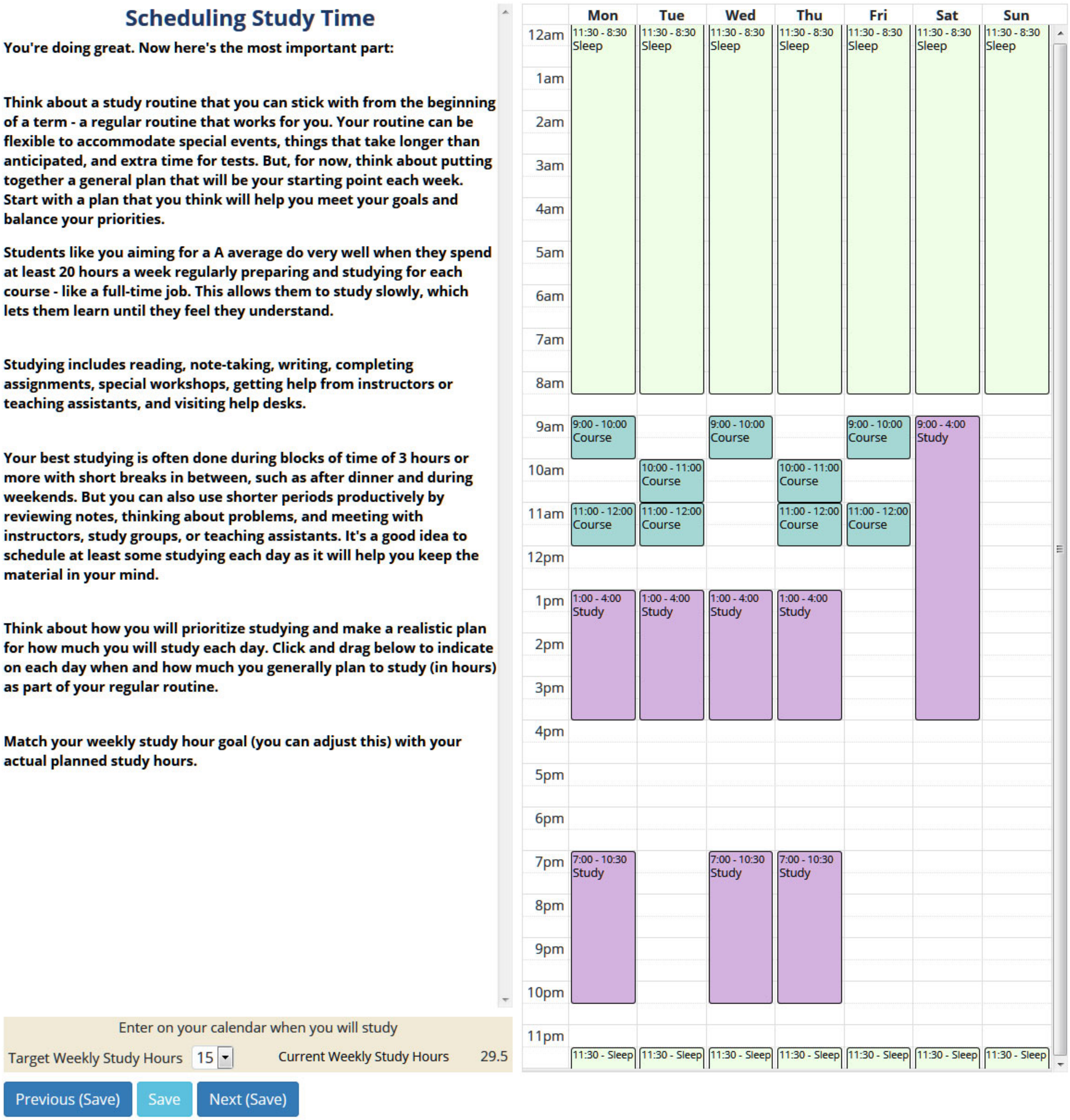
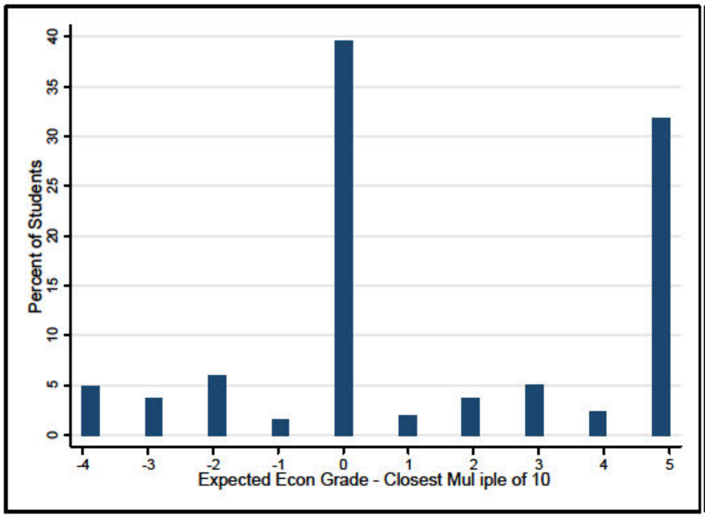
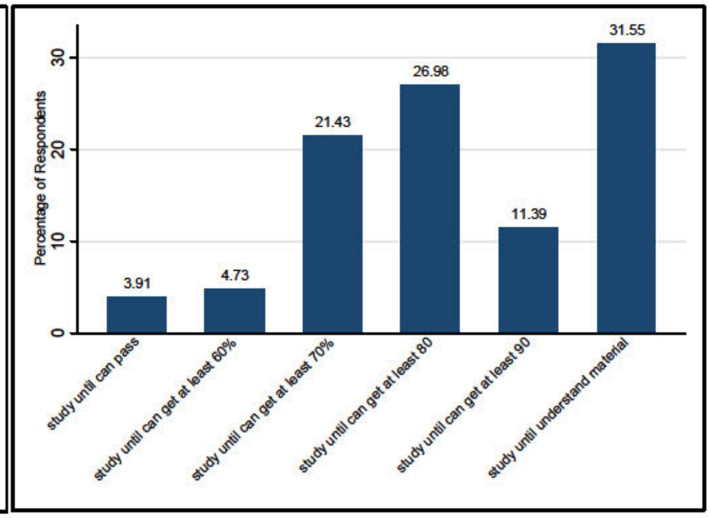


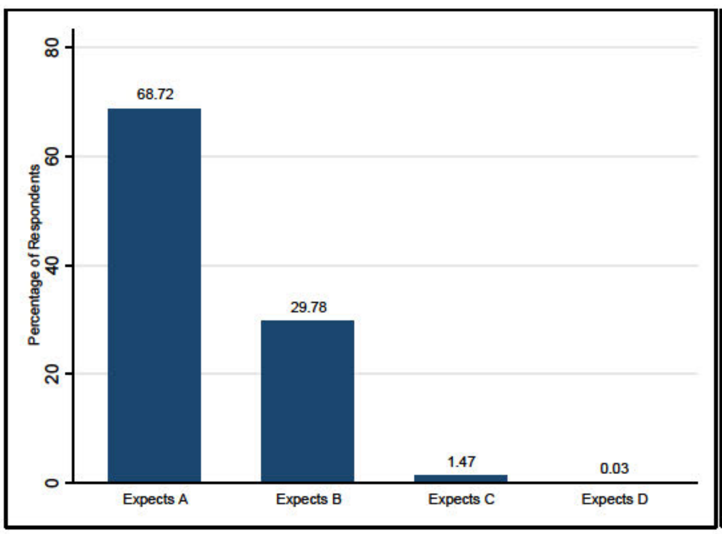
Figure A6: Supporting the Modelling Assumptions



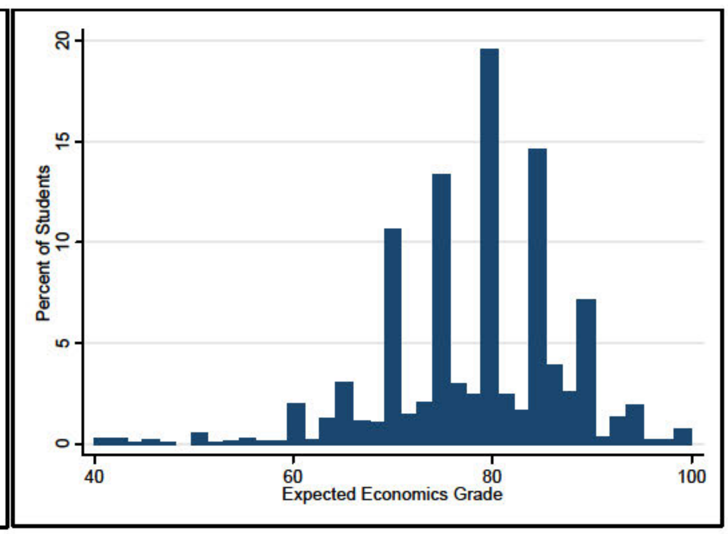
(a): Distance Between Expected Econ Grade and Multiple of Ten



(b): Test Preparation Strategies



(c): Grade Expectations Over All Courses



(d): Distribution of Expected Grades in Economics

Notes: Panel (a) shows the percentage of students whose expected economics grade (reported during the baseline survey) is a distance from a multiple of 10 that is indicated by the values on the horizontal axis. Panel (b) shows the percentage of students who report (during the follow-up survey) having each of the test preparation strategies listed on the horizontal axis. Panel (c) shows the percentage of students who expect to earn each letter grade on the horizontal axis in their economics course (reported during the baseline survey). Panel (d) shows the full distribution of expected economics grades (at baseline).

Appendix B: Model and Supporting Evidence

In this appendix, we first present evidence in support of our modelling assumptions in Section IV. We then present formal proofs for the propositions discussed in Section IV.

I. Motivating the Assumptions

In this subsection, we present evidence to support our modelling assumptions that students perceive the benefits of earning higher grades in categories and that they do not discriminate between grades that are equal to a letter grade of C or below.

Categorical Thinking

We first provide evidence that students think about the benefits of earning higher grades in discrete categories. Using survey data from the fifth year of experiment, panel (a) of Figure A.6 shows that nearly 40 percent of all students expect to earn an economics grade that is an exact multiple of ten—a far larger fraction than for expected grades at any other (integer) distance from the closest multiple of ten. At UofT, grade multiples of ten always indicate a change of letter grades, suggesting that students are bunching their grade expectations around clear letter grade cutoffs. Panel (b) of Figure A.6 presents direct evidence that students approach their studies by thinking about grade categories, showing the distribution of student test preparation strategies. Only 30 percent of students report studying until they completely understand the material, while the remaining 70 percent report studying only until they feel confident that they will earn a specific percentage grade that is a multiple of ten. Students' tendencies to think about their performance in grade categories is perhaps not surprising, given that most institutions (UofT included) produce transcripts that report letter grade performance (or GPA categories) for each course—measures that do not vary continuously with students' underlying percentage grades and only change when these grades cross specified thresholds.

Grouping All Grades Up To And Below a C Into One Category

Panel (b) of Figure A.6 shows that only about 9 percent of students approach preparing for a test by studying enough to earn only a C or below; 5 percent of students study just enough to earn a C (60 percent average grade) and 4 percent study just enough to pass. Panel (c) shows that only 1.5 percent of students expect to earn a grade of C or below across all their courses at the start of the semester, while Panel (d) shows the full distribution of expected percentage grades in economics, revealing that a very small mass of students expect to earn a C or less (60 percent or below). In summary, it appears that very few students expect to earn a grade below a B, and even among those who do, most do not expect to earn less than a C.

II. Proofs

In this section, we present formal proofs of the propositions made in Section IV in the main text. To begin, recall that the optimal study choice of student i in period t is given by equation (7) in the main text:

$$s_{it}^{*} = \begin{cases} s_{it}^{A} \text{ if } \theta_{it}^{A} - \theta_{it}^{B} \geq c\left(s_{it}^{A}\right) - c(s_{it}^{B}) \\ s_{it}^{B} \text{ if } \theta_{it}^{A} - \theta_{it}^{B} < c\left(s_{it}^{A}\right) - c(s_{it}^{B}) \end{cases},$$

where $s_{it}^{j} = \frac{y^{j} - \hat{\alpha}_{it}}{\hat{\beta}_{it}}$ for $j = A$ and B. We first establish that the RHS of (7) is decreasing in both $\hat{\alpha}_{it}$ and $\hat{\beta}_{it}$.

Lemma 1: Define $k\left(\hat{\alpha}_{it}, \hat{\beta}_{it}\right) = c\left(s_{it}^{A}\right) - c(s_{it}^{B})$. $k\left(\hat{\alpha}_{it}, \hat{\beta}_{it}\right)$ is decreasing in both $\hat{\alpha}_{it}$ and $\hat{\beta}_{it}$.

Proof. Taking the partial derivative of $k(\hat{\alpha}_{it}, \hat{\beta}_{it})$ with respect to each object and noting that $c(\cdot)$ is strictly increasing and convex gives the desired result.

We now present a proof for Proposition 1 in the main text, establishing how the behavior of students who are originally aiming for an A changes as they learn new information.

Proposition 1: *Suppose student i is originally studying enough to expect to earn a letter grade of A. Hold fixed the difference between the perceived benefit of earning an A and the benefit of earning a B, $\theta_{it}^{A} - \theta_{it}^{B}$. If student i receives a positive update about her academic ability (α_i) or return to studying (β_i), she continues aiming for an A but with less study effort. If she receives a small negative update, she continues aiming for an A but with more study effort; if she receives an intermediate negative update, she lowers her expected grade to a B but decreases or does not change study effort; if she receives a large negative update, she lowers her expected grade to a B and increases study effort.*

Proof. Suppose student i is studying enough to expect to earn an A at time 0 such that $\theta_{it}^{A} - \theta_{it}^{B} \geq k(\hat{\alpha}_{i0}, \hat{\beta}_{i0})$ in equation (7).

Case 1: Suppose student i receives a positive information shock, such that $\hat{\alpha}_{i1} > \hat{\alpha}_{i0}$ or $\hat{\beta}_{i1} > \hat{\beta}_{i0}$. Because $k(\hat{\alpha}_{i0}, \hat{\beta}_{i0})$ is decreasing in both objects, the RHS of (7) falls, ensuring the inequality remains satisfied. The student responds by continuing to study enough to expect an *A* but reduces study time, such that $s_{i1}^{*} = s_{i1}^{A} < s_{i0}^{*} = s_{i0}^{A}$.

Case 2: Suppose student i receives a negative information shock, such that $\hat{\alpha}_{i1} < \hat{\alpha}_{i0}$ or $\hat{\beta}_{i1} < \hat{\beta}_{i0}$. Because $k(\hat{\alpha}_{i0}, \hat{\beta}_{i0})$ is decreasing in both objects, the RHS of (7) increases. For the remainder of the proof, we consider a decrease in $\hat{\alpha}_i$, such that $\hat{\alpha}_{i1} < \hat{\alpha}_{i0}$, assuming that $\hat{\beta}_{i1} = \hat{\beta}_{i0}$. (Following analogous steps would establish the results when $\hat{\beta}_{i1} < \hat{\beta}_{i0}$ and $\hat{\alpha}_{i1} = \hat{\alpha}_{i0}$.) Let $\hat{\alpha}_i^*(\hat{\beta}_{it}, \theta_{it}^A, \theta_{it}^B)$ denote the value of $\hat{\alpha}_i$ that, for a given return to studying and benefits to grades, ensures that equation (7) is satisfied as an equality – that is, student i is indifferent between studying enough to earn an A or B (in which case we assume she aims for an *A*). Let the student i's new belief over her academic ability be $\hat{\alpha}_{i1} = \hat{\alpha}_{i0} - \Delta\hat{\alpha}_i$ for some $\Delta\hat{\alpha}_i > 0$.

Case 2(i): Suppose the change in $\hat{\alpha}_i$ is relatively small, such that $\hat{\alpha}_{i1} = \hat{\alpha}_{i0} - \Delta\hat{\alpha}_i > \hat{\alpha}_i^*(\hat{\beta}_{it}, \theta_{it}^A, \theta_{it}^B)$. Then it is still that case that $\theta_{it}^A - \theta_{it}^B > k(\hat{\alpha}_{i1}, \hat{\beta}_{i1})$. Student i continues aiming for an *A* but increases study effort, such that $s_{i1}^* = s_{i1}^A > s_{i0}^* = s_{i0}^A$.

Case 2(ii): Suppose the change in $\hat{\alpha}_i$ is such that $\hat{\alpha}_{i1} = \hat{\alpha}_{i0} - \Delta\hat{\alpha}_i < \hat{\alpha}_i^*(\hat{\beta}_{it}, \theta_{it}^A, \theta_{it}^B)$ but that the downward revision $\Delta\hat{\alpha}_i$ is not too big, such that $\Delta\hat{\alpha}_i \leq y^A - y^B$. In this case, because $\theta_{it}^A - \theta_{it}^B < k(\hat{\alpha}_{i1}, \hat{\beta}_{i1})$, student i switches to aiming for a *B* but either reduces or does not change study time. The change in study time is given by $s_{i1}^* - s_{i0}^* = \frac{y^B - y^A - (\hat{\alpha}_{i1} - \hat{\alpha}_{i0})}{\hat{\beta}_{i0}}$, which is negative when $\Delta\hat{\alpha}_i = \hat{\alpha}_{i0} - \hat{\alpha}_{i1} < y^A - y^B$ and zero when $\Delta\hat{\alpha}_i = \hat{\alpha}_{i0} - \hat{\alpha}_{i1} = y^A - y^B$.

Case 2(iii): Suppose the change in $\hat{\alpha}_i$ is such that $\hat{\alpha}_{i1} = \hat{\alpha}_{i0} - \Delta\hat{\alpha}_i < \hat{\alpha}_i^*(\hat{\beta}_{it}, \theta_{it}^A, \theta_{it}^B)$ and that the downward revision $\Delta\hat{\alpha}_i$ is large, such that $\Delta\hat{\alpha}_i > y^A - y^B$. In this case, because $\theta_{it}^A - \theta_{it}^B < k(\hat{\alpha}_{i1}, \hat{\beta}_{i1})$, student i switches to aiming for a *B* but increases study time. The change in study time is given $s_{i1}^* - s_{i0}^* = \frac{y^B - y^A - (\hat{\alpha}_{i1} - \hat{\alpha}_{i0})}{\hat{\beta}_{i0}}$, which is positive because the downward revision to beliefs about academic ability is sufficiently large, such that $\Delta\hat{\alpha}_i = \hat{\alpha}_{i0} - \hat{\alpha}_{i1} > y^A - y^B$.

We now present a proof for Proposition 2 in the main text, establishing how the behavior of students who are originally aiming for a B changes as they learn new information.

Proposition 2: *Suppose student i is originally studying enough to expect to earn a letter grade of B. Hold fixed the difference between the perceived benefit of earning an A and the benefit of earning a B, $\theta_{it}^{A} - \theta_{it}^{B}$. If student i receives a negative update about her academic ability (α_i) or return to studying (β_i), she continues aiming for a B but with more study effort. If she receives a small positive update, she continues aiming for a B but with less study effort; if she receives an intermediate positive update, she increases her expected grade to an A and increases or does not change study effort; if she receives a large positive update, she raises her expected grade to an A but decreases study effort.*

Proof. Suppose student i is studying enough only to expect to earn a B at time 0 such that $\theta_i^A - \theta_i^B < k(\hat{\alpha}_{i0}, \hat{\beta}_{i0})$ in equation (7).

Case 1: Suppose student i receives a negative information shock, such that $\hat{\alpha}_{i1} < \hat{\alpha}_{i0}$ or $\hat{\beta}_{i1} < \hat{\beta}_{i0}$. Because $k(\hat{\alpha}_{i0}, \hat{\beta}_{i0})$ is decreasing in both objects, the RHS of (7) increases, ensuring the inequality remains satisfied. The student responds by continuing to study enough to expect to earn only a *B* but increases study time, such that $s_{i1}^{*} = s_{i1}^{B} > s_{i0}^{*} = s_{i0}^{B}$.

Case 2: Suppose student i receives a positive information shock, such that $\hat{\alpha}_{i1} > \hat{\alpha}_{i0}$ or $\hat{\beta}_{i1} > \hat{\beta}_{i0}$. Because $k(\hat{\alpha}_{i0}, \hat{\beta}_{i0})$ is decreasing in both objects, the RHS of (7) decreases. For the remainder of the proof, we consider an increase in $\hat{\alpha}_i$, such that $\hat{\alpha}_{i1} > \hat{\alpha}_{i0}$, assuming that $\hat{\beta}_{i1} = \hat{\beta}_{i0}$. (Following analogous steps would establish the results when $\hat{\beta}_{i1} > \hat{\beta}_{i0}$ and $\hat{\alpha}_{i1} = \hat{\alpha}_{i0}$.) As above, let $\hat{\alpha}_i^{*}(\hat{\beta}_{it}, \theta_{it}^{A}, \theta_{it}^{B})$ denote the value of $\hat{\alpha}_i$ that, for a given return to studying and benefits to grades, ensures that equation (7) is satisfied as an equality and let the student i's new belief over their academic ability be $\hat{\alpha}_{i1} = \hat{\alpha}_{i0} + \Delta\hat{\alpha}_i$ for some $\Delta\hat{\alpha}_i > 0$.

Case 2(i): Suppose the change in $\hat{\alpha}_i$ is relatively small, such that $\hat{\alpha}_{i1} = \hat{\alpha}_{i0} + \Delta\hat{\alpha}_i < \hat{\alpha}_i^*(\hat{\beta}_{it}, \theta_i^A, \theta_i^B)$. Then it is still that case that $\theta_{it}^A - \theta_{it}^B < k(\hat{\alpha}_{i1}, \hat{\beta}_{i1})$. Student i continues aiming for a B but decreases study effort, such that $s_{i1}^* = s_{i1}^B < s_{i0}^* = s_{i0}^B$.

Case 2(ii): Suppose the change in $\hat{\alpha}_i$ is such that $\hat{\alpha}_{i1} = \hat{\alpha}_{i0} + \Delta\hat{\alpha}_i > \hat{\alpha}_i^*(\hat{\beta}_{it}, \theta_{it}^A, \theta_{it}^B)$ but that the upward revision $\Delta\hat{\alpha}_i$ is not too big, such that $\Delta\hat{\alpha}_i \leq y^A - y^B$. In this case, because $\theta_{it}^A - \theta_{it}^B > k(\hat{\alpha}_{i1}, \hat{\beta}_{i1})$, student i switches to aiming for an A and either increases or does not change study time. The change in study time is given by $s_{i1}^* - s_{i0}^* = \frac{y^A - y^B - (\hat{\alpha}_{i1} - \hat{\alpha}_{i0})}{\hat{\beta}_{i0}}$, which is positve when $\Delta\hat{\alpha}_i = \hat{\alpha}_{i0} - \hat{\alpha}_{i1} < y^A - y^B$ and zero when $\Delta\hat{\alpha}_i = \hat{\alpha}_{i0} - \hat{\alpha}_{i1} = y^A - y^B$.

Case 2(iii): Suppose the change in $\hat{\alpha}_i$ is such that $\hat{\alpha}_{i1} = \hat{\alpha}_{i0} + \Delta\hat{\alpha}_i > \hat{\alpha}_i^*(\hat{\beta}_{it}, \theta_{it}^A, \theta_{it}^B)$ and that the upward revision $\Delta\hat{\alpha}_i$ is large, such that $\Delta\hat{\alpha}_i > y^A - y^B$. In this case, because $\theta_{it}^A - \theta_{it}^B > k(\hat{\alpha}_{i1}, \hat{\beta}_{i1})$, student i switches to aiming for a A but decreases study time. The change in study time is given $s_{i1}^* - s_{i0}^* = \frac{y^A - y^B - (\hat{\alpha}_{i1} - \hat{\alpha}_{i0})}{\hat{\beta}_{i0}}$, which is negative because the upward revision to beliefs about academic ability is sufficiently large, such that $\Delta\hat{\alpha}_i = \hat{\alpha}_{i0} - \hat{\alpha}_{i1} > y^A - y^B$.

Proposition 3: *Holding* $\hat{\alpha}_i$ and $\hat{\beta}_i$ *fixed, the maximum amount of time a student is willing to study for an A is increasing in the difference between the perceived benefit of earning an A and the perceived benefit of earning a B,* $\theta_{it}^A - \theta_{it}^B$.

Proof. As above, let $\hat{\alpha}_i^*(\hat{\beta}_{it}, \theta_{it}^A, \theta_{it}^B)$ denote the value of $\hat{\alpha}_i$ that, for a given return to studying and benefits to grades, ensures that equation (7) is satisfied as an equality. Because the RHS of (7) $k(\hat{\alpha}_{it}, \hat{\beta}_{it})$ is decreasing in $\hat{\alpha}_{it}$, for a given return to studying $(\hat{\beta}_{it})$ and values of $\theta_{it}^A - \theta_{it}^B$, the maximum amount of study time that student i is willing to put forward to earn an A occurs at the level of $\hat{\alpha}_{it}$ when $\hat{\alpha}_{it} = \hat{\alpha}_{it}^*(\hat{\beta}_{it}, \theta_{it}^A, \theta_{it}^B)$ and is given by $s_{it}^A(\hat{\alpha}_{it}^*) = \frac{y^A - \hat{\alpha}_{it}^*}{\hat{\beta}_{it}}$. Because LHS of equation (7) is increasing in $\theta_{it}^A - \theta_{it}^B$, it follows that $\hat{\alpha}_{it}^*(\hat{\beta}_{it}, \theta_{it}^A, \theta_{it}^B)$ is decreasing in $\theta_{it}^A - \theta_{it}^B$, as higher values of $\theta_{it}^A - \theta_{it}^B$ allow equation (7) to hold as an equality for smaller values of $\hat{\alpha}_{it}$ (when the RHS, $k(\hat{\alpha}_{it}, \hat{\beta}_{it})$, is larger). Because the maximum amount one is willing to study for an A, $s_{it}^A(\hat{\alpha}_{it}^*)$, is decreasing in $\hat{\alpha}_{it}^*$, and $\hat{\alpha}_{it}^*$ is decreasing in $\theta_{it}^A - \theta_{it}^B$, it follows that $s_{it}^A(\hat{\alpha}_{it}^*)$ is increasing in $\theta_{it}^A - \theta_{it}^B$.