Improving Online Learning through Course Design: A Microeconomic Approach

Hammad Shaikh*
University of Toronto

December 12, 2021

Latest version of the paper can be found here.

Abstract

Online education has expanded dramatically over the past two decades, yet significant learning challenges remain. In light of these, my paper provides the first microeconomic analysis to examine how the quality of online university courses can be enhanced through course design. First, I gather rich data covering 3,700 undergraduates at a large public university taking an online introductory programming course that has a cumulative structure. The data allow me to monitor students' study time precisely and to characterize important dimensions of heterogeneity: student attentiveness and whether they are forward-looking. I then conduct two randomized interventions that nudge students to utilize an online discussion board more fully and to complete online assignments. I find that an additional 4.5 weeks of discussion board utilization increases final exam grades by 0.07 SD and completing one extra online assignment (out of 10 in total) raises final grades by 0.18 SD. I then develop and estimate a behavioural model of student effort supply, credibly identifying the marginal benefits and costs of effort at each stage of the cumulative learning process using the two field experiments. The estimated model allows me to explore the efficacy of changing assignment grading weights to improve student learning. In contrast to the actual (equally-weighted) grading scheme, simulated weights that maximize learning are decreasing across assignments, serving to increase effort by myopic students early in the course when they acquire foundational skills. My course-design approach is applicable more generally in other online and traditional course settings.

*I would like to thank Robert McMillan, Aloysius Siow, and Román Andrés Zárate for their guidance and support. Thanks also to Victor Aguirregabiria, Isaiah Andrews, Carolina Arteaga, Stephen Ayerst, Gabriel Carroll, Marc-Antoine Chatelain, Cheok In Fok, Jiaying Gu, Elaine Guo, Alexander Hempel, Tanjim Hossain, Faisal Ibrahim, Daniel Indacochea, Raji Jayaraman, Catherine Michaud Leclerc, Étienne Makdissi, David Price, Swapnika Rachapalli, Baxter Robinson, Marc-Antoine Schmidt, Eduardo Souza-Rodrigues, Eva Vivalt, Jessica Wagner, Matt Walshe, Joseph Jay Williams, Farhan Yahya, Ruizhi Zhu, Jiaqi Zou, and participants in the University of Toronto's Public Economics Discussion Group and Empirical Microeconomics Seminar for their comments. Sameul Khan, Sam Maldanado, and Qi Yin Zheng provided valuable data support, and anonymous course instructors helped me navigate the online learning platforms. The experiments and data collection are approved by the Research Ethics Board of the anonymized institution in this study. Financial support from an Ontario Graduate Scholarship is greatly appreciated. All remaining errors are my own. Contact: Department of Economics, University of Toronto, 150 St. George Street, Toronto, Ontario, Canada, M5S 3G7. Please send comments to hammy.shaikh@mail.utoronto.ca.

1 Introduction

Online education has expanded rapidly over the past two decades,[1] and dramatically so in response to the COVID-19 pandemic. As post-secondary institutions around the world were forced to implement distance-learning alternatives to in-person course delivery, the vast majority of students have taken online courses during the past academic year. On the positive side, online education brings with it low per-student costs and easy scalability (Deming et al., 2015). Yet significant learning challenges remain. In particular, many students who learn online find it difficult to obtain individualized support or stay engaged with online course work (Bowen, 2012); in turn, a lack of support and low engagement may lead students to fall permanently behind, especially in courses with a cumulative structure. Consistent with these difficulties, a large literature documents worse academic outcomes in online courses relative to traditional in-person education (Bettinger et al., 2017).[2]

In light of the challenges that come with online learning, this paper examines how the quality of online university courses can be enhanced through course design. It considers two prominent features of online courses. First, online peer discussion boards offer one appealing way of supporting students while providing education at scale, allowing students to discuss course concepts asynchronously. Second, online assignments are intended to reinforce student learning in the course; by setting grading incentives associated with online assignments appropriately, the instructor can help guide students to allocate effort in a way that increases their skill accumulation.

The approach I develop to assess the efficacy of these common features of online courses consists of the following components. At the outset, I gain access to unusually rich administrative data from a large pre-existing foundational online STEM course that has a cumulative learning structure (in common with many technical subjects); I supplement these data with dedicated student surveys. I then conduct two field experiments in the course: a randomized intervention that informs students about the course's online discussion board, and a randomized intervention sending homework reminders throughout the course. The exogenous variation in students' utilization that results from the first experiment allows me to establish the value-added of the peer learning environment; the exogenous variation in students' study time at each stage of the learning process arising from the second experiment allows me to estimate the cumulative learning technology. Next I develop and estimate a multi-stage behavioral model of student effort supply, credibly identifying the marginal

[1] The share of students enrolled in at least one online course in the US, for example, rose from 15% in 2004 to 43% by 2018 (US National Centre for Education Statistics).

[2] Coates et al. (2004) and Figlio et al. (2013) find students who study introductory economics online perform worse; Bueno (2020) documents lower college graduation for students who received primary education virtually.

benefits and costs of effort at each stage of the cumulative learning process using the two field experiments. Since conducting a further field experiment to randomize grading schemes across students is potentially unethical and logistically difficult, I then use the estimated model to simulate the effects of changing the profile of assignment grading weights counterfactually, exploring whether doing so improves student learning. I take these components of my approach in turn, starting with the context and data.

The specific setting for my analysis is a large online introductory programming course offered each 12-week semester at a research-intensive Canadian university. The course uses an open-source online learning platform where students learn content on their own by watching videos and doing practice problems posted on a weekly basis. In addition to weekly low-stakes homework assessments, the course also includes two high-stakes assessments: a midterm and a final exam. Given that students in the course learn most of the material through self-study, the course also employs an online student discussion board to further support students. The discussion board facilitates learning by allowing students to interact with each other, discussing the course material and collaborating on assignments in an instructor-moderated online environment. Utilization of the discussion board is completely voluntary, whereas the homework assessments are incentivized through their inclusion in a student's overall course grade.

I collect data on nearly 3,700 students who consented to participate in the research.[3] The rich administrative data include time-stamped student interactions with the online homework environment throughout the entire semester, and the week (if ever) when students register for the discussion board. These data enable me to measure total online study time at each stage of the learning process precisely. Evidence from the administrative data suggests a lack of online participation activity by a non-trivial proportion of students in the control groups of the field experiments. Each week, around $15-20\%$ of students spend no time whatsoever doing the homework. Furthermore, around 30% of students never sign up for the discussion board. I supplement the administrative data with survey data (as mentioned), collecting demographic information from students and further eliciting their behavioural characteristics, such as a their attentiveness and forward-looking (versus myopic) perspectives. I find that more attentive students have a higher propensity to register for the discussion board, and students who are more myopic tend to be less likely to do optional (ungraded) homework problems that are available throughout the course. The survey evidence also reveals substantial heterogeneity in effort allocation according to a student's behavioural 'type.'

To examine the efficacy of the online peer discussion board and online homework assignments,

[3]The student consent rate is around 87%.

I conduct the two field experiments introduced above. The interventions considered in this study can both be characterized as 'targeted informational reminders nudges,' as they prompt a student to take a specific action, provide simple instructions for doing so, and lastly, remind the student to complete the task. The first nudge the students are subjected to is an online sign-up activity which encourages students to register for the discussion board near the start of the course. The sign-up activity uses screenshots and an animated GIF to provide clear, step-by-step instructions to register for the discussion board, and informs students about its functionality. The second nudge is a homework reminder message which is deployed repeatedly across several weeks throughout the course and is aimed at promoting further participation in online homework. The reminder message informs students of the upcoming homework deadline, prompts them to set aside time in their schedule to work on the homework, and provides a direct link to the homework.

I find the sign-up activity is successful in encouraging registration early on in the course, increasing student utilization of the online discussion board by approximately 4.5 weeks. I then use the random assignment to the sign-up activity as an instrument for the number of weeks students are registered: an additional five weeks utilizing the discussion board increases homework and final exam achievement by 0.14 SD and 0.08 SD, respectively. I also use the survey data to investigate the mechanisms underlying this intervention, finding that the sign-up activity is highly effective in informing inattentive students about the online peer discussion platform – that is, students who are less aware of the discussion board and who are unfamiliar with its functionality at the outset of the semester.

The deployment of randomized homework reminders throughout the course provides an opportunity to estimate the cumulative learning technology as a function of students' study time allocation. I find that receiving an additional reminder message, on average, induces students to spend an extra 23 minutes on the corresponding homework assessment. The reminder messages are most useful for inattentive students, who are less likely to be aware of upcoming homework deadlines. By using random assignment to the number of homework reminders as an instrument for total homework study time, I examine the causal learning benefits from participation in low-stakes homework. Here I find that an additional hour spent studying course material through the online assignments increases final exam grades by 0.09 SD.

To further explore the student effort allocation process, I develop an estimable model of online learning. The model features a single instructor (i.e., the principal) and multiple students (i.e., the agents), the latter exhibiting heterogeneity in their baseline knowledge, English language proficiency, and whether they are forward-looking or not. Students allocate effort across three learning stages: basic, intermediate, and advanced. Myopic students set effort at each learning stage independently without internalizing that the productivity of studying in the future is increasing in current knowledge accumulation, while forward-looking students internalize the cumulative benefits when allocating their study time earlier in the course. The instructor's objective is to maximize the learning of a representative student net of effort costs, whereas students exert effort throughout the course to maximize their expected course grade. The solution to the students' problem indicates that myopic students misallocate their effort and underinvest in low-stakes homework assessments covering foundational programming skills. Moreover, the solution to the instructor's problem demonstrates that, when designing an optimal grading scheme for a course in the presence of a cumulative learning technology in which dynamic complementarities are strong, a myopic (forward-looking) student's learning is best served by assessments whose weights decrease (increase) throughout the course.

Next, I use the administrative dataset including students' precise study time allocation together with the two field experiments to estimate the model (also showing empirical evidence that the model implications just outlined are consistent with the data). The benefit of effort in a learning stage depends on the cumulative technology, which has two endogenous variables – the total study time in the current learning stage and the knowledge accumulated in the previous learning stage. I construct instruments for both endogenous inputs using the number of randomly assigned homework reminders at each learning stage, allowing me to identify the parameters of the cumulative learning technology. I find a positive marginal return to effort at each stage of the learning process and document evidence of dynamic complementarities. For example, in the advanced learning stage, I find an additional hour of online study time increases final exam grades by 0.11 SD, and this marginal benefit increases by 0.07 SD for every 1 SD increase in intermediate stage knowledge.

I then estimate the cost function using maximum likelihood estimation. Here I assume the cost function is convex and linearly separable across the learning stages. Considering each stage independently, the convexity parameter of the cost function can be identified. To see how, note that I already found the sign-up nudge, deployed at the start of the course, had persistent effects of increasing students' study time at each learning stage. The availability of exogenous variation in study time resulting from the sign-up activity across students who have the same marginal cost of effort in each learning stage then identifies the convexity of the cost function.[4] I estimate a convexity parameter of 2.1 in the basic stage, implying a cost function that is closely quadratic in effort. The convexity parameter estimate of the cost function gradually increases across stages of

[4]See Section 8.8 for details underlying the identification argument for the effort cost function.

the course, and is 2.5 by the advanced stage.

The estimated model enables me to project student effort allocation and corresponding learning outcomes as a function of the grading scheme implemented by the instructor. In a course with a cumulative structure and a large portion of myopic students, the estimated optimal assignment grading weights are gradually decreasing as the course progresses. Relative to the grading scheme used in the existing course with equal assignment weights, I find that implementing the optimal grading scheme is predicted to increase final exam performance for myopic students by 0.11 SD. The achievement gain arises from myopic students effectively allocating their study time by investing more effort at earlier learning stages of the course under the optimal weights, thereby obtaining proficiency in the basic concepts that serve as foundational building blocks for rest of the course.

Overall, the findings in this paper contribute to our understanding of how the quality of online university courses can be improved through course design, following the systematic approach I develop. I provide the first causal evidence showing that access to an online discussion board and weekly online assignments helps support learning at scale. The experimental variation in student engagement allows me to estimate and credibly identify a behavioural model of effort allocation which sheds new light on the optimal design of course grading schemes. For a course with a cumulative structure and many myopic students, the simulations indicate that instructors should assign more weight to assessments given earlier in the course. Doing so will lead myopic students to appropriately front-load their effort (as noted).

Although the analysis in this paper focuses on a single online STEM course, there are good reasons to think the results can inform the design of foundational university courses more broadly. Many students enrolled in first-year core courses with large class sizes are going to find it difficult to obtain individualized support. While online learning exacerbates the problem of supporting students as they can be in a different time zone from the instructor, an analogous issue arises in universities where many students commute, some over long distances, to attend classes and office hours. Additionally, many first-year university courses have a cumulative structure, especially in STEM, with considerable heterogeneity in the student body. As most popular learning management systems are equipped with an online discussion board and online homework hosting capabilities, these online facilities can be used to support student learning. Furthermore, most university courses have multiple assessments, and the instructor can distribute the grading incentives unevenly, as informed by the model simulations. In these ways, the microeconomic course-design approach considered in my paper can help alleviate widespread challenges prevalent in both online courses and traditional in-person foundational university courses.

The rest of the paper is organized as follows. The next section places my analysis in the context of the related literature. Section 3 provides information about the sample and describes the online homework and discussion board environments. Section 4 outlines sources of data collection and also presents descriptive statistics. The experimental design and the key features of the interventions are discussed in Section 5, and corresponding results are presented in Section 6. A model of student online learning participation and corresponding theoretical implications are set out in Section 7. I estimate the proposed model and discuss identification in Section 8, Section 9 presents counterfactual analyses using the estimated model, and Section 10 concludes.

2 Literature Review

This paper builds on several prior literatures. These include the structural behavioral economics literature, papers using structural models to inform the design of public policies, and research that estimates cumulative education production functions and uses field experiments to evaluate the efficacy of various educational interventions, among others. In this section, I discuss these areas of prior research in turn and highlight the ways in which my course-design approach contributes to each of them.

First, my paper relates to papers in structural behavioral economics that use experimental variation together with a model of decision making to estimate behavioural parameters. Allcott and Taubinsky (2015) conduct an experiment involving store customers to measure the extent to which customers are inattentive to the energy savings of LED light bulbs. The authors elicit willingness to pay for various light bulbs through an iPad survey and randomize a portion of the participants to receive information about the corresponding energy costs. They find modest evidence of inattention that would justify a subsidy of 3 dollars per LED light bulb. Taubinsky and Rees-Jones (2018) conduct an online shopping experiment in which consumers face various tax rates, and find substantial evidence of inattention to taxation. Augenblick (2018) measures the present bias of study participants who are paid to complete unpleasant tasks by asking them repeatedly on different days the number of blurry greek letter transcriptions they would like to complete on a given date. By randomizing wages and the number of days after the decision the tasks are carried, the author finds some evidence of present-bias exhibited by the study participants. A large fraction of the literature that identifies structural behavioural parameters uses experiments that are conducted in laboratory settings (Card et al., 2011). Stringent experimental configurations are typically required to credibly identify parameters such as the attentiveness or forward-looking perspective in a population (see Appendix A.1 for a more extensive list of related studies). Since my paper focuses on studying student effort allocation in a pre-existing online STEM course, I use surveys to elicit students' behavioural characteristics. Then heterogeneity in students' attentiveness and myopia/forward-looking perspective is accounted for when seeking to improve the quality of online courses through course design.

Second, my paper is related to recent papers using structural models to inform the design of public policies, especially in education. Macartney et al. (2021) build a quantitative framework for analyzing the distributional effects for student learning when altering teaching accountability incentives. The authors find credible evidence that incentives influence how teachers allocate effort across heterogenous students. They estimate a model of teacher effort to rationalize teacher effort choices under the No Child Left Benefit (NCLB) accountability system. Through counterfactual analysis of different incentive schemes, the authors identify alternative policies that can generate achievement gains (and reduce black-white achievement gaps) for the same cost. Gilraine (2016) builds a model informing the incentive design for schools, accounting for the presence of dynamic complementarities in school inputs. Relative to NCLB, the author finds an alternative policy that uses student-specific achievement targets using baseline test scores that can improve student achievement while decreasing test score inequality. In line with the approaches in these papers, my study estimates a model of effort supply, using it to inform the design of course grading schemes. In contrast, my course-design approach uses administrative data on students' study time allocation, incorporates behavioural elements into the student problem, and uses experimental variation to identify the structural parameters that feature in the effort-setting process.

Third, my paper relates to the literature estimating cumulative education production functions. Such cumulative technology maps present and historical inputs to current learning outcomes. Todd and Wolpin (2007) estimate a cumulative production function for children as a function of child ability and history of family inputs. The authors find that lagged family inputs are significant predictors of cognitive achievement. Consistent with a cumulative technology, Aizer and Cunha (2012) find larger IQ gains from preschool enrolment for children with higher stocks of early human capital. Gilraine (2016) uses year-to-year variation in school accountability to identify dynamic complementarities in school inputs. The author finds a 0.18 SD increase in test scores for students who are in schools that were subject to school accountability in two consecutive periods relative to those subject to accountability only in the previous period. Appendix A.2 describes the education production function literature in more depth. I contribute to this literature by estimating the cumulative learning technology as a function of students' study time allocation. It uses exogenous variation in student study time at each learning stage to identify dynamic complementarities in student effort inputs.

Fourth, the paper also contributes to a body of research studying the incentive design of course assignments. Grodner and Rupp (2013) conduct an experiment in which students in an introductory microeconomic course are randomly assigned to one of two grading schemes. The control group faces four tests worth 25% each, with an optional assignment; the treatment group faces four tests worth 22.5% and an additional assignment worth 10%. The authors find that students who are given incentives to complete the assignment obtained $5 - 6\%$ higher on the term tests, and are 6 percentage points more likely to complete the course. Other studies, such as Pozo and Stull (2006), Artés and Rahona (2013), and Latif and Miles (2020), find that introducing low-stakes graded homework improves performance on high-stakes assessments. Appendix A.3 includes a more extensive list of papers focusing on course design. Although the literature establishes the importance of graded homework assessments, little is known about how grading weights should be distributed across multiple homework assessments that are assigned throughout the course. My paper uses experimental variation to estimate a multi-stage model of student effort supply to inform the optimal dynamic design of such a grading weights profile.

Fifth, my analysis of extrinsic grading incentives relates to the literature using monetary incentives to encourage educational investment. Fryer Jr (2011) investigates the impact of financial incentives on student achievement using school-based field experiments involving over 200 schools in Dallas, New York, and Chicago. The author finds statistically insignificant effects on achievement regardless of whether the monetary incentives were based on reading books or test performance. Angrist et al. (2014) conduct a field experiment at a Canadian commuter college to investigate the impact of offering cash incentives to students for obtaining course grades above 70 percent. Although the monetary incentives increase the number of courses in which students achieved at least a 70 percent, there are no significant impacts on their GPA. Barrow and Rouse (2018) randomize over 5000 high school students to performance-based post-secondary scholarships. The authors find that monetary incentives increase students' time spent on educational activities while decreasing their time allocated to work and leisure. In relation to this literature, my paper studies the role of grading scheme design in improving student learning outcomes by extrinsically incentivizing students to allocate their study time effectively.

Finally, the two field experiments in my paper contribute to a large body of recent work investigating the efficacy of behavioural nudges in promoting desirable academic behaviours in higher education.[5] Smith et al. (2018) conduct a field experiment to evaluate the efficacy of a personalized email message which informed students how their assignment grade will influence their final grade, based on their current grade in the course. The authors find that students who received the message achieved a 4 percentage point higher grade on the assignment. Oreopoulos et al. (2018) investigate the effectiveness of a planning module which involved a group of randomly selected students building a weekly calendar and receiving follow-up reminders from an upper-year coach. The authors find that the online planning module marginally increased self-reported weekly study time, but the increase in weekly study time did not result in an increase in academic performance outcomes. Clark et al. (2020) conduct an experiment to test whether college students who set goals exert more effort and achieve improved learning outcomes. The authors find setting task-based goals increased task completion and subsequent course performance. However, setting performance-based goals does not result in significant increases in learning outcomes. Appendix A.4 includes a more extensive list of related papers exploring student effort choices. I add to this literature by using administrative data on student effort and showing that targeted informational reminders can improve achievement by nudging inattentive students to participate further in learning activities.

3 Institutional Background

The setting for the study is a first year undergraduate online introductory programming course offered at a large research-intensive public university in Canada. This section describes the course structure and the platforms used to facilitate student learning in the course.

Cumulative Course Structure. The course assumes no prior programming knowledge and teaches the fundamentals of programming using Python (see Appendix B.1 for the course outline). It is offered every semester and typically enrols around 1000-1500 students in the Fall and Winter terms, and around 200-400 students in the Summer term. Although the course is offered at the first year level in the Computer Science (CS1) department, it consists of CS-majors and non-majors alike and is not exclusive to first year students; many students who enrol have no programming experience.

The course content can be naturally partitioned into three segments: basic, intermediate, and advanced. Weeks 1 - 4 cover the foundational concepts of programming such as variable declaration and loops. Then weeks 5 - 8 cover intermediate concepts such as nested loops and dictionaries.

[5] Kizilcec et al. (2020) and Harackiewicz and Priniski (2018) discuss a variety of behavioural interventions in the literature that are focused on improving academic outcomes in higher education.

Finally, building on the basic and intermediate learning stages, the course concludes by covering advanced concepts such as algorithms and objected oriented programming. Although the content in week 1 requires no prior programming experience, the topics covered in all others weeks are cumulative as they build on concepts covered in past weeks.[6]

The coursework consists of low-stakes weekly homework assessments and also higher stakes assessments, which include a midterm, and a final exam. The homework in the first two weeks is optional (i.e., they are ungraded) to allow students to practice interacting with the online learning environment.[7] The midterm is typically written in week 5, and the final exam is typically conducted in week 13. In addition, students can obtain course credit by participating in two research surveys that are deployed at the start and end of the course.[8] The graded homework assessments count for 25% of students' overall course grade. The midterm counts for 30% of the students' course grade, and the final exam count for the remaining 45%.

The Online Learning Platform. The weekly homework modules are hosted on an open-source online learning environment created by the computer science department at this institution. The environment is an interactive online platform that allows education providers to bundle video instruction together with multiple choice and open-ended programming problems. Distinct content on the online homework platform is separated by weeks, and each week students are assigned to watch videos and complete follow-up problems. Appendix B.2 further illustrates the user interface of this learning environment. The online learning platform for the course contains around 133 instructional videos and 401 problems, assigned across 12 weeks through homework assessments.

The Online Peer Discussion Board. The course offers an interactive online course discussion board where students can discuss course material. All students in the course are able to sign up for the online discussion board, where they can participate by asking questions, help their peers by writing answers, and engage in discussion with peers by commenting on existing questions and/or answers. Posts on the discussion board are organized per week as new content is introduced weekly. Appendix B.3 describes the user interface of the discussion board in more detail. Although encouraged, participation in the discussion board is completely voluntary, and students are not

[6] For example, learning nested lists and nested loops in week 6 requires that students understand the basics of loops and lists covered in weeks 4 and 5, respectively.

[7] Students can enrol into the course up until the second week. Making the homework in the first two weeks optional also reduces the logistical burden on instructors, as otherwise students who enrol late may demand alternative make-up assessments, or request a grading scheme adjustment.

[8] Both surveys were deployed online using the Qualtrics survey platform. Students earn a 1% bonus credit for each research survey they complete.

given any course credit for registering for it or creating content.

Learning Management System. Canvas is the learning management system (LMS) employed by the course involved in this study. It is used to set up and organize a digital learning environment. In my setting, Canvas is used by instructors to post announcements, manage course deadlines, and release student grades. Students also have a message inbox on Canvas, separately from their institutional email. Instructors can send student messages through Canvas, and such messages are received in students' Canvas messages inbox, and also automatically forwarded to students' institutional email.

4 Data and Descriptive Statistics

This study uses a combination of rich student-level administrative and survey data to characterize online learning participation behaviour for different types of students. All data are gathered from the introductory programming course (introduced above), offered during the Winter 2020, Fall 2020, Winter 2021, and Summer 2021 academic semesters in a large research-intensive Canadian university. The data are collected and merged together from the following sources: online surveys, a learning management system, an online homework platform, and the online peer discussion board. Pilot data were also gathered from the Summer 2019 and Fall 2019 cohorts and served to finalize the design of the primary data collection that is the focus of this section.[9] The timeline of the complete data collection exercise is shown in Figure 10.

Student Survey Data. The baseline survey collects information about students' demographics and elicits information about their behavioural characteristics; the final survey gathers data on various course inputs, interactions with peers, and elicits student feedback about different components of the course. Each survey takes around 20 - 25 minutes to fill out, and is voluntary, although students are given around 1% course credit for completing each survey. The response rate is around 91% for the baseline survey and 86% for the end-line survey.[10] The baseline survey contains a consent form, which asks students to participate in the study by allowing their data to be used for the purposes of academic analysis and research. In addition to the baseline survey, students are also given an opportunity to consent to be a participant of the research study on the online homework

[9] Pilot data collection involved having 30-minute recorded interviews with several students, conducting online surveys using various software, and prototyping various interventions.

[10] The research surveys are announced through the learning management system, and students who did not complete the survey 2 days before the deadline receive a reminder to do so.

environment. Overall the consent rate is around 87%. The sample of total consenting students who completed the course consists of 3686 students.

Student Activity Data from the Learning Management System. I collect student activity data from student interaction reports captured in Canvas. This includes the total number of announcement views, aggregate page views, and a daily list of all students enrolled in the course. The list of students enrolled in the course is retrieved daily to track attrition of students from the sample over the study period.[11]

Student Achievement Data. Student achievement data are collected from the weekly online homework, the midterm, and the final exam. The availability of this high frequency achievement data allows me to assess student learning throughout the course. The primary measure of learning is the student's grade on the cumulative final exam, which is standardized to be mean 0 and standard deviation 1.[12]

Student Discussion Board Activity Data. Student discussion board registration status is collected at the weekly level. Thus, I observe the number of weeks a student is registered for the discussion board. Additionally, I also observe time-stamped data on all contributions (question, answers, or comments), and the number of unique posts viewed by students each week. Overall, I observe discussion board registration, contributions, and 'consumption' decisions.

4.1 Student Study Time Data

Numerous student interactions with the online platforms are observed in the administrative data to the nearest second. That is, the online platforms serve as a monitoring device in terms of students' learning activities. For example, observed interactions include the times when students login or logout, play or pause an instructional video, submit problem solutions, and write in the discussion board. The availability of such rich time-stamped interaction-level data enables me to construct a precise measure of online study time at each stage of the learning process. The study time measure includes minutes spent watching instructional videos, working on homework problems, and reading and writing posts on the discussion board. I will now outline the construction of study time for the online homework, and for the online peer discussion board separately.

[11] The data are retrieved using the Canvas Application Program Interface.

[12] The final exam is a 3-hour comprehensive assessment that evaluates overall understanding of introductory programming in Python.

Online Homework Study Time. I couple the students' time-stamped online interactions together with a basic clustering algorithm to identify periods of learning activity at each stage of the course. The time-stamped interaction data are used to measure the minutes spent watching instructional videos and doing homework problems. The procedure is built around the empirical observation that students tend to study in approximately 30-minute blocks throughout the week (e.g., Tuesday from 6 - 6:30 pm). Each block of homework activity begins with students interacting with the online learning platform for at least 5 minutes, and concludes after 5-minutes of inactivity. Video watching time is computed based on when students play or pause the instructional videos. Students' time spent attempting homework problems is measured using information when students submit problems and click to view the next problem. Then online homework study time is constructed by aggregating all blocks of learning activity for each stage of the course.

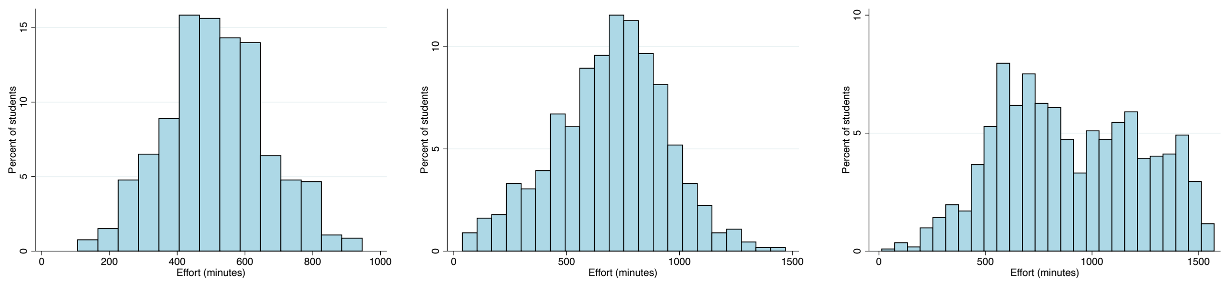
Online Discussion Board Study Time. Although the administrative dataset includes the number of posts written and read at each stage of the learning process, the time spent on these activities is not observed. To fill this gap, the final survey asks students the minutes they spend on average writing and reading a post (see Appendix C.4 for survey questions). The administrative data on engagement, and corresponding student-level survey data on average time use, is used together to measure the minutes spent on the discussion board at each learning stage.[13]

Distribution of Online Study Time. Total online study time at each learning stage aggregates minutes spent on the online learning environment together with the minutes spent on the online discussion board. Figure 2 below shows the distribution of online study time at each learning stage of the course. This figure shows that on average, students' study time increases as the course progresses. Students exerting more effort at later stages of the course is consistent with the grading incentives, as the homework in the first two weeks is optional, the midterm falls in the intermediate stage, and the final exam is at the end of the advanced stage. Additionally, the apparent left skewness of the study time distribution in the advanced stage suggests that many students exert a large fraction of their effort at the end of the course.

[13]For example, suppose a student views 11 posts and writes 5 questions. If in the survey data this student reports to taking 3 minutes to view a post, and 6 minutes to write a question, then their estimated time spent on the discuss board is $11 \times 3 + 5 \times 6 = 63$ minutes.

Figure 1: Distribution of Study Time by Learning Stage

(a) Basic Stage (b) Intermediate Stage (c) Advanced Stage



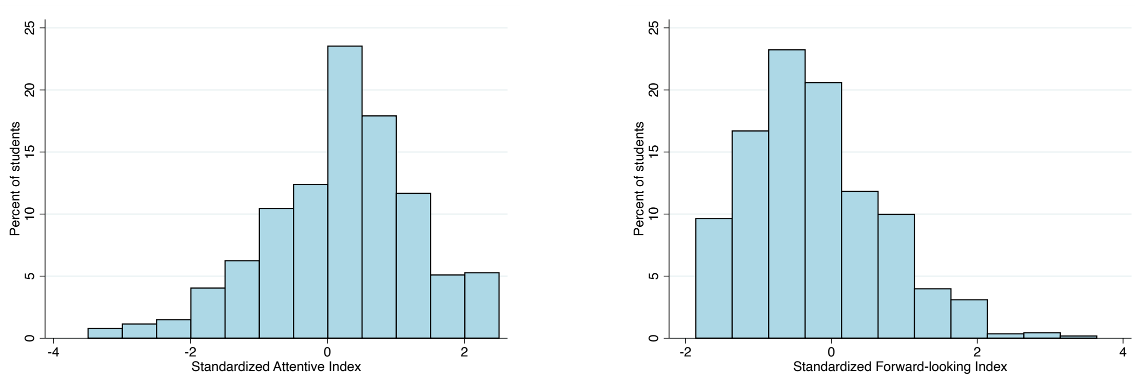
Notes: The figure presents the distribution of total online study time for each stage of the course: basic, intermediate, and advanced. All histograms use a bin width of 60-minutes.

4.2 Student Behavioural Characteristics

The baseline survey elicits students' attentiveness and forward-looking perspectives through a series of questions. Each question is measured on a 7-point Likert scale, and a student's response can vary between strongly disagree (i.e., a response value of 1) to strongly agree (i.e., a response value of 7). For example, to measure students' attentiveness, they are asked whether "I tend to read all the instructor announcements for this course." Similarly, to elicit whether students are forward-looking, one question they are asked is "I consider myself to be a forward-looking person who has clear plans about the future." Both behavioural characteristics are measured using 5 separate questions; all questions relating to students' attentiveness and forward-looking perspective are included in Appendix C.2. To construct a continuous index of the behavioural responses, replies to all relevant questions are aggregated together so that they are increasing in the attribute of interest. The distribution of students' attentiveness and forward-looking view are shown the following figure:

Figure 2: Distribution of Study Time by Learning Stage

(a) Attentiveness (b) Forward-looking View



Notes: The figure presents the distribution of the standardized behavioural index variables: attentiveness and forward-looking view. All histograms use a bin width of 0.5 SD.

The apparent left skew of the attentiveness distribution suggests that most students are attentive. In contrast, the right skew exhibited in the distribution of students' forward-looking view suggests a large fraction of students are myopic.

I also dichotomize the attentiveness and forward-looking behavioural measures for expositional simplicity in later sections. The questions pertaining to students' forward-looking perspective have their response values increasing with students' forward-looking view. I then characterize students as being forward-looking (versus myopic) if they responded with at least 5 (out of 7) to each question. I analogously define whether a student is attentive (versus inattentive).

4.3 Summary Statistics

Table 1 presents a rich set of summary statistics related to student demographics, characteristics, behavioural information, homework participation, and discussion board activity. Although computer science graduates are primarily male (Baer and DeOrio, 2020), there is no significant gender disparity in my sample as 49% of the students are female. The course is offered as a first-year course, but is not exclusive to first-year students, as around 38% of students are beyond their first year. Additionally, around 28% of students are pursuing non-STEM majors. Consistent with only 53% of students being domestic Canadians, only around 29% of students speak English at home. Appendix C.1 contains the survey questions used for gathering the student demographics and other characteristics.

Panel B of Table 1 shows that 87% of students do not have any programming experience prior to taking the course. As shown in Panel C, around 76% of students are attentive, and 32% of students are forward-looking. As indicated in Panel D, around 16% of students do not attempt the low-stakes homework each week. On average, students spend around 25 minutes watching videos each week, and 2 hours working on problems. Panel E shows that around 79% of students in the data (including treatment group) signed up for the discussion board, and from those who registered, the average student spends around 29 minutes on the peer discussion board per week.

5 Experimental Design and Description of Interventions

This section describes the rationale behind the interventions that were deployed and outlines the experimental design for allocating students to treatment.

5.1 Design of Experiments

The sample frame eligible to receive the nudges consists of all students who consented to participate in research during the Winter 2020, Fall 2020, Winter 2021, and Summer 2021 academic terms. As discussed in the previous section, the data collection results in a sample of 3686 study participants. The study followed a double-blind protocol for implementing the randomized interventions. That is, students were not informed of their treatment status but were aware that a study was being conducted for the purposes of improving course design. The course instructors were aware of the interventions that were being deployed but were not informed about the students' treatment status. I performed the randomizations on an anonymized dataset, and I was not part of the instructional team. Prototyping interventions during the pilot data collection in Fall 2020 informed the design of the two interventions presented in this section.

5.2 Description of Interventions

The interventions considered in this study can be categorized as 'targeted informational reminders' as their design includes the following elements: 1) they prompt students to take a specific action (e.g., registering for the discussion board), 2) they provide information on how to clearly execute the action (e.g., instructions to sign-up for discussion board), and 3) they serve as a reminder for the specified task (e.g., course uses a discussion board). The design of the nudges is inspired by insights from psychology and behavioural economics research (Damgaard and Nielsen, 2018). In particular, the interventions are designed to nudge inattentive students who may be less aware of the discussion board, and have a tendency to forget homework deadlines.

Discussion Board Sign-up Activity. The sign-up activity is designed to promote discussion board registration. The activity is composed of the following elements: 1) it presents a link to the discussion board sign up page, 2) it uses screen shots to illustrate key steps for sign up, 3) it summarizes all steps into an animated GIF, and 4) students are disclosed information about the discussion board. The online activity consists of two pages: the first page contains the instructions, and the second includes information about the discussion board. The informational page discusses the functionality of the discussion board, and also discloses the proportion of students' questions that have been answered by either a peer or the instructor. The activity had an interactive component as students were prompted to reflect on the information they were presented (see Appendix D.1 for the self-reflection questions).

For all the students who completed the baseline survey and did not register for the discussion board within the first week of the course, half of them received the sign-up activity. Such a balanced assignment will maximize power if the variance in potential outcome distributions across control and treatment is identical (Tabord-Meehan, 2018). Additionally, students in the treatment group are randomly assigned to receive either sign-up instructions only (i.e., page 1), discussion board information (i.e., page 2), or both. Table 2 shows the assignment of students to control and treatment. The factorial design helps assess which elements of the sign-up activity are most effective for nudging students to utilize the discussion board. The primary analysis combines the three conditions into a single treatment group that is compared with the control group.

Homework Reminder Messages. The reminder messages were aimed at promoting students to further participate in their weekly low-stakes homework. Reminders are only sent for the graded homework assessments after week 2.[14] The homework reminder is composed of the following three elements: 1) reminding students of the upcoming homework deadline, 2) prompting them to set aside time in their schedule to next make progress on the homework, and 3) including a direct link to the homework assessment. Appendix D.2 shows the template of the homework reminder message. The reminder messages were sent within 48 hours after the homework assignment was released and are deployed using the learning management system (i.e., Canvas). Students would receive the reminder both in their Canvas and institutional email inbox.

For students who had not completed the homework before the deployment of the reminder message, half of them are randomly assigned to receive a homework reminder.[15] The reminder messages were sent throughout the course and were re-randomized with each deployment. Consequently, the number of total homework reminders a student receives follows a binomial distribution with 10 trials and a 0.5 probability of success. Figure 11 illustrates the assignment of students to the number of homework reminders.

5.3 Statistical Validity of Experiments

I now discuss the statistical validity of the experimental design by showing the following: 1) pre-treatment characteristics are balanced across the control and treatment group, 2) there is no differential attrition by treatment status, and 3) results are robust to spillovers.

[14]The courses instructors would make important announcements in the first two weeks to get students started with the course. Consequently, the reminders were not sent during this week to avoid crowding out the instructors' announcements.

[15]Each week, only around $5\% - 10\%$ of students completed the homework within 48 hours of release.

Independence of Treatment Assignment. The aim of the experiments is to identify Intent to Treat (ITT) effects of interest. The ITT is identified as students are randomly assigned to either the control or treatment group across all interventions. I investigate the validity of the random assignment by testing whether the pre-treatment student demographics and characteristics are balanced across the control and treatment group. I do so by standardizing each pre-treatment control and regressing these on the treatment status. Figure 12 shows a well-balanced control and treatment group for the sign-up activity. Similarly, Figure 13 shows that students who are assigned to receive an extra homework reminder are statistically identical in their demographics and characteristics at baseline.

Student Attrition. Student attrition is natural in my setting as students who initially enrolled and consented to participate in the study can choose to drop out from the course afterwards. In my sample of 4091 students who initially agreed to participate in the study, around 90% of them completed the course. Table 3 reports the differences in the proportion of attriters in the control and treatment groups across both interventions studied in this paper. The analysis suggests that none of the interventions caused students to dropout of the course directly as all treatment coefficients are close to 0 and the corresponding p-values are larger than 0.1.

Well-defined Treatment Assignment. For the treatment allocation to be well-defined, the following two assumptions must hold true: 1) the treatment level is unique so that potential outcomes are well defined, and 2) the treatment applied to one student does not affect learning outcomes of other students. The intensity of the sign-up activity and homework reminders is homogenous across the treatment groups as all students within a treatment condition receive the same intervention. Therefore, the potential outcomes corresponding to the experimental conditions are well defined.

Next, I discuss the possibility of spillover effects across students. Since students can interact with each other on the discussion board and work towards solving problems, it is possible that students in the treatment groups who received some nudge will interact with the control group who did not receive any intervention. Assuming the nudge increases an outcome of interest (e.g. more participation on homework problems), that can result in positive spillovers to the control group through information sharing (e.g., answering questions of control group students) or peer effects (e.g., control group student mimicking behaviour of treatment group student). Such positive spillover effects will result in downward biased effect sizes.

Although the experimental design does not guard against such spillovers in this setting, the online nature of the course mitigates standard in-person student interactions that would typically be present. Additionally, I am able to leverage certain features of the data collection for robustness analysis. The baseline survey collected data on whether students are in a study group, the number of other students in the course they study with, and how frequently they meet. The final survey also directly asked students whether they discussed information shown in the sign-up activity or reminder messages with other students. I use this survey data to discuss the robustness of my primary results to potential spillovers in the next section. Appendix C.5 includes the survey questions about student peer interactions.

6 Empirical Framework and Results

This section discusses the results from the fields experiments described in Section 5, and outlines the corresponding empirical methodology.

6.1 The Effect of Sign-up Activity on Discussion Board Accessibility

To measure the effect of the sign-up activity on discussion board accessibility, I estimate the following specification:

$$Y_i = \alpha_0 + \alpha_1 SignupActivity_i + X_i'\gamma + \epsilon_i,$$

where Y_i denotes either an indicator for signing up for the discussion board, or the number of weeks a student utilizes the online peer forum; $SignupActivity_i$ is an indicator denoting whether a student receives the sign-up activity nudge. Control variables in X_i include student demographics and pre-treatment characteristics listed in Panels A and B of Table 1.

Figure 14 illustrates the effect of being assigned the sign-up activity nudge on discussion board utilization. Showing clear registration instructions and/or providing information to students about the discussion board increased utilization by around 4.5 weeks relative to the control group. The magnitude is large when considering that none of the students who were eligible for treatment had registered for the discussion board at the start of the course, and the intervention more than doubled the duration of utilization (i.e. from around 3 weeks to 7.5 weeks).[16] This effect size is also statistically significant at the 1% significance level with an F-statistic exceeding 100. Table 4 shows that the sign-up activity increases discussion board registration by end of the course by

[16] Among students that sign up for the discussion board, only 3% are inactive. The vast majority of students registered to the discussion board at least view a few posts each week – behavior that can be tracked by the researcher.

17 percentage points. Additionally, the treatment effects are also stable following the inclusion of pre-treatment control variables and cohort fixed effects. Looking ahead, assignment to the sign-up activity nudge will provide a strong first stage for inducing exogenous variation in discussion board utilization.

Mechanisms Underlying the Sign-up Activity. I investigate the mechanisms underlying the sign-up activity by examining how the nudge affects the extent to which different types of students are informed about the discussion board. I investigate whether the sign-up activity informs inattentive students about the online discussion board. To do so, I use student responses from a question that was embedded at the end of the baseline survey (i.e., post sign-up activity) which elicited whether students are well informed about the existence and functionality of the discussion board. Figure 15 displays the propensity to be well informed about the discussion board as a function of students' attentiveness by treatment status. Clearly, the nudge was successful in closing the information gap between inattentive and attentive students.

Next, I investigate whether the component of the sign-up activity that informs students about the functionality of the discussion board induces utilization more than when only registration instructions are provided. Figure 16 shows that receiving only the functionality information increases discussion board utilization by an extra week relative to receiving only registration instructions. (The increase in discussion board utilization for students who receive both the instructions and information is not statistically different from receiving the functionality information only.)

6.2 Learning Value-Added from Discussion Board Utilization

Identifying the causal effect of discussion board utilization on final exam achievement is challenging, given students are likely to self-select into registering for the online discussion board based on their expected learning benefit. For example, attentive and motivated students are going to be more likely to utilize the discussion board and also likely to perform better on the final exam, resulting in upwards bias. To circumvent such endogeneity issues, I use the sign-up activity as an instrument for discussion board utilization. This is a valid instrumental variable as it is randomly assigned to students (i.e., it is independent), does not directly affect the exam grade (i.e., is excludable), and directly affects discussion board utilization (i.e., is relevant).

To evaluate the learning gains from discussion board utilization, I estimate the following 2SLS model:

$$\begin{cases} A_i = \beta_0 + \beta_1 WeeksRegistered_i + X_i'\theta + \epsilon_i, \\ WeeksRegistered_i = \tau_0 + \tau_1 SignupActivity_i + X_i'\pi + \epsilon_i \end{cases}$$

where A_i denotes achievement outcome such as the final exam grade or the average homework grade; $WeeksRegistered_i$ is the number of weeks a student is registered for the discussion board.

Table 5 presents the results for the effect of discussion board accessibility on student achievement. The results indicate that an extra 5 weeks of discussion board accessibility increases mean homework and final exam achievement by 0.14 SD and 0.07 SD, respectively. The different effect sizes across the homework and final exam are consistent with the course rule that allows students to discuss the homework problems on the online peer discussion board, but that forbids students from discussing questions from their online exam with one another. These magnitudes are large, equivalent to increasing the course grade by half a letter grade (e.g., B+ to A-). A back-of-the-envelope calculation predicts that 31% of students in the control group would have received half a letter higher course grade had they also been assigned to the sign-up activity.

6.3 The Effect of Homework Reminders on Homework Participation

To measure the effect of receiving reminder messages on students' homework participation, I estimate the following specification:

$$D_i = \delta_0 + \delta_1 RemindersFreq_i + X_i'\Delta + \epsilon_i,$$

where D_i is either the number of homework assessments completed or the total hours a student spends watching videos and doing problems; $RemindersFreq_i$ represents the total number of homework reminders a student receives.

Table 6 presents the results from estimating the above specification. On average, receiving 5 additional reminders induces students to complete an extra homework assessment. Additionally, the estimates show that receiving an extra reminder message increases the time spent on homework by 23 minutes. Since students spend around 2.4 hours each week on homework, receiving a homework reminder increases corresponding homework study time by around 16%. This effect size is also statistically significant at the 1% significance level, with an F-statistic exceeding 100.

Figure 17 illustrates the average number of homeworks completed by the number of reminder messages received. Clearly, receiving more reminders encourages students to complete more homework. The figure also suggests that the marginal increase in homework completion is decreasing with more reminders, although the apparent diminishing returns to homework reminders are not statistically significant.[17]

Mechanisms Underlying the Homework Reminders. Similar to the investigation of mechanisms for the sign-up activity, I examine whether the reminder messages are more helpful for less attentive students. The final survey asked whether the students found the reminder emails to be helpful in keeping on track with the homework assessments. Figure 18 illustrates the relationship between finding the reminders useful and a student's attentiveness. The significant negative linear association suggest that less attentive students are more likely to be helped by the homework reminders.

6.4 The Effect of Homework Participation on Learning

Since students choose their level of homework participation, associating homework participation with learning outcomes will likely result in biased estimates due to omitted variable bias. For example, students who have a higher programming ability will obtain better grades on course assessments, while exerting less effort than students with lower innate programming ability. As a result, the returns to homework participation will be downward-biased through this unobserved programming ability channel. To circumvent such issues of endogeneity, I use random assignment to the number of homework reminders received as an instrument for homework participation. Email reminders are a valid instrument for homework participation as they are randomly assigned to students (i.e., independent), do not directly affect learning outcomes (i.e., are excludable), and promote students to complete homework successfully (i.e., they are relevant).

I estimate the causal effects of study time on homework performance using the following 2SLS model:

$$\begin{cases} ExamGrade_i = \lambda_0 + \lambda_1 D_i + X_i'\Pi + \epsilon_i, \\ D_i = \phi_0 + \phi_1 RemindersFreq_i + X_i'\Gamma + \epsilon_i \end{cases}$$

where $ExamGrade_i$ denotes the final exam grade. Table 7 presents the 2SLS results. The estimates show that completing one extra homework increases the final exam grade by around 0.18 SD. Additionally, an extra hour spent studying through doing online homework increases final exam grade by 0.09 SD. These estimates are statistically significant at the 1% level. The large effects reflect that fact that the homework is the primary source of learning the course material in this

[17] Regressing homework completion on the reminder frequency and the square of the reminder frequency results in a negative, but statistically insignificant, coefficient on the quadratic term.

online STEM course.

6.5 Robustness to Spillover Effects

I now present two pieces of evidence supporting the view that the main results presented in this section are not severely affected by spillover effects from treated to control students. First, only around 12% of students in the final survey attested to discussing contents of either the sign-up activity or reminder messages with their peers. Therefore, information spillovers from the treatment to the control group would be expected to be small. Second, around 17% of the students in the course are in study groups, where they meet at least once a month and discuss course material. I investigate whether the treatment effects for either the sign-up activity or reminder messages vary according to whether students are in a study group at baseline. The analysis is presented in Table 8. The results there suggest that the sign-up activity is less effective for students in a study group, although the relevant estimates are only marginally statistically significant. In contrast, the efficacy of the reminder messages does not vary according to whether a student is in a study group.

7 Theoretical Framework

This section presents a conceptualization of an online course with a cumulative structure where students exert effort and accumulate knowledge across multiple learning stages. This framework formalizes the underinvestment of effort by myopic students in learning foundational skills, and serves to guide the design of a grading scheme to optimize student learning.

7.1 The Environment

Consider N students in a course, who allocate total study time or 'effort' (e) across three learning stages $t \in \{basic, int, adv\}$. Then, let L_i^t denote the amount of learning for student i during stage t. Students can vary in their baseline human capital (h), English proficiency (E), and whether they are forward-looking (f). Human capital h_i and English proficiency E_i are standardized to have mean 0 and standard deviation 1. The variable $f_i = 0$ denotes a myopic student, and $f_i = 1$ represents a forward-looking student. Let p_m be the proportion of myopic students in the course. The instructor sets the grading weight w_t for each learning period.

The timeline of the model is as follows. First, the instructor specifies the grading weights $(w_t)_t$. Then, given the grading scheme, students allocate their effort (i.e., study time) across the course

$(e_i^t)_t$. For forward-looking students, I solve this model by backwards-induction, and therefore begin by discussing the students' effort choice problem first, and then outline the instructor's problem.

7.2 The Student Effort Choice Problem

Forward-looking students internalize the cumulative learning process and allocate their effort to maximize their course grade net of effort costs:

$$\max_{(e_i^t)_t} \sum_t w_t L_i^t(e_i^t; L_i^{t-1}) - C(e_i^{basic}, e_i^{int}, e_i^{adv}; E_i), \tag{1}$$

where $C(\cdot)$ is a convex function representing the cost of effort exertion. The learning technology at each stage of the learning process is cumulative. As a result, the amount of learning L_i^t in a given period depends on present effort e_i^t, and previous knowledge L_i^{t-1}.[18] I assume the learning production functional is concave, increasing in effort and baseline human capital, whereas the cost function is convex and increasing in effort. Additionally, the cost of effort decreases in English proficiency.[19]

In contrast, the myopic students do not internalize the cumulative course structure, and focus on each learning stage separately. As a result, they allocate effort to maximize their grade in each learning stage net of effort costs:

$$\max_{e_i^t} w_t L_i^t(e_i^t; L_i^{t-1}) - C(e_i^t; E_i) \text{ for each } t \in \{basic, int, adv\}. \tag{2}$$

Let $e_i^{t,*}$ denote the optimal effort of student i in stage t resulting from solving their effort-choice problem.

Incorporating Additional Assessments. The effort allocation model is flexible and can be adapted to incorporate beyond three assessments. In the actual course, students are assigned online assignments at each learning stage, a midterm based on material covered in the basic stage, and a cumulative final exam. Then the forward-looking students allocate effort to maximize:

$$\max_{(e_i^t)_t} \sum_t w_t L_i^t(e_i^t; L_i^{t-1}) + w_{mid} L_i^{mid}(e_i^{basic}; h_i) + w_{final} L_i^{final}(e_i^{adv}; L_i^{int}) - C(e_i^{basic}, e_i^{int}, e_i^{adv}; E_i). \tag{3}$$

The myopic students effort choice problem in the basic and intermediate stage is as represented in equation 2. For the advanced stage, however, the myopic students internalize that studying

[18] Since the course structure is assumed to be cumulative, L_i^{t-1} is used as a sufficient statistic for all prior knowledge accumulation. The prior knowledge at the basic stage is denoted by L_i^{-1} and is the baseline knowledge h_i.

[19] Students who are proficient in English are going to have an easier time understanding the contents of the videos and interpreting the homework problems.

during the advanced stage benefits both their assignments and final exam performance.[20] The final exam has the highest weight of all assessments and the university announces the exam schedule at the advanced stage. Therefore even the myopic students internalize the importance of exerting effort during the advanced stage and do so by maximizing:

$$\max_{e_i^{adv}} w_{adv} L_i^{adv}(e_i^{adv}; L_i^{int}) + w_{final} L_i^{final}(e_i^{adv}; L_i^{int}) - C(e_i^{adv}; E_i). \tag{4}$$

7.3 Instructor's Grading Scheme Design Problem

Suppose primary evaluation in the advanced learning stage is a comprehensive and cumulative final exam. Then, the instructor's goal is to have the representative student allocate their effort throughout the course to maximize their grade on the final exam net of effort costs. Since p_m proportion of the students in the course are myopic, the instructor weights learning and effort costs across representative myopic and forward-looking students as follows:

$$\begin{aligned} p_m[L_i^{final}(e_i^{adv}; L_i^{int}, \bar{h}, f_i = 0) - C(e_i^{basic}, e_i^{int}, e_i^{adv}; \bar{E}, f_i = 0)] - \\ (1 - p_m)[L_i^{final}(e_i^{adv}; L_i^{int}, \bar{h}, f_i = 1) - C(e_i^{basic}, e_i^{int}, e_i^{adv}; \bar{E}, f_i = 1)], \end{aligned} \tag{5}$$

where $\bar{h}$ is the average baseline human capital, and $\bar{E}$ is the average English proficiency.[21] The instructor uses the final exam grade as the primary outcome measure as it is cumulative and best represents the totality of the course material relative to other assessments in earlier stages. To induce students to efficiently allocate their effort, the instructor sets the grading scheme $(w_t)_t$ to maximize the above objective.

7.4 Stylized Example

To intuitively illustrate the implications of the model, consider a course with a cumulative structure and two learning stages. Suppose in the first half of the course that students learn basic concepts, and advanced concepts for the remaining half. That is, $t \in \{basic, adv\}$. Additionally, students only vary according to whether they are forward-looking ($f_i = 1$) or myopic ($f_i = 0$).

Parameterization of Learning Technology and Cost Function. Let the following simple learning technologies represent the cumulative learning process:

[20] Students in this setting do not study for the midterm and final exam distinctly from the assignments. The assignments also serve as the main source of preparation for the high-stakes assessments.

[21] The average baseline knowledge and English proficiency are similar across myopic and forward-looking students in the data. Then in the model I assume student characteristics are independant of students' forward-looking view.

$$L_i^{basic} = \alpha_1 e_i^{basic},$$

$$L_i^{adv} = \beta_1 e_i^{adv} + \beta_2 L_i^{basic} + \beta_3 e_i^{adv} \times L_i^{basic}.$$

A positive marginal benefit of effort at both learning stages implies that $\alpha_1 > 0$ and $\beta_1 > 0$. Since the advanced learning stage is cumulative, then clearly $\beta_2 > 0$. Finally, assuming effort exertion in the basic stage increases the productivity of advanced stage effort (i.e., there are dynamic complementarities in effort), then $\beta_3 > 0$.

The cost of effort is assumed to be linearly separable and represented by a quadratic cost function:

$$c(e_i^t) = \kappa \frac{(e_i^t)^2}{2} \text{ for } t \in \{basic, adv\},$$

where κ represents the steepness of the cost function.

The Students' Optimal Effort Choice. Let us assume there is no bonus credit, and so $w_{basic} + w_{adv} = 1$. The forward-looking student sets effort by maximizing her expected course grade net of costs using backwards-induction, beginning at the advanced stage. The resulting effort allocation across the basic and advanced learning stages are:

$$e_i^{basic,*}(f_i = 1) = \frac{\alpha_1[\kappa(1 - w_{adv}(1 - \beta_2)) + \beta_1\beta_3 w_{adv}^2]}{(\kappa - w_{adv}\alpha_1\beta_3)(\kappa + w_{adv}\alpha_1\beta_3)},$$

$$e_i^{adv,*}(f_i = 1) = \frac{w_{adv}[\beta_3\alpha_1^2[\kappa(1 - w_{adv}(1 - \beta_2)) + \beta_1\beta_3 w_{adv}^2] + \beta_1(\kappa - w_{adv}\alpha_1\beta_3)(\kappa + w_{adv}\alpha_1\beta_3)]}{\kappa(\kappa - w_{adv}\alpha_1\beta_3)(\kappa + w_{adv}\alpha_1\beta_3)}.$$

In contrast, the myopic students do not internalize the cumulative learning process, and supply effort as follows:

$$e_i^{basic,*}(f_i = 0) = \frac{\alpha_1(1 - w_{adv})}{\kappa},$$

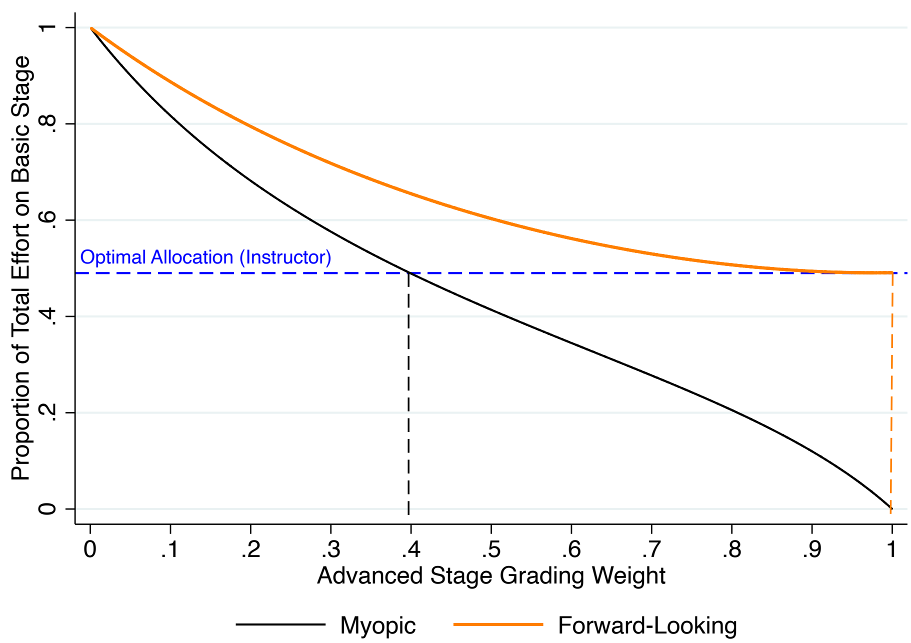
$$e_i^{adv,*}(f_i = 0) = \frac{w_{adv}[\beta_3\alpha_1^2(1 - w_{basic}) + \kappa\beta_1]}{\kappa^2}.$$

Clearly, myopic students exert less effort than forward-looking students in the basic learning stage as long as the assessment in the advanced stage is cumulative (i.e., $\beta_2 > 0$), and there are dynamic complementarities in effort inputs (i.e., $\beta_3 > 0$). Otherwise, if the course covers distinct and unrelated topics (i.e., $\beta_2 = \beta_3 = 0$) then both myopic and forward-looking students allocate effort identically:

$$e_i^{basic,*}(\beta_2 = \beta_3 = 0) = \frac{\alpha_1(1 - w_{adv})}{\kappa}, \text{ and } e_i^{adv,*}(\beta_2 = \beta_3 = 0) = \frac{\beta_1 w_{adv}}{\kappa}.$$

The Instructor's Optimal Grading Scheme. Suppose the instructor wants students to allocate their effort to maximize learning (i.e., their advanced stage grade) net of effort cost. The optimal grading scheme to incentivize the students to effectively allocate their study time varies according to whether the student is myopic or forward-looking. The following figure summarizes the model implications, simulating the proportion of total effort that is exerted in the basic stage as a function of the advanced grading weight; reasonable parameter values are used to represent a course with a cumulative structure.

Figure 3: Proportion of Total Effort Expended in the Basic Stage by Grading Scheme



The figure shows that myopic students' relative basic effort supply decreases sharply as less weight is assigned to the basic stage. In contrast, forward-looking students always allocate effort to the basic stage even if it does not count towards any course credit. The instructor's problem and the forward-looking student problem align when all weight is placed on the advanced stage. Therefore assigning 100% weight to the advanced stage is the optimal grading scheme for forward-looking students. However, for the myopic students, the instructor needs to assign appropriate weight to the basic stage to incentivize students to adequately exert effort during that stage. If the course structure is sufficiently cumulative, then the instructor should assign more weight to the basic stage than the advanced stage. Figure 3 also shows that the learning outcomes for myopic students are much more influenced by the grading scheme design than those of the forward-looking students.

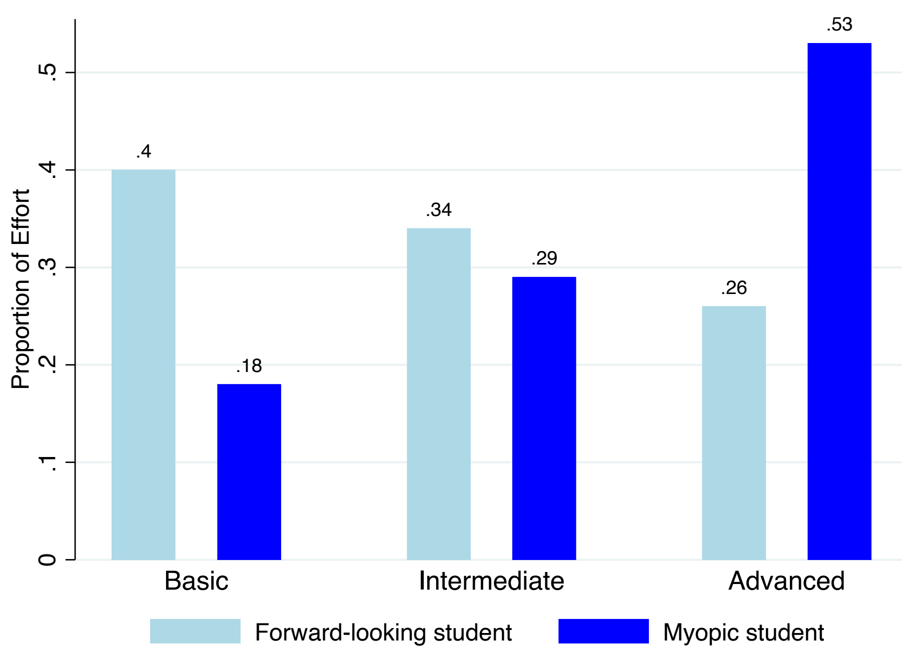
8 Estimating a Model of Online Learning

In this section, I describe the estimation of the model introduced in the previous section. The model is informed by the actual structure of the introductory programming course under consideration, noting that students learn across three distinct learning stages: basic (e.g., loops), intermediate (e.g., nested loops), and advanced (e.g., algorithms). Further, the model takes advantage of the unique data in this setting, the administrative data allowing me to observe in a precise way both the total online study time spent on each learning stage and the corresponding learning associated with each stage. Students' English proficiency, and forward-looking status are inferred from the survey data (see Appendix C.2 and C.3 for details).

8.1 Validating Theoretical Model Implications

The model presented above predicts that myopic students invest less effort early on in the course than forward-looking students. To test whether this implication holds in the data, Figure 4 shows the average proportion of total study time allocated to each learning stage by forward-looking students and myopic students.

Figure 4: Allocation of Effort Across Learning Stages



The figure makes clear that on average, myopic students' effort increases as the course progresses, whereas forward-looking students front-load their study time allocation. The myopic students' effort allocation is consistent with the actual course grading scheme, which is increasing as the course progresses, as the initial two online assignments are ungraded, followed by the graded midterm in the middle, and the final exam with the most weight at the end. This evidence is consistent with the model prediction that myopic students exert less effort learning foundational skills, as they do not internalize the cumulative benefits of exerting effort early on, given the course structure, when choosing effort.

8.2 Estimating the Model

To rationalize the multi-stage effort-setting process by myopic and forward-looking students under a given grading scheme, I now estimate the three-learning stage model discussed in the previous section. Since students in the model choose effort optimally in order to balance their expected benefit against the costs, I estimate the cumulative learning technology and the implied learning benefits as well as a convex cost to effort exertion. The estimated model will serve as a foundation for conducting counterfactual experiments that inform the optimal grading scheme design, as anticipated in the Introduction. This approach is analogous to the one followed in Macartney et al. (2021). Those authors estimate a model of teacher effort as a function of accountability incentives under the No Child Left Behind Act of 2001, and use the estimated model to conduct counterfactual analyses of alternative incentive schemes.

8.3 Specifying the Learning Technology

The benefit of effort exertion in a given learning stage depends on the learning technology. The technology maps effort inputs into contemporaneous learning for a given stage of the learning process. While the true technology is unknown, I impose minimal structure on the learning technology to serve as a first-order approximation, using the following assumptions.

Assumption 1: The learning technology is linear and additive in inputs

First, I assume the learning technology is linear and additive in effort and prior knowledge. The linear structure allows me to identify the marginal benefit of effort using instrumental variable estimation. Consistent with the cumulative nature of programming, I also assume the learning technology is cumulative.

Assumption 2: The learning technology is cumulative

Second, I assume the technology is cumulative, allowing learning beyond the basic stage to be cumulative. For example, learning in the intermediate stage is increasing in the knowledge accumulated in the basic stage. The cumulative technology reflects the cumulative course structure.

Assumption 3: The learning technology includes dynamic complementarities in effort

Third, I allow for the productivity of study time in the present stage to depend on the knowledge accumulated in the previous stage. Putting all three assumptions together, the learning technology in the basic stage is as follows:

$$L_i^{basic} = \alpha_0 + \alpha_1 e_i^{basic} + \alpha_2 h_i + \alpha_3 e_i^{basic} \times h_i + \epsilon_i^{basic}, \tag{6}$$

where L_i^{basic} is the basic stage homework performance, e_i^{basic} is the total online study time at the basic learning stage, h_i denotes baseline programming experience, and ϵ_i^{basic} is a mean 0 stochastic error term. In equation 6, $\alpha_1 > 0$ implies a positive marginal benefit of effort. The extent to which prior programming knowledge persists to the basic stage is captured by $\alpha_2 > 0$. For $\alpha_3 > 0$, the marginal learning gains from basic effort exertion are increasing in baseline knowledge. The learning technology at the intermediate and advanced learning stages are analogously defined as:

$$L_i^{int} = \beta_0 + \beta_1 e_i^{int} + \beta_2 L_i^{basic} + \beta_3 e_i^{int} \times L_i^{basic} + \epsilon_i^{int}, \tag{7}$$

$$L_i^{adv} = \lambda_0 + \lambda_1 e_i^{adv} + \lambda_2 L_i^{int} + \lambda_3 e_i^{adv} \times L_i^{int} + \epsilon_i^{adv}, \tag{8}$$

where L_i^{int} is a sufficient statistic for previously accumulated knowledge in equation 8, reflecting the cumulative course structure. Given a positive marginal benefit of effort at each learning stage, dynamic complementarities in effort across the stages implies $\beta_3 > 0$ and $\lambda_3 > 0$. As the midterm is based on the basic stage material, the technology mapping basic stage effort to midterm performance is defined analogously to equation 6. Similarly the technology that maps effort inputs to the cumulative final exam performance is defined analogously to equation 8.

8.4 Specifying the Cost Function

I also impose some minimal structure on the cost function for tractability, making the following assumptions.

Assumption 4: The cost function is linearly separable across learning stages

The assumption that the cost function is linearly separable across the learning stages amounts to:

$$C(e_i^{basic}, e_i^{int}, e_i^{adv}; E_i) = C^{basic}(e_i^{basic}; E_i) + C^{int}(e_i^{int}; E_i) + C^{adv}(e_i^{adv}; E_i).$$

The assumption implies that exerting effort in a given learning stage does not affect the marginal cost of studying in any future period. A violation to this would be if students 'burn out' from exerting too much effort early in the course, impairing performance subsequently. As most students in the data allocate the majority of their study time in the advanced stage (as shown in Figure 4), the separability assumption is reasonable.

Assumption 5: The cost function is convex

The cost function for a given stage is specified as a power function:

$$C^t(e_i^t) = \kappa_t \exp(-\gamma_t E_i) \frac{(e_i^t)^{1+\gamma_t}}{1+\gamma_t} \text{ for } t \in \{basic, int, adv\},$$

where κ_t is the steepness, and $\gamma_t > 0$ represents the convexity. Intuitively, a convex cost of effort reflects the tendency for students to become fatigued the longer time they spend studying.

8.5 Estimating the Marginal Benefit of Effort Parameters

The marginal benefit of effort parameters are estimated using 2SLS using the number of randomly assigned reminders a student receives at each learning stage to construct the relevant instruments. For the basic learning technology, the number of reminders received at the basic stage is used to instrument for total basic stage study time. The intermediate learning technology has two endogenous variables: the intermediate stage effort and basic stage knowledge. I use the number of reminders received at the basic and intermediate stages separately as instruments to estimate the intermediate learning technology. The advanced learning technology is estimated analogously.

The marginal benefit parameter estimates are shown in Table 9. The estimates show a positive marginal benefit of effort at each learning stage. Complementarities in present effort exertion and previous knowledge are present in both the intermediate and advanced learning stages. The results are consistent with most students having no prior programming experience, and also reflect the cumulative learning structure of programming.

8.6 Estimating Marginal Cost of Effort Parameters

After estimating the cumulative learning process, I estimate the marginal cost of effort parameters $\Theta = (\kappa, \gamma)$ using maximum likelihood estimation (MLE) at each learning stage. I construct the likelihood (l) using an 'implementations error' approach (Bernheim et al., 2019). That is, I assume students implement the optimal effort choice with error:

$$\underbrace{e_i^t}_{\text{Observed effort}} - \underbrace{e_i^{t,*}(\Theta^t)}_{\text{Optimal effort from model}} \sim \underbrace{N(0, \sigma_t^2)}_{\text{Deviation from optimal distribution}}, t \in \{basic, int, adv\}.$$

Then the resulting log-likelihood function is:

$$l(\Theta^t; (e_i^t)_i) = -n \log(2\pi) + \frac{n}{2} \log(\sigma_{\epsilon t}^2) - \frac{1}{2\sigma_{\epsilon t}^2} \sum_{i=1}^{N} (e_i^t - e_i^{t,*}(\Theta^t))^2.$$

The parameter estimates are then uncovered numerically by carrying out the following steps, iteratively maximizing the likelihood function:

1. Start with an initial value of $\tilde{\Theta}^t$.
2. Compute $e_i^*(\tilde{\Theta}^t)$ for students $i = 1, \ldots, n$.
3. Use $\tilde{\Theta}, e_i^*(\tilde{\Theta}^t)$, and e_i for each student to compute likelihood $l(\tilde{\Theta}^t)$.
4. Update $\tilde{\Theta}^t$ to $\tilde{\Theta}^{t'}$ using Newton's method to take the next step.
5. Iterate through steps 2-4 until convergence: $|\tilde{\Theta}^t - \tilde{\Theta}^{t'}| < 10^{-6}$.

The estimation routine results in parameter estimates that maximize the likelihood function:

$$\widehat{\Theta}^t_{MLE} = \arg\max_{\Theta^t} l(\Theta^t).$$

Maximum likelihood estimation is carried out in three learning stages since the cost function is linearly separable.[22] The marginal parameter estimates are shown in Table 10. The increasing estimates for the convexity parameter across learning stages are consistent with the course becoming progressively more difficult, increasing the rate at which fatigue from studying is accumulated.

8.7 Goodness of Fit

To evaluate the model fit, I start by comparing the distribution of observed effort to the effort distribution predicted by the model. Figure 19 shows clearly that the model fits the data well, as the mean and variance of effort distribution predicted by the model at each learning stage are closely aligned with the observed data. Additionally, I examine the association between observed effort and the corresponding effort implied by the model. Figure 20 shows that there is a strongly

[22] I also repeat the iterative MLE for several initial values and find the resulting estimates are fairly stable, suggesting a global optimum is achieved.

linear relationship between the model implied effort and observed effort at each learning stage. That is, students who exert effort well beyond the average are also predicted by the model to be exerting higher amounts of effort.

8.8 Identification of Model Parameters

Next I discuss the features of the data and sources of exogenous variation that drive the values of the parameter estimates, following best practices for structural research outlined in Andrews et al. (2020). I use rich micro-data and the two field experiments to identify the marginal benefit and marginal cost of effort parameters at each learning stage.

Identification of Marginal Benefit Parameters. Identifying the cumulative technology requires randomization in student effort, learning stage by learning stage. The marginal benefit parameters are identified using the exogenous variation in online learning participation within a student across the learning stages induced by the randomly assigned homework reminders throughout the course. Consistent with the cumulative course structure, the learning technology at each stage of the learning process is a function of present period total study time and previously accumulated knowledge. I can instrument for both endogenous variables by using the number of randomly assigned homework reminders a student receives at each learning stage. Therefore the repeated homework reminders identify marginal benefit parameters.

Identification of Marginal Cost Parameters: The marginal cost parameter vector is $\Theta_t = (\gamma_t, \kappa_t)$ where t indexes the relevant learning stage. While these parameters are estimated jointly using MLE, I will discuss moments in the data induced through the sign-up activity that identify the cost of effort function at each stage. The key cost parameter is γ_t, which represents the convexity of the cost function. Since the cost function is assumed to be linearly separable, identification can be considered at each stage separately. Through log-linearizing the first-order conditions of the students' problem, it can be shown that γ_t is proportional to the effort elasticity with respect to the marginal learning benefit after conditioning on English proficiency. Identification of the convexity parameter therefore requires exogenous variation in effort across students within each learning stage. Since the sign-up activity was deployed at the start of the course, it had persistent effects in terms of increasing average study time at each learning stage. Consequently, I can use random assignment to the sign-up activity to identify the convexity parameter of the cost function at each stage.

Intuitively, consider two distinct myopic students i and j, who have the same English proficiency, i.e., $E_i = E_j$. Then both students have the same marginal cost to exerting effort at each stage. Suppose, however, that $L_i^{basic} > L_j^{basic}$, as student i further utilizes the discussion board at the basic stage due to randomly receiving the sign-up activity. Then differencing the log-linearized first order condition across the two students within the intermediate stage results in the following equation:

$$log\left(\frac{e_i^{int,*}}{e_j^{int,*}}\right) = \frac{1}{\gamma_{int}} log\left(\frac{\beta_1 + \beta_3 L_i^{basic}}{\beta_1 + \beta_3 L_j^{basic}}\right).$$

Then, the above equation makes clear that γ_{int} can be uniquely recovered from the data given that exogenous variation in basic stage knowledge between students with identical English proficiency is available. Given linear separability of the cost function, the identification argument is analogous for γ_{basic} and γ_{adv}.[23] Then κ_t at each learning stage is identified from the corresponding effort-setting first order condition as it is a function of identified quantities.

In summary, as the cost of effort function is characterized by two unknown parameters, at least two moments of the data at each stage are required for identification. Such moments can be constructed using students average effort allocation under the control and treatment groups of the sign-up activity at each stage, conditioning on students' English proficiency.

9 Counterfactual Experiments

This section conducts a set of counterfactual experiments that would be difficult to carry out in the field at a large scale. It does so using the estimated marginal benefit and cost of effort parameters from the behavioural model of student effort supply. First, I simulate the grading scheme that maximizes learning for a given distribution of myopic and forward-looking students. Second, I simulate the optimal grading scheme by students' behavioural type and quantify the learning gains from personalizing incentives. For each counterfactual, I compare the simulated final exam grade to the learning acquired under the existing course structure. Both counterfactuals can be seen as a modification of students' incentives to exert effort throughout the course. Aside from total effort exertion, the simulations show that the specific allocation of effort across the stages of the course is also relevant, as learning in programming is cumulative with strong dynamic complementarities. Third, after establishing the optimal grading scheme for a course where all students are either myopic or forward-looking, I simulate the optimal grading scheme for any share of myopic students.

[23] I treat baseline knowledge h_i as exogenous within the model.

Identifying the Student Effort Response to Different Grading Schemes. To study how student effort allocation changes with the design of the grading incentives, the ideal experiment would randomize a large number of students to a variety of different grading schemes. Although this experimental variation in incentives is unavailable in practice, the model leverages within-student variation in grading weights along with the observed differential responses to incentives by myopic and forward-looking students to identify how changing the grading weights influences the effort-setting process. All marginal benefit and cost of effort parameters in the previous section are estimated using unweighted learning measures and are held constant in the counterfactuals.

9.1 Optimal Assignment Weights

Setup of Optimal Assignment Weights Counterfactual. It should be noted that many universities require instructors to place more weight on the final exam than the midterm. Assigning a large fraction of the weight to the midterm may decrease course completion as students who perform unexpectedly poorly on the midterm through receiving a negative shock (e.g., they become sick on midterm day) may drop out early. I consider a realistic counterfactual exercise where the weights on the midterm and final exam remain fixed, and the assignment weights can vary. Aligned with the actual course grading scheme, I set the midterm weight to be 30% and the final exam weight, 45%. Since myopic and forward-looking students respond differentially to grading incentives, the instructor also internalizes the proportion of myopic students in the course when setting the grading scheme; using survey data, I infer that around 68% of the students are myopic.[24] The grading scheme is also catered to the representative student, with average English language proficiency and average prior programming experience. Aside from the marginal benefit and cost parameters, the remaining fixed parameters of the simulation setup are as follows:

$$w_{midterm} = 0.30, w_{final} = 0.45, h_i = \bar{h}, E_i = \bar{E}, p_m = 0.68.$$

Then, the instructor's problem is to set assignment weights to incentivize the representative student to allocate effort to maximize their grade in the advanced learning stage, captured by their performance on the cumulative final exam:

$$\max_{w_{basic}, w_{int}, w_{adv}} p_m[L_i^{final}(e_i^{adv,*}; e_i^{int,*}, e_i^{basic,*}, \bar{h}, f_i = 0) - \sum_t C(e_i^{t,*}; \bar{E}, f_i = 0)] -$$

[24] I explore the sensitivity of the main simulations to changes in this proportion below.

$$(1-p_m)[L_i^{final}(e_i^{adv,*}; e_i^{int,*}, e_i^{basic,*}, \bar{h}, f_i = 1) - \sum_t C(e_i^{t,*}; \bar{E}, f_i = 1)],$$

$$\text{subject to } w_{basic} + w_{int} + w_{adv} = 0.25,$$

where the cumulative learning technology and cost functions are as estimated in the previous section. Since myopic students' effort-choice is considered independently at each stage, their effort at each learning stage is a function of the stage-level assignment weight and prior knowledge in the basic and intermediate stage. That is, $e_i^t(f_i = 0) = e_i^t(w_t; f_i = 0, L_i^{t-1})$ for $t \in \{basic, int\}$. Given the large weight on the final exam, myopic students do internalize that doing homework in the advanced stage helps prepare for the final exam, i.e., $e_i^{adv}(f_i = 0) = e_i^{adv}(w_{adv}, w_{final}; f_i = 0, L_i^{int})$. As the forward-looking students internalize the cumulative course structure, their effort at each stage is a function of current and future assessment weights, and prior knowledge:

$$e_i^{basic,*}(f_i = 1) = e_i^{basic,*}(w_{basic}, w_{int}, w_{adv}, w_{mid}, w_{final}; f_i = 1, h_i),$$

$$e_i^{int,*}(f_i = 1) = e_i^{int,*}(w_{int}, w_{adv}, w_{final}; f_i = 1, L_i^{basic}),$$

$$e_i^{adv,*}(f_i = 1) = e_i^{adv,*}(w_{adv}, w_{final}; f_i = 1, L_i^{int}).$$

(See Section 7.2 for a fuller discussion of how myopic and forward-looking students choose effort by balancing the benefits and costs of effort exertion.) As the course includes a fixed number of instructional videos and assignment problems at each learning stage, there is a natural upper bound on effort exertion for the average student.[25] In the counterfactual simulations, the students can only exert effort up to a reasonable amount of study time $\bar{e}^t$, where the upper bound is inferred from the administrative data. That is, the model implied effort at each learning stage is $\max\{e_i^{t,*}, \bar{e}^t\}$ for $t \in \{basic, int, adv\}$. Then the optimal assignment weights are determined through the following iterative routine:

1. Start with an initial value of $w_{assign} = (w_{basic}, w_{int}, w_{adv})$ and set $p_m = 0.68$.

2. Compute effort $e_i^{t,*}(w_{assign})$ for myopic ($f_i = 0$) and forward-looking ($f_i = 1$) students for each stage $t \in \{basic, int, adv\}$.

3. Use effort allocations to compute learning $L_i^{final}(e_i^{adv,*}; e_i^{int,*}, e_i^{basic,*}, \bar{h})$ for myopic ($f_i = 0$) and forward-looking ($f_i = 1$) students.

[25] It is possible for a student to watch videos repeatedly, however only a few students watch all the videos more than once.

4. Use effort allocations to compute total cost of effort $\sum_t C(e_i^{t,*})$ for myopic and forward-looking students.

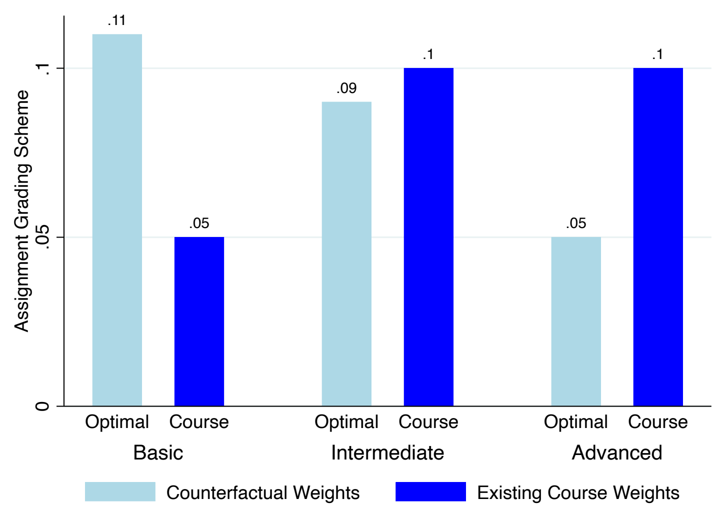
5. Evaluate instructor's objective function using knowledge accumulated in the advanced stage and the total effort cost as computed in steps 3 and 4, respectively.

6. Update w_{assign} to w'_{assign} using Newton's method to take the next step.

7. Iterate through steps 2-6 until convergence.

The above procedure is repeated for numerous initial values of w_{assign}. The simulated assignment weights that maximize learning of a representative student net of effort costs are then selected.

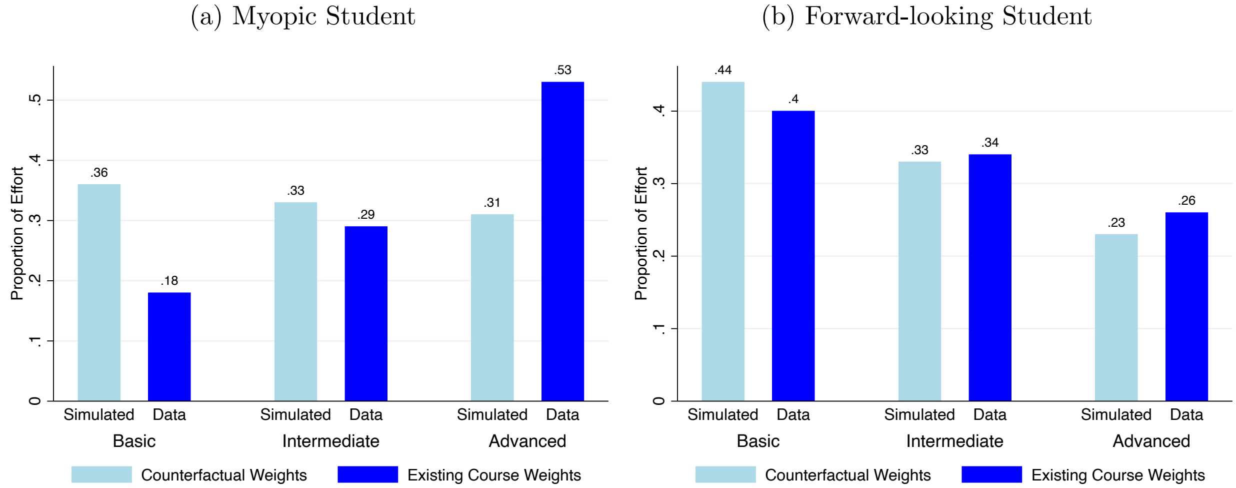
Figure 5: Optimal Assignment Weights



Notes: The figure presents the optimal assignment weights together with the actual course grading scheme. Within each learning stage, the assignment weights are shown in the following order: (1) optimal simulated weights, and (2) the actual course assignment weights.

Simulation of Optimal Assignment Weights. Figure 5 compares the optimal assignment weights with the existing course grading scheme. In contrast to the existing grading scheme, the optimal counterfactual grading scheme weights are decreasing across the learning stages: $w_{basic} = 0.11, w_{int} = 0.09, w_{adv} = 0.05$. The optimal grading scheme is expected to increase myopic students' final exam grade by 2.5 pp ($0.11\sigma^{**}$) while having no significant impact on the forward-looking students. Under the optimal grading scheme, myopic students are incentivized to front-load their effort allocation, while mitigating distortions introduced to the optimal effort allocation of forward-looking students. The following figure shows the proportion of effort exerted by the representative myopic and forward-looking students under the optimal grading scheme.

Figure 6: Allocation of Effort Across Learning Stages by Grading Scheme



Notes: The figure presents the relative distribution of total study time allocation across the learning stages. Effort allocation under the optimal grading scheme and the actual course grading weights are presented separately for myopic (left) and forward-looking (right) students.

Figure 6a clearly shows that under the simulated grading scheme a myopic student exerts a larger fraction of their total effort in the basic stage relative to the actual course grading scheme. Myopic students effort supply can be easily influenced in the basic stage by assigning it more weight. In contrast, a forward-looking student's effort supply in the basic stage is less malleable to changes in the assignment grading weights. This is because, the forward-looking student internalizes the cumulative course structure and the increased productivity of study time in later stages resulting from exerting effort in the basic stage. Figure 6b shows forward-looking students only slightly increase their effort in earlier stages under the front-loaded counterfactual grading scheme. Comparing Figures 6a and 6b also shows that forward-looking students in the actual course front-load their effort to a greater degree than myopic students under the simulated optimal grading scheme. The evidence thus suggests that student learning can be further improved if both myopic and forward-looking students are incentivized to allocate their type-specific effort most efficiently through receiving personalized grading schemes.

9.2 Personalizing Optimal Assignment Weights by Student Type

Setup of the Personalized Assignment Weights Counterfactual. Next I consider a counterfactual where myopic and forward-looking students receive personalized assignment grading weights that best incentivize them to allocate effort throughout the course. That is, the instructor designs a separate grading scheme for a representative myopic and forward-looking student, respectively. Then the instructor sets the assignment grading scheme by optimizing the following objective:

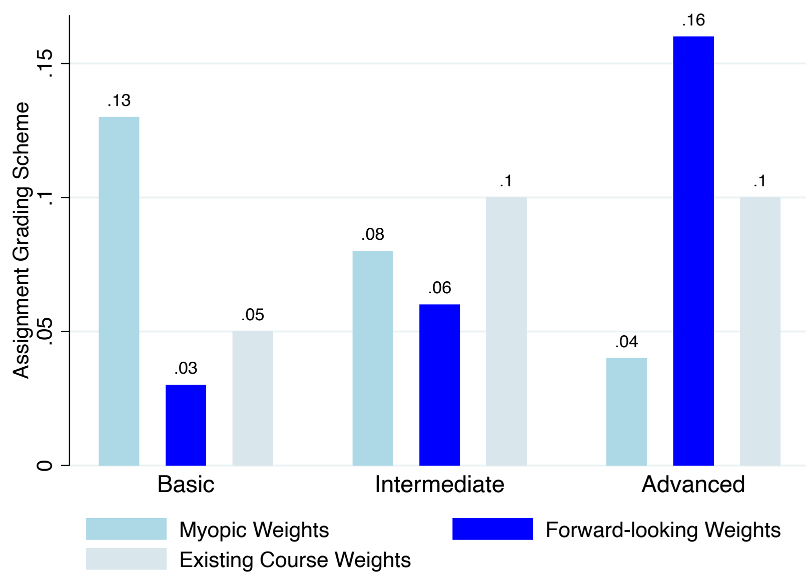
$$\max_{w_{basic}, w_{int}, w_{adv}} L_i^{final}(e_i^{adv,*}; e_i^{int,*}, e_i^{basic,*}, \bar{h}, f_i) - \sum_t C(e_i^{t,*}; \bar{E}, f_i)$$

$$\text{subject to } w_{basic} + w_{int} + w_{adv} = 0.25,$$

for myopic ($f_i = 0$) and forward-looking ($f_i = 1$) students. The optimization procedure outlined for the previous counterfactual is repeated separately for both type of students using this objective function to determine the personalized assignment grading weights.

Personalized Assignment Weights Simulation. The figure below compares the optimal personalized assignment grading weights with the existing course grading scheme.

Figure 7: Personalized Optimal Assignment Weights

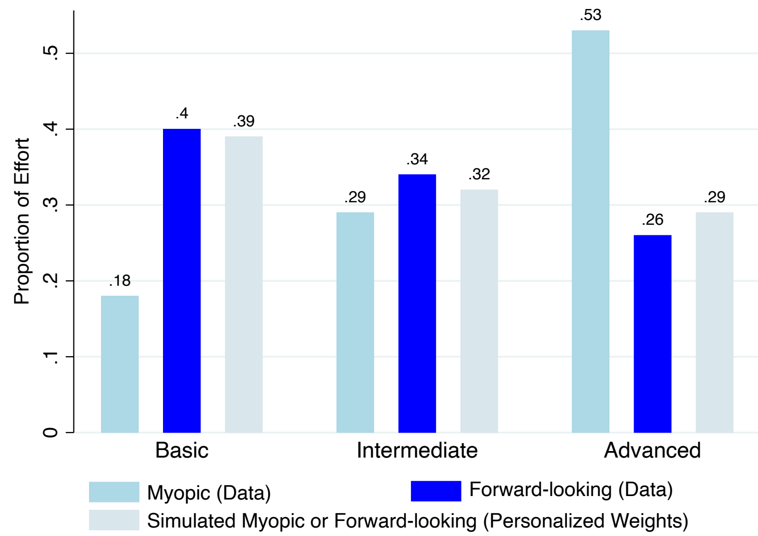


Notes: The figure presents the personalized assignment weights together with the actual course grading scheme. Within each learning stage, the assignment weights are shown in the following order: (1) optimal weight for a myopic student, (2) optimal weight for a forward-looking student, and (3) the actual course assignment weight.

For the myopic student, the optimal grading schemes is decreasing across the learning stages: $w_{basic} = 0.13, w_{int} = 0.08, w_{adv} = 0.04$. In contrast, for the forward-looking student, the assignment weights are sharply increasing across the course: $w_{basic} = 0.03, w_{int} = 0.06, w_{adv} = 0.16$. Relative to the actual course grading scheme, the personalized weights for myopic students increase their final exam grades by 2.8 pp ($0.13\sigma^{***}$). Forward-looking students' final grades also increase by around

1.5 pp, but this change is not statistically significant. Overall, the personalized optimal grading schemes improve students' final exam grades on average by 2.2 pp ($0.1\sigma^{**}$). The following figure shows the distribution of total effort allocation under the personalized assignment weights.

Figure 8: Effort Allocation Under Personalized Optimal Assignment Weights

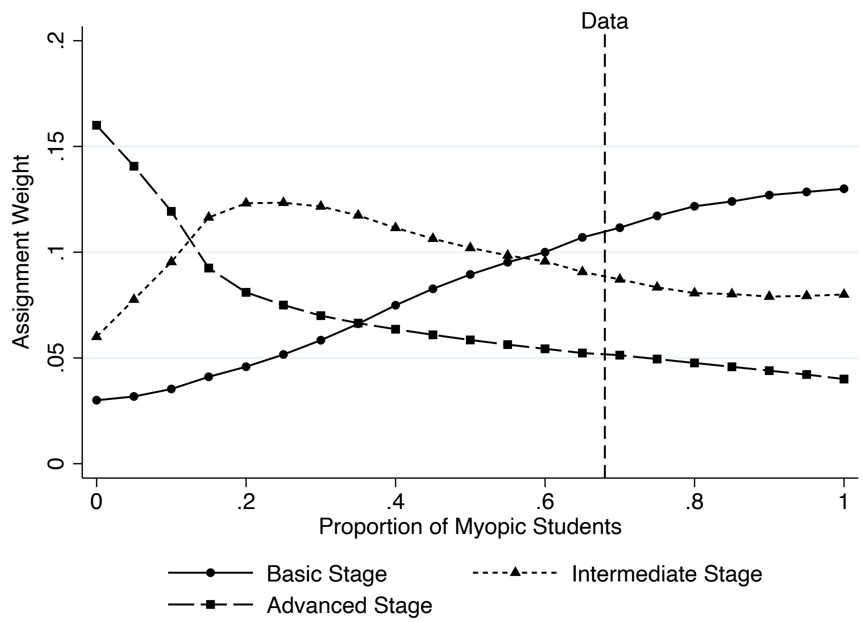


Notes: The figure presents the relative distribution of total study time across the learning stages for different type of students and grading incentives. Within each learning stage, the proportion of effort allocated is shown in the following order: (1) myopic student under the actual course grading scheme, (2) forward-looking student under the actual course grading scheme, and (3) both student types under their personalized grading scheme.

In the simulation, both representative myopic and forward-looking students have average baseline knowledge, average English proficiency, and are in the same course. Therefore the actual marginal benefit and cost of effort are identical across both types of student, but effort allocation can vary as the myopic student does not internalize the cumulative course structure. Then the efficient distribution of total effort across stages of the course is identical by student type. The simulation indicates that the personalized assignment weights induce both myopic and forward-looking students to identically distribute effort as presented in Figure 8. That is, the instructor can design a personalized grading scheme so that the representative myopic and forward-looking students allocate effort most effectively in the course. Relative to the actual course grading scheme, Figure 8 shows that the effort allocation for a representative myopic student is drastically different under the personalized grading scheme. In contrast, the representative forward-looking student distributes effort close to the optimal allocation even under the actual course grading scheme.

Optimal Assignment Weights by Proportion of Myopic Students. The personalized assignment weights are informative of a universal course grading scheme when all students in the course are either myopic or forward-looking. I will now generate the optimal grading scheme for any distribution of myopic and forward-looking students. To do so, I simulate the optimal assignment weights as the proportion of myopic students increases from 0 to 1. The following figure shows how the optimal grading weights transition from being loaded on the advanced stage, intermediate stage, and then the basic stage as the share of myopic students increases.

Figure 9: Optimal Assignment Weights by Share of Myopic Students



Notes: The figure presents the optimal assignment weight for each learning stage by 0.05 increments in the proportion of myopic students. The intersection of the assignment weights with the dashed vertical line indicates the optimal grading scheme in the actual course with a 68% share of myopic students.

Figure 9 shows that as the share of myopic students in a course increases from 0 to 20%, there is a rapid decline in the weight assigned to the advanced stage and a gradual increase in the basic assignment weight. The corresponding steep increase in the intermediate assignment weight suggests that the advanced stage weight is substituted more on the intermediate stage than the basic stage. That is, for a course with a modest 20% share of myopic students, the optimal assignment weights are loaded towards the middle of the course. Assigning the majority of the weight to the middle of the course caters to both myopic and forward-looking students. The myopic students are incentivized to exert effort prior to entering the advanced stage, while mitigating losses in learning to forward-looking students by deviating the incentives more on the intermediate rather than the basic stage. Next, as the share of myopic students exceeds 55% as observed in the actual course, the assignment weights are loaded most to the basic stage. As a result, front-loading the assignment weights is optimal when a modest majority of students are myopic.

10 Conclusion

The share of students pursuing education online has been increasing rapidly over the past two decades. In an especially graphic way, over the past academic year, the vast majority of students and education providers around the world have experienced both the benefits and challenges of online learning, as educational institutions have been forced to roll-out online frameworks in response to the COVID-19 pandemic. Although online learning is flexible and easy to access for students, the learning challenges that arise can outweigh the obvious benefits. In particular, students in online courses find it difficult to obtain individualized support or to stay engaged with online course work. In light of these well-known challenges, this paper presented a microeconomic approach to course design to address the twin needs of providing personalized support and encouraging students to stay engaged with course work.

The course-design approach I developed involves several components. I began by assessing the efficacy of two prominent features of online courses: online peer discussion boards and online assignments. To do so, I employed rich administrative data from a large pre-existing foundational online STEM course that has a cumulative structure. I then conducted two randomized interventions that were successful in nudging students to more fully utilize the discussion board and complete more online assignments. The administrative dataset, including precise measures of student study time allocation, together with the two field experiments were then used to estimate a behavioural model of student effort supply. Exogenous experimental variation arising from the two field experiments served to credibly identify the marginal benefit and cost to effort at each stage of the cumulative learning process built into the STEM course. The estimated model allowed me to conduct counterfactual experiments that would be difficult to implement in practice, such as randomly assigning observationally equivalent students to drastically varying grading schemes.

I presented three main findings that help to inform the design of large foundational online and traditional in-person courses. First, the discussion board serves as an effective tool for supporting student learning: an additional 5-weeks of utilization increases final exam grades by 0.08 SD. Second, completing low-stakes online assignments throughout the course is essential for student learning: spending an extra hour doing online assignments increases final exam grades by 0.09 SD (noting that online homework is the key means of learning in the course). Third, given a cumulative course structure and the presence of many myopic students, an instructor can usefully adjust the assignment grading weights. That is, in contrast to the actual equally-weighted course grading scheme, I find that the simulated assignment weights that maximize learning are decreasing as the course progresses. The optimal weights serve to encourage myopic students to adequately invest effort to learn foundational skills early in the course, increasing their final exam grade by 0.11 SD.

While this paper considered an online STEM course, the analysis can be adapted to many traditional in-person course settings. The paper's findings contribute to our understanding of how to support the learning of students at scale, and also to design incentives that induce heterogenous students to allocate their effort more efficiently; heterogeneity in the student population arises naturally in large courses. Additionally, students within large foundational courses which typically have a cumulative structure can find it difficult to obtain individualized support. The microeconomic course-design approach considered in this paper helps to alleviate these widespread challenges, having broader applicability to other online and traditional course settings.

References

Aizer, Anna and Flavio Cunha (2012) "The production of human capital: Endowments, investments and fertility," Working Paper 18429, National Bureau of Economic Research, http://www.nber.org/papers/w18429.

Allcott, Hunt and Dmitry Taubinsky (2015) "Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market," *American Economic Review*, 105 (8), 2501–38.

Andrews, Isaiah, Matthew Gentzkow, and Jesse M Shapiro (2020) "Transparency in structural research," *Journal of Business & Economic Statistics*, 38 (4), 711–722.

Angrist, Joshua, Philip Oreopoulos, and Tyler Williams (2014) "When opportunity knocks, who answers? New evidence on college achievement awards," *Journal of Human Resources*, 49 (3), 572–610.

Artés, Joaquín and Marta Rahona (2013) "Experimental evidence on the effect of grading incentives on student learning in Spain," *The Journal of Economic Education*, 44 (1), 32–46.

Augenblick, Ned (2018) "Short-Term Time Discounting of Unpleasant Tasks," working paper, http://faculty.haas.berkeley.edu/ned/Augenblick_ShortTermDiscounting.pdf.

Baer, Amy and Andrew DeOrio (2020) "A longitudinal view of gender balance in a large computer science program," in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, 23–29.

Barrow, Lisa and Cecilia Elena Rouse (2018) "Financial incentives and educational investment: The impact of performance-based scholarships on student time use," *Education Finance and Policy*, 13 (4), 419–448.

Bernheim, B Douglas, Stefano DellaVigna, and David Laibson (2019) *Handbook of Behavioral Economics-Foundations and Applications 2*: Elsevier.

Bettinger, Eric P, Lindsay Fox, Susanna Loeb, and Eric S Taylor (2017) "Virtual classrooms: How online college courses affect student success," *American Economic Review*, 107 (9), 2855–75.

Bowen, Jose Antonio (2012) *Teaching naked: How moving technology out of your college classroom will improve student learning*: John Wiley and Sons.

Bueno, Carycruz (2020) "Bricks and mortar vs. computers and modems: The impacts of enrollment in K-12 virtual schools," *Computers and Modems: The Impacts of Enrollment in K-12 Virtual Schools (July 3, 2020)*.

Card, David, Stefano DellaVigna, and Ulrike Malmendier (2011) "The role of theory in field experiments," *Journal of Economic Perspectives*, 25 (3), 39–62.

Clark, Damon, David Gill, Victoria Prowse, and Mark Rush (2020) "Using goals to motivate college students: Theory and evidence from field experiments," *Review of Economics and Statistics*, 102 (4), 648–663.

Coates, Dennis, Brad R Humphreys, John Kane, and Michelle A Vachris (2004) "No significant distance between face-to-face and online instruction: Evidence from principles of economics," *Economics of Education Review*, 23 (5), 533–546.

Damgaard, Mette Trier and Helena Skyt Nielsen (2018) "Nudging in education," *Economics of Education Review*, 64, 313–342.

Deming, David J, Claudia Goldin, Lawrence F Katz, and Noam Yuchtman (2015) "Can online learning bend the higher education cost curve?" *American Economic Review*, 105 (5), 496–501.

Figlio, David, Mark Rush, and Lu Yin (2013) "Is it live or is it internet? Experimental estimates of the effects of online instruction on student learning," *Journal of Labor Economics*, 31 (4), 763–784.

Fryer Jr, Roland G (2011) "Financial incentives and student achievement: Evidence from randomized trials," *The Quarterly Journal of Economics*, 126 (4), 1755–1798.

Gilraine, Michael (2016) "School accountability and the dynamics of human capital formation," http://tinyurl.com/Gilraine-JMP.

Grodner, Andrew and Nicholas G Rupp (2013) "The role of homework in student learning outcomes: Evidence from a field experiment," *The Journal of Economic Education*, 44 (2), 93–109.

Harackiewicz, Judith M and Stacy J Priniski (2018) "Improving student outcomes in higher education: The science of targeted intervention," *Annual review of psychology*, 69, 409–435.

Kizilcec, Rene F, Justin Reich, Michael Yeomans et al. (2020) "Scaling up behavioral science interventions in online education," *Proceedings of the National Academy of Sciences*, 117 (26), 14900–14905.

Latif, Ehsan and Stan Miles (2020) "The Impact of Assignments and Quizzes on Exam Grades: A Difference-in-Difference Approach," *Journal of Statistics Education*, 28 (3), 289–294.

Macartney, Hugh, Robert McMillan, and Uros Petronijevic (2021) "A Quantitative Framework for Analyzing the Distributional Effects of Incentive Schemes," Working Paper 28816, National Bureau of Economic Research, https://www.nber.org/papers/w28816.

Oreopoulos, Philip, Richard W Patterson, Uros Petronijevic, and Nolan G Pope (2018) "When studying and nudging don't go as planned: Unsuccessful attempts to help traditional and online college students," Working Paper 25036, National Bureau of Economic Research, https://www.nber.org/papers/w25036.

Pozo, Susan and Charles A Stull (2006) "Requiring a math skills unit: Results of a randomized experiment," *American Economic Review*, 96 (2), 437–441.

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

Tabord-Meehan, Max (2018) "Stratification trees for adaptive randomization in randomized controlled trials," *arXiv preprint arXiv:1806.05127*, https://arxiv.org/abs/1806.05127.

Taubinsky, Dmitry and Alex Rees-Jones (2018) "Attention variation and welfare: theory and evidence from a tax salience experiment," *The Review of Economic Studies*, 85 (4), 2462–2496.

Todd, Petra E and Kenneth I Wolpin (2007) "The production of cognitive achievement in children: Home, school, and racial test score gaps," *Journal of Human capital*, 1 (1), 91–136.

Tables

Table 1: Student Level Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel A: Demographics					
I(Female)	0.491	0.499	0	1	3686
I(First year of university)	0.618	0.485	0	1	3686
I(Domestic student)	0.531	0.499	0	1	3686
I(Speaks English at home)	0.286	0.452	0	1	3686
I(First generation university)	0.173	0.378	0	1	3686
I(Mother at least college graduate)	0.668	0.459	0	1	3686
I(Father at least college graduate)	0.700	0.458	0	1	3686
Panel B: Characteristics					
I(Has some prior programming experience)	0.134	0.341	0	1	3686
I(Course required for major)	0.736	0.441	0	1	3686
I(Pursuing STEM major)	0.717	0.448	0	1	3686
Panel C: Behavioural Characteristics					
I(Student is attentive)	0.762	0.426	0	1	3686
I(Student is forward-looking)	0.315	0.465	0	1	3686
Panel D: Online Homework Participation					
I(Started weekly homework)	0.843	0.367	0	1	3686
I(Completed weekly homework)	0.671	0.471	0	1	3686
Weekly unique minutes of videos watched	24.52	9.211	0	38.35	3686
Weekly minutes spent doing problems	122.49	63.107	0	434.13	3686
Panel E: Discussion Board Participation					
I(Registered for course discussion board)	0.791	0.407	0	1	3686
No. of total contributions	3.57	14.087	0	237	2911
No. of unique posts viewed	121.88	149.28	0	1022	2911
Weekly minutes spent on discussion board	28.61	14.51	0	187.66	2911

Notes: Table presents descriptive statistics related to student demographic and characteristics, discussion board participation and online homework activity. Statistics shown in Panel A and B are formulated using self-reported student responses on the baseline survey (see Appendix C.1). Panel C uses survey data to characterize students as attentive or forward-looking (see Appendix C.2). Panel D statistics are formulated from the administrative data of the online homework platform. Finally, the statistics shown in Panel E are computed using data gathered from the discussion board.

Table 2: Assignment of Students to Sign-Up Activity Control and Treatment Groups

Assignment Group	Number of Students	Percent of Students
Control	1673	49.8
Instructions only	560	16.7
Information only	558	16.6
Both instructions and information	563	16.9

Table 3: Student Attrition and Treatment Allocation

	(1) I(Dropped course)	(2) I(Dropped course)	(3) I(Dropped course)	(4) I(Dropped course)
I(Assignment to sign-up activity)	-0.0125		-0.0103	
	(0.0270)		(0.0183)	
No. of reminder messages received		0.0116		-0.0061
		(0.0898)		(0.0192)
Controls	No	No	Yes	Yes
No. of Students	3712	4091	3712	4091
R-squared	0.0019	0.0015	0.13	0.13

Notes: Table shows differences in attrition rate across the control and treatment group for both interventions considered in this study. Controls include pre-treatment student demographics and characteristics included in Panels A and B of Table 1, and cohort fixed effects. Only students who did not register for the discussion board prior to sign-up activity were eligible to receive the treatment. Significance levels are represented by $^*p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

Table 4: Effect of Sign-up Activity on Discussion Board Utilization

	(1) I(Signed Up)[a]	(2) I(Signed Up)	(3) Weeks Registered[b]	(4) Weeks Registered
I(Receives sign-up activity)	0.174***	0.181***	4.532***	4.441***
	(0.0216)	(0.0223)	(0.2234)	(0.2135)
Control Mean	0.69	0.69	2.913	2.913
Controls	No	Yes	No	Yes
Adjusted R-square	0.119	0.223	0.216	0.331
F-stat for treatment	64.89	62.93	412.02	414.41
No. of Students	3354	3354	3354	3354

Notes: [a]The outcome variable is whether students signed up for the course discussion board by end of the course. [b]Outcome is the number of weeks the student is registered for the discussion board. Student demographics and other characteristics include variables present in Panel A and B in Table 1 respectively. Controls also include cohort fixed effects. Only students who had not registered for the discussion board prior to the baseline survey were eligible for treatment. Significance levels are represented by $^*p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

Table 5: Effect of Discussion Board Utilization on Learning Outcomes (2SLS)

	(1) Exam Performance[a]	(2) Exam Performance	(3) Homework Performance[b]	(4) Homework Performance
No. of weeks registered	0.014^{**}	0.016^{**}	0.025^{***}	0.028^{***}
	(0.0066)	(0.0068)	(0.0088)	(0.0086)
Controls	No	Yes	No	Yes
Adjusted R-square	0.091	0.215	0.104	0.362
No. of Students	3354	3354	3354	3354

Notes: [a]Standardized final exam grade. [b]Standardized average performance across all online homework. Assignment to the sign-up activity is used as an instrument for the number of weeks registered. Controls include pre-treatment student demographic and characteristics included in Panels A and B of Table 1, and cohort fixed effects. Significance levels are represented by $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

Table 6: Homework Reminders and Online Homework Participation

	(1) Homework Completion[a]	(2) Homework Completion	(3) Study Time[b]	(4) Study Time
No. of reminders received	0.184^{***}	0.187^{***}	23.83^{***}	22.71^{***}
	(0.0126)	(0.0128)	(2.1663)	(2.0783)
Controls	No	Yes	No	Yes
F-stat for treatment	213.15	217.61	121.31	120.24
Adjusted R-square	0.125	0.358	0.133	0.326
No. of Students	3686	3686	3686	3686

Notes: [a]Homework completion is defined as the student attempting all problems with a positive score. [b]Total minutes spent watching videos and working on homework problems. Students can receive at most 10 reminder messages. Controls include pre-treatment student demographic and characteristics included in Panels A and B of Table 1, and cohort fixed effects. Significance levels are represented by $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

Table 7: Student Online Learning Participation and Final Exam Grade (2SLS)

	(1) Exam Performance[a]	(2) Exam Performance	(3) Exam Performance	(4) Exam Performance
Homework Completion[b]	0.181***	0.176***		
	(0.0540)	(0.0516)		
Total Study Time (Hours)[c]			0.084***	0.091***
			(0.0312)	(0.0322)
Controls	No	Yes	No	Yes
Adjusted R-square	0.092	0.262	0.083	0.241
No. of Students	3686	3686	3686	3686

Notes: [a]Standardized final exam grade. [b]Homework completion is defined as the student attempting all problems with a positive score. [c]Time spent watching videos and working on homework problems. Total number of reminders received is used as an instrument for homework participation. Controls include pre-treatment student demographic and characteristics included in Panels A and B of Table 1, and cohort fixed effects. Significance levels are represented by $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

Table 8: Efficacy of Nudges and Study Group Involvement

	(1) I(Signed Up)[a]	(2) I(Signed Up)	(3) Homework Completed[b]	(4) Homework Completed
I(Study group) × I(Receives sign-up activity)	-0.0217*	-0.0193*		
	(0.0126)	(0.0104)		
I(Study group) × No. of reminders			0.061	0.043
			(0.0537)	(0.0317)
Controls	No	Yes	No	Yes
Adjusted R-square	0.132	0.243	0.142	0.371
No. of Students	3354	3354	3686	3686

Notes: Indicator for whether a student is in a study group and treatment status are also included in the estimation. Student demographics and other characteristics include variables present in Panel A and B in Table 1 respectively. Only students who had not registered for the discussion board prior to the baseline survey were eligible for sign-up activity. Significance levels are represented by $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

Table 9: Benefit of Effort Parameter Estimates (2SLS)

Parameter	Estimate (SE)
Panel A: Basic Learning Stage Performance	
$\widehat{\alpha_1}$ (minutes study)	0.00241***
	(0.00061)
$\widehat{\alpha_2}$ (baseline knowledge)	0.114**
	(0.0553)
$\widehat{\alpha_3}$ (basic effort $\times$ baseline knowledge)	0.00126*
	(0.00072)
Panel B: Midterm Performance	
$\widehat{\theta_1}$ (basic minutes study)	0.00221***
	(0.00059)
$\widehat{\theta_2}$ (baseline knowledge)	0.093*
	(0.0492)
$\widehat{\theta_3}$ (basic effort $\times$ baseline knowledge)	0.00214
	(0.00358)
Panel C: Intermediate Learning Stage Performance	
$\widehat{\beta_1}$ (minutes study)	0.00212***
	(0.00052)
$\widehat{\beta_2}$ (basic knowledge)	0.166***
	(0.0442)
$\widehat{\beta_3}$ (int. effort $\times$ basic knowledge)	0.00111**
	(0.00054)
Panel D: Advanced Learning Stage Performance	
$\widehat{\lambda_1}$ (minutes study)	0.00178**
	(0.00087)
$\widehat{\lambda_2}$ (intermediate knowledge)	0.183***
	(0.0441)
$\widehat{\lambda_3}$ (adv. effort $\times$ int. knowledge)	0.00137**
	(0.00069)
Panel E: Final Exam Performance	
$\widehat{\pi_1}$ (adv. minutes study)	0.00189**
	(0.00095)
$\widehat{\pi_2}$ (intermediate knowledge)	0.155***
	(0.0418)
$\widehat{\pi_3}$ (adv. effort $\times$ int. knowledge)	0.00123**
	(0.00061)
No. of students	3686

Notes: All assessment performances are standardized. Baseline knowledge is a standardized measure that aggregates prior programming knowledge and cGPA. The number of reminders received at each stage are the instrumental variables for study time and prior stage knowledge. Significance levels are represented by $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

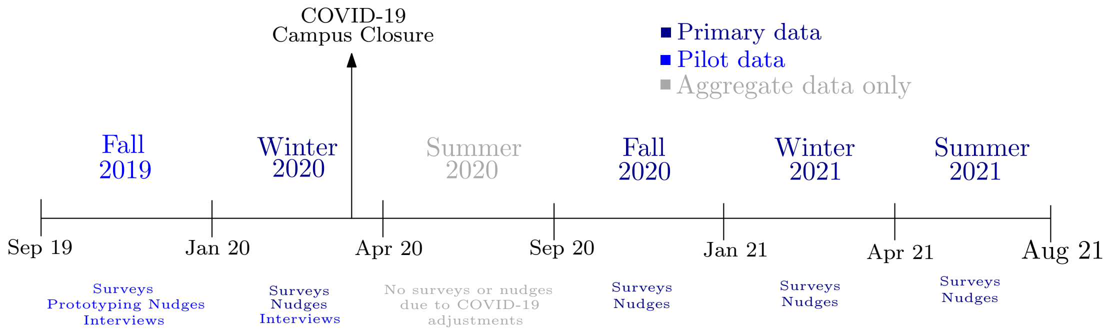
Table 10: Cost of Effort Parameter Estimates (MLE)

Parameter	Estimate (SE)
Panel A: Basic Learning Stage	
$\widehat{1+\gamma_1}$ (Convexity)	2.1367***
	(0.5014)
$\widehat{\kappa_1}$ (Steepness)	0.00038***
	(0.00014)
Panel B: Intermediate Learning Stage	
$\widehat{1+\gamma_2}$	2.3531***
	(0.4424)
$\widehat{\kappa_2}$	0.00038***
	(0.00013)
Panel C: Advanced Learning Stage	
$\widehat{1+\gamma_3}$	2.5121***
	(0.5430)
$\widehat{\kappa_3}$	0.00039***
	(0.00014)
No. of students	3686

Notes: Performance within each stage is standardized. Cost parameters estimated using maximum likelihood estimation. Standard errors appear in parentheses and are calculated using the outer-product of gradients method. Significance levels are represented by $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

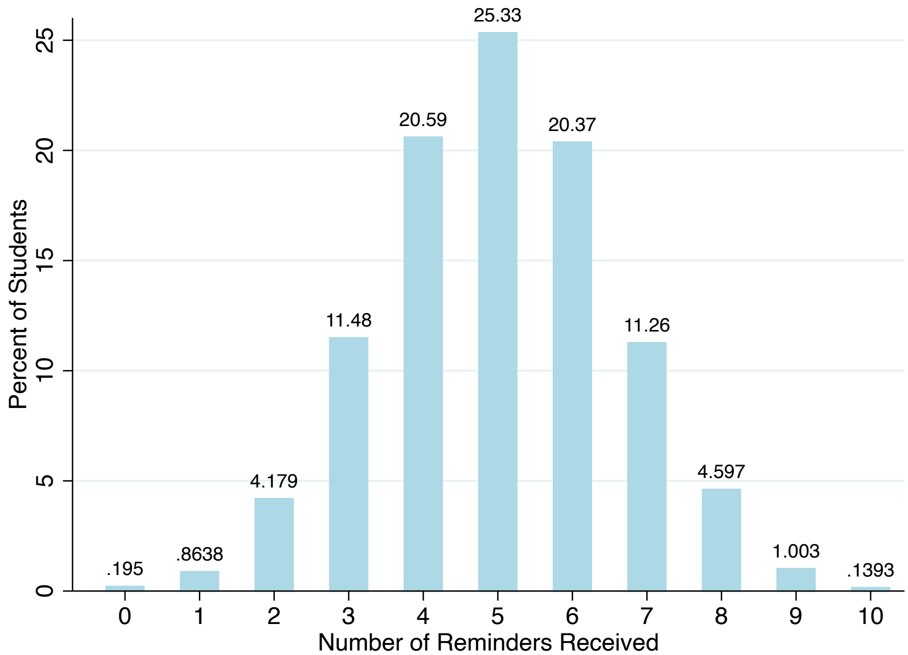
Figures

Figure 10: Timeline of Data Collection



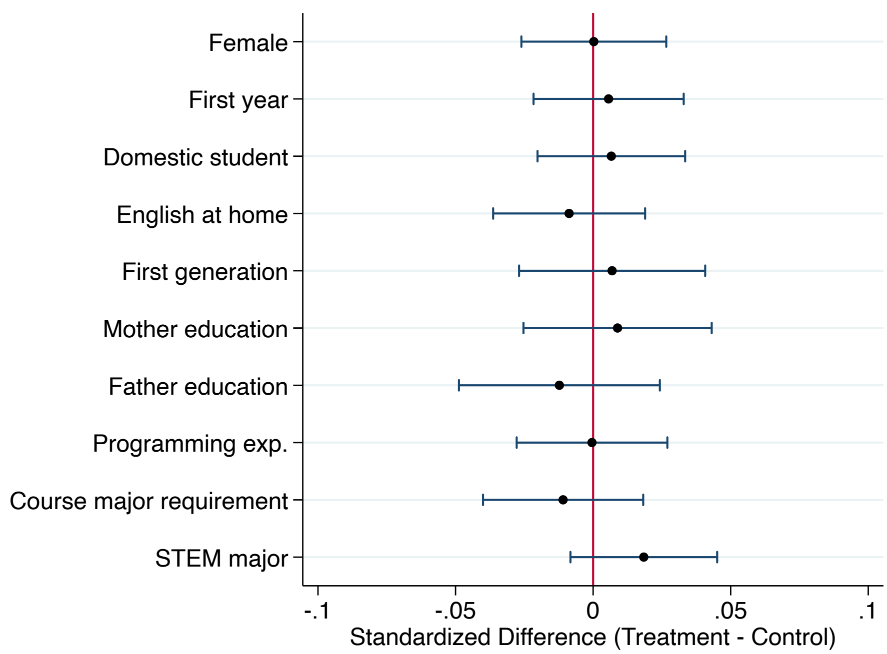
Notes: The figure illustrates the timeline of the data collection. The primary data collection is gathered from the Winter 2020, Fall 2020, Winter 2021, and Summer 2021 cohorts. Pilot data is gathered from the Fall 2019 cohort and involved conducting interviews with students and instructors, surveying students, and prototyping interventions.

Figure 11: Distribution of Homework Reminders Received



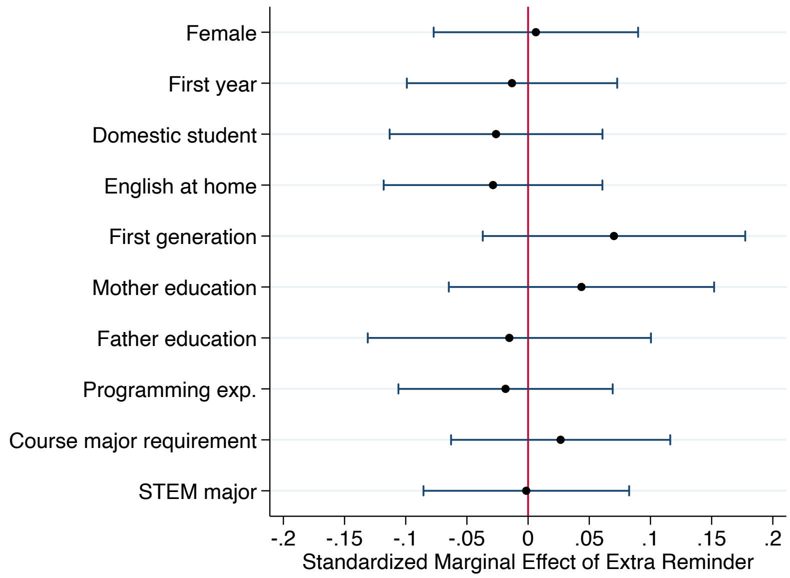
Notes: The figure shows the distribution of reminders received by students. For each of the 10 graded homework assessments, half of the students are randomly selected to receive a reminder. Then each student is eligible to receive between 0 and 10 homework reminders in total.

Figure 12: Student Demographic and Characteristics Balance Check for Sign-up Activity



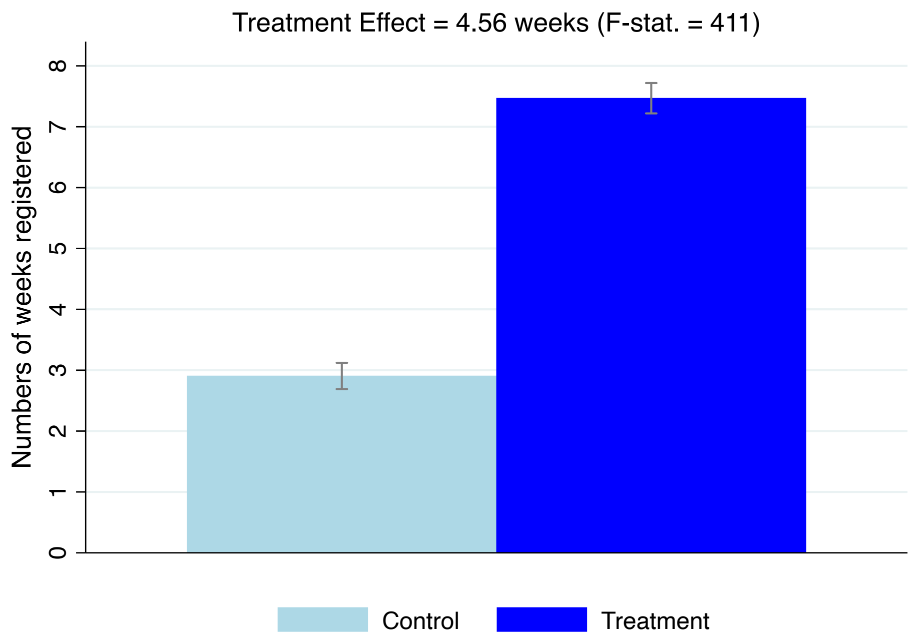
Notes: The figure reports the mean differences in the baseline student demographics and characteristics across the control and treatment groups for the sign-up activity. The estimates displayed are computed by regressing each standardized variable presented on the vertical axis on an indicator for receiving the sign-up activity. The error bars represent 95% confidence intervals.

Figure 13: Student Demographic and Characteristics Balance Check for Homework Reminders



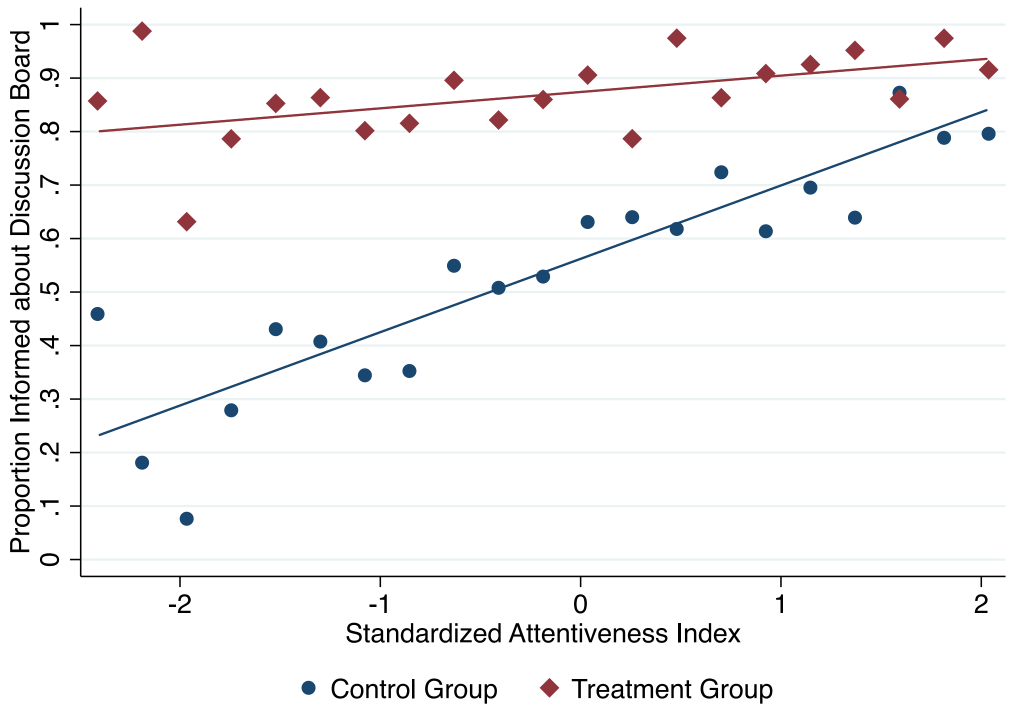
Notes: The estimates displayed are computed by regressing each standardized variable presented on the vertical axis on the number of reminders received. The error bars represent 95% confidence intervals.

Figure 14: Sign-up Activity and Discussion Board Utilization



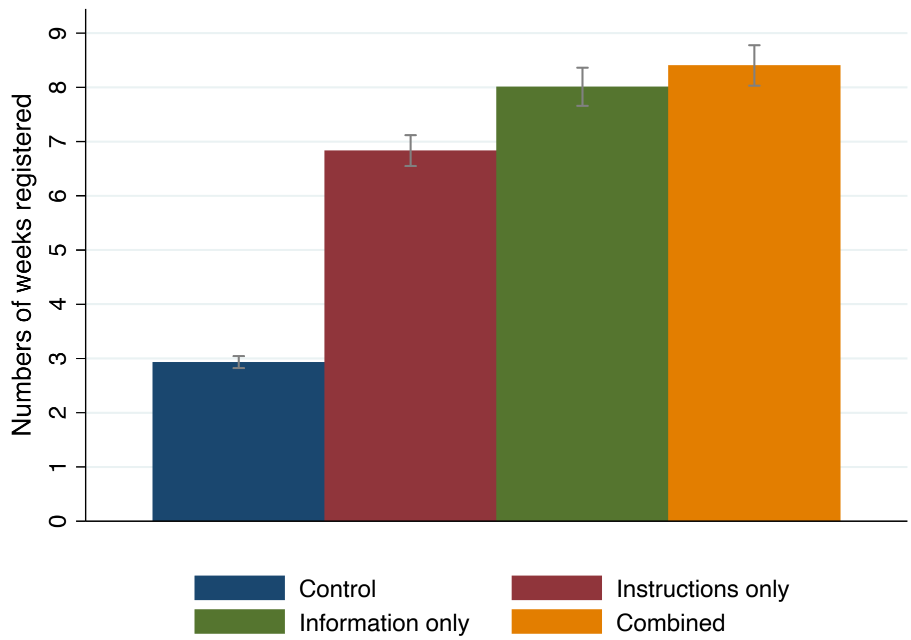
Notes: The figure shows the average number of weeks students are registered on the discussion board across the control (left) and treatment (right) group. The maximum number of weeks a student can be registered is 12-weeks. The treatment group consists of the students who were randomly assigned to receive the sign-up activity near the start of the course. The error bars represent the 95% confidence intervals.

Figure 15: Sign-up Activity and Being Informed About Discussion Board



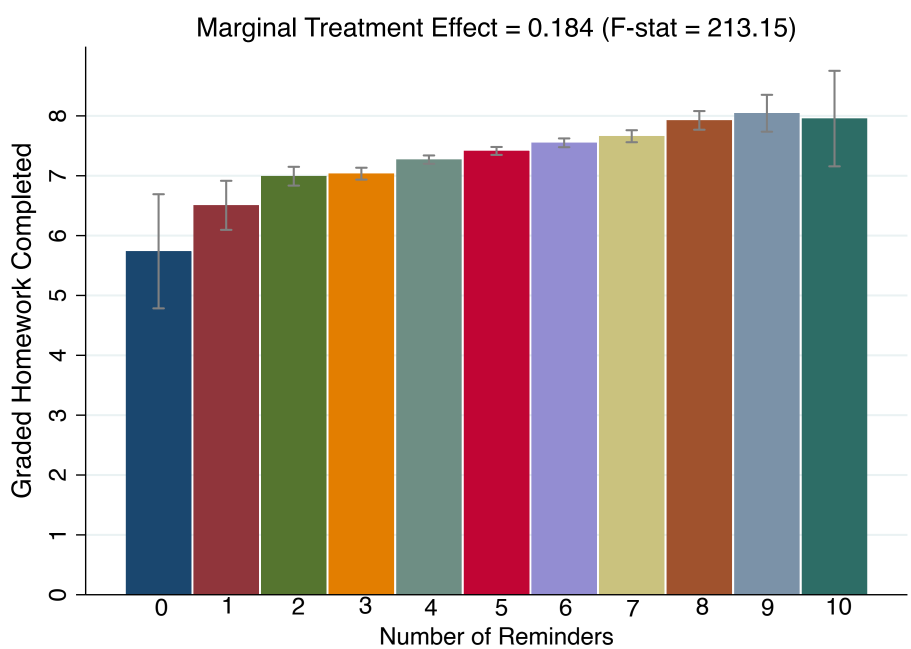
Notes: The figure presents a binned scatter plot showing the relationship between being informed about the discussion board and students' attentiveness by their sign-up activity treatment status. Students in the treatment group were randomly assigned to receive the sign-up activity near the start of the course. Whether a student is informed about the discussion board is inferred through a question on the baseline survey. The student attentiveness index is constructed using a series of survey questions (see Appendix C.2).

Figure 16: Variations of the Sign-up Activity and Discussion Board Utilization



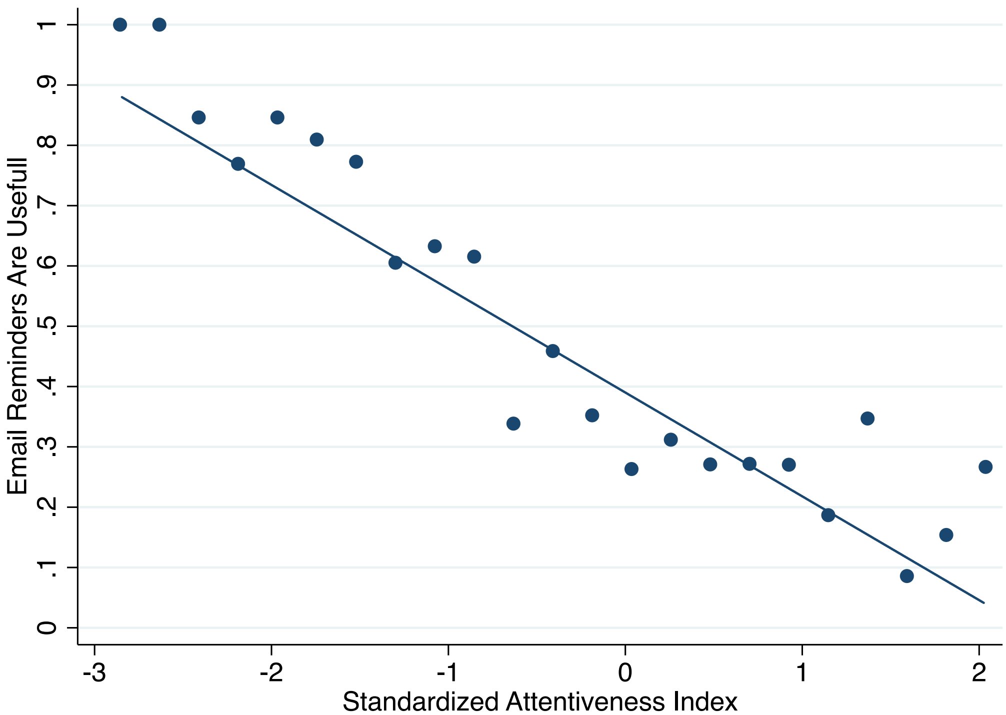
Notes: The figure shows the average number of weeks students are registered on the discussion board across the control and the variations of the sign-up activity. The maximum number of weeks a student can be registered is 12-weeks. The mean estimates are presented in the following order: (1) control, (2) instructions only, (3) information only, and (4) both instructions and information. The error bars represent the 95% confidence intervals.

Figure 17: Homework Completed and Reminder Messages



Notes: The figure shows the average number of online homework assessments completed by the number of reminders received. There are 10 graded homework assessments. The error bars represent the 95% confidence intervals.

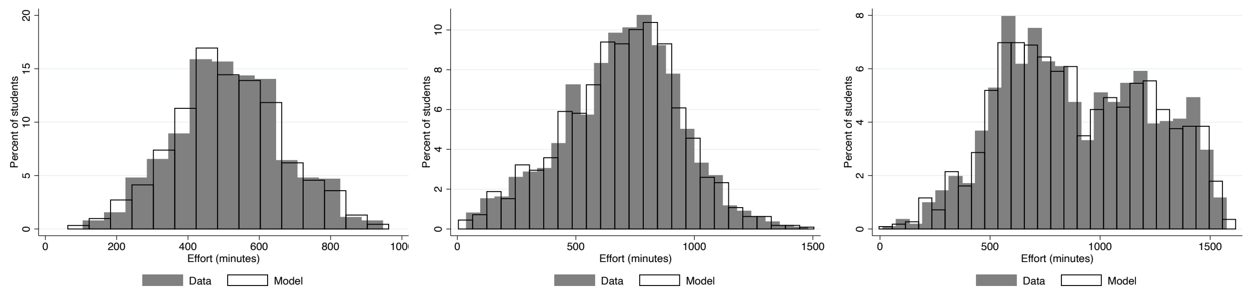
Figure 18: Reminder Messages Attentiveness Mechanism



Notes: The figure presents a binned scatter plot showing the relationship between finding the reminders helpful for keeping on track with homework and student attentiveness. Whether a student finds reminders to be useful is inferred from the survey data. The student attentiveness index is constructed using a series of survey questions (see Appendix C.2).

Figure 19: Distribution of Model Implied Study Time and Observed Study Time

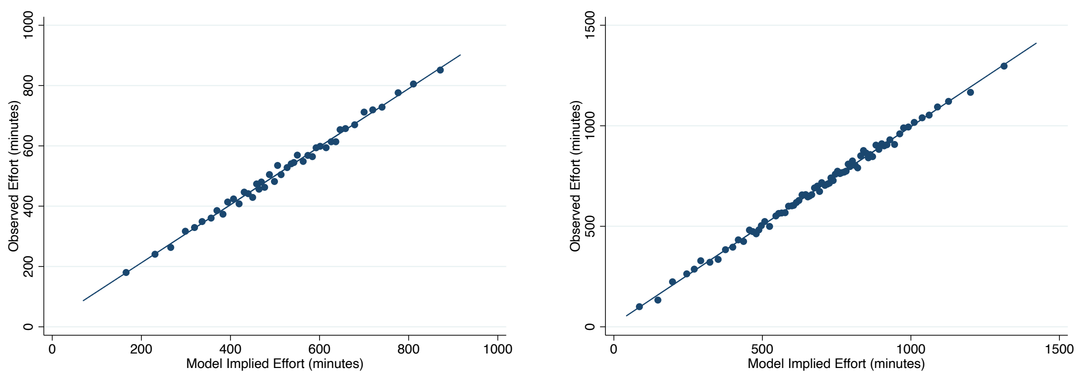
(a) Basic Stage (b) Intermediate Stage (c) Advanced Stage



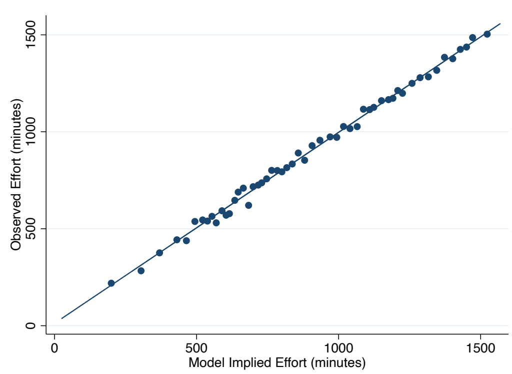
Notes: The figure presents the distribution of observed study time (grey) overlaid together with the distribution of model implied study time (white) at each learning stage.

Figure 20: Model Implied Study Time and Observed Study Time

(a) Basic Stage (b) Intermediate Stage



(c) Advanced Stage



Notes: The figure presents a binned scatter plot showing the relationship between observed study time (vertical axis) and model implied study time (horizontal axis).

A Appendix: Summary of Related Literature

A.1 Related Structural Behavioural Economics Literature

Table 11: Research Estimating Relevant Behavioural Parameters

Title	Authors (Date, Publisher)	Data	Research Design	Main Results
Short-Term Time Discounting of Unpleasant Tasks	Augenblick (2018, WP)	N = 79 subjects who make several decisions overtime resulting in 8875 observations.	Ask repeatedly within a week how many unpleasant real effort tasks they want to do. Task to transcribe blurry greek letters. Randomized wages and number of days from decision task is to be carried out.	Choose 42 tasks start time one week away, and drops to 40 (one day away), 38 (1 hour), and 36 (imminent) as deadline approaches. Estimate discount factor $\delta = 0.85$ when task 1-week away, and some present bias as $\beta \in [0.9, 0.95]$.
What Motivates Effort? Evidence and Expert Forecasts	Dellavigna and Pope (2018, ReStud)	Around N = 9861 Mturkers. Additionally had N = 213 experts make forecast for each treatment condition.	Mturkers randomly allocated to one of 18 treatment conditions ($N \approx 550$ each). Task was to press A-B consecutively on keyboard fast as possible within 10 minutes.	Psychological treatments more effective than baseline, but less impactful than monetary incentives. No evidence of present bias ($\beta = 1$) and altruism ($a = 0$).
Salience and Taxation: Theory and Evidence.	Chetty, Looney and, Kroft (2009, AER)	Focused on cosmetics, hair care accessories, and deodorant. Around 750 distinct products. Intervention was for three weeks.	Treatment group was tax-inclusive prices for the chosen toiletries within a store. Control group 1 are similar products like toothpaste within same aisle and store. Control group 2 are toiletries sold in other stores.	Estimate a degree of intention parameter of $\theta = 0.75$, where $\theta = 1$ is full awareness of the taxation. Although, the result is imprecise due to the quasi-experimental design that relies on week-to-week varition.
Evaluating Behaviourally Motivated Policy: Experimental Evidence from the Lightbulb Market.	Allcott and Taubinsky (2015, AER)	N = 1087 survey participants who were customers at the store.	Participants of iPad survey randomized into information treatment. Treatment group included energy costs for CFLs vs. incandescents. Control group did not get any information on energy costs.	Evidence for small inattention to energy savings with $\theta = 0.05$. Justifies a subsidy of 3 dollar per LED bulb due to inattention.
Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment.	Taubinsky and Rees-Jones (2018, ReStud)	N = 2998 online consumers purchase common household products.	Research participants given 20 dollars to buy one of the randomly chosen products. Randomize into 1) no tax, 2) standard sales tax, 3) triple the sales tax.	Find substantial inattention to taxes and vast heterogeneity in attentiveness. Individuals react to non-salient taxes as if they were 25% their size (i.e., $\theta = 0.25$). Less inattention of $\theta = 0.5$ when taxes are tripled.

A.2 Related Education Production Function Literature

Table 12: Research Exploring Education Production and Dynamic Complementarities

Title	Authors (Date, Publisher)	Data	Research Design	Main Results
The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps	Todd and Wolpin (2007, JHC).	N = 7700 individuals who are aged 14-21 in the NLSY79-CS.	Estimate cumulative production function using the mothers ability, child ability, and history of family and school inputs.	Lagged home inputs are significant predictors of present achievement. Overall, estimates suggest the learning process is cumulative.
The Technology of Skill Formation	Cunha and Heckman (2007, AER).	NA (Conceptual Framework)	Develop model of human capital accumulation which features dynamic complementarities in parental inputs.	Model suggests it is important to invest during early childhood stage (e.g. pre-school), more so than later stages (e.g. tuition reduction programs).
The Production of Human Capital: Endowments, Investments, and Fertility	Aizer and Cunha (2012, NBER WP).	N = 30,039 children from 1963 - 1970 whose mothers were involved in the National Collaborative Perinatal Project (NCPP).	Use introduction of Head Start in 1996 as instrument for investment.	Consistent with dynamic complementarities, authors find larger IQ gains from preschool for children with the highest stock of early human capital.
School Accountability and the Dynamics of Human Capital Formation	Gilraine (2018, Working Paper).	N = 3,310 school-year observations from public schools in North Carolina.	Leverages year-to-year variation in school accountability resulting from whether there are at least forty students belonging to a specific demographic group.	Author finds a 0.18σ increase in test scores for students who are in schools that were subject to school accountability in two consecutive periods relative to those in schools subject to accountability only in the previous period.
Does EdTech Substitute for Traditional Learning? Experimental Estimates of Educational Production Function	Bettinger et al. (2020, NBER WP).	N = 6253 grade 3 students in Russia. Teachers had access to computer assisted learning software to help students learn math and language by solving assigned problems.	Students randomized to 1) no computer assisted learning (control), 2) 45-minute computer assisted learning, 3) 90-minute computer assisted learning. Time spent learning using the software was a direct substitute for traditional learning.	Education production function is concave in computer assisted learning. Estimates suggest a hybrid of computer assisted learning and traditional learning is optimal.

A.3 Related Course Design Literature

Table 13: Research Exploring Course Design and Implications of Homework Participation

Title	Authors (Date, Publisher)	Data	Research Design	Main Results
Procrastination, Deadlines, and Performance: Self-Control by Precommitment	Ariely and Wertenbroch (2002, PS).	N = 169 students in a semester (14 weeks) course at MIT.	Students randomly assigned into 1) equally spaced deadlines, 2) end deadlines, and 3) free-choice (set own deadlines).	Performance in equally spaced deadlines dominates self-imposed deadline. However, self-imposed deadlines enhanced performances more than maximally delayed deadlines.
Requiring a Math Skills Unit: Results from a Randomize Experiment	Pozo and Stull (2006, AER:P&P).	N = 273 students in principles of macroeconomics from the spring of 2004 from Western Michigan University.	Randomly assign students to one of two sections with the treatment group requiring the completion of a math unit as a part of the final grade.	Requiring a graded math unit increases participation in homework and raise overall course performance by 2 percentage points, implying an increased letter grade for 26 percent of the class.
Experimental Evidence on the Effect of Grading Incentives on Student Learning in Spain	Artes and Rahona (2013, JEE)	N = 289 students in Public Finance offered in fall of 2019 at University of Madrid.	Students enrolled in morning and afternoon version of class alphabetically by their last name. Experimenters randomized 2 out of 4 problem sets to be graded in each section. Exam questions could be linked to a problem set.	Graded problem sets increased final exam score by 8 percentage points. Students with lower baseline knowledge benefit most from graded homework.
The Role of Homework in Student Learning Outcomes: Evidence from a Field Experiment	Grodner and Rupp (2013, JEE); Citations = 81.	N = 423 students enrolled in microeconomics in the spring of 2008 in a mid-sized university in North Carolina.	Students randomized to two grading schemes 1) 10% homework and four 22.5% tests, and 2) four 25% term tests.	The homework-required group had 5-6% higher test averages; 10-14% higher for those who failed the first test. Students in the homework-required group are also 6 p.p. more likely to complete the course.
The Impact of Assignments and Quizzes on Exam Grades: A Difference-in-Difference Approach	Latif and Miles (2020, JSE)	N = 124 students enrolled in an introductory statistics course in a small business school in Canada.	Authors use data over time across 3 course sections and employ differences-in-differences approach to leverage introduction of quizzes and assignments in only some of the sections. Control section does not have either quiz nor assignment.	Introduction of homework assessment significantly improves midterm grade. Although, no significant results found with the introduction of quizzes.

A.4 Related Student Effort Literature

Table 14: Research on Exploring Student Effort

Title	Authors (Date, Publisher)	Data	Research Design	Main Results
The Effect of Time Spent Online on Student Achievement in Online Economics and Finance Courses	Calafiore and Damianov (2011, JEE).	N = 438 students enrolled in online Economics and Finance courses during the Spring and Fall 2008 in large public university in south Texas.	Multiple and logistic regression analysis using prior cGPA, age, gender, and major as a control variables. Use sessions logs from Blackboard to track time usage.	Even after conditioning on prior cGPA, time spent on course activities is a significant predictor of performance and earning a better letter grade in the course.
"Making it count": incentives, student effort and performance	Chevalier, Dolton, and Luhrmann (2018, JRS).	N = 424 introductory economic students across two cohorts enrolled at a large college of the University of London. Students are followed across 20 weeks.	Variation in incentives across weeks of either 1) additional study material conditional on quiz participation, 2) 20 GBP book voucher for best quiz performance, or 3) quiz grade counts towards course grade.	Additional study material for participation and book vouchers are ineffective in increasing quiz participation. Grade incentives significantly increases quiz participation and also results in improved exam grades.
Financial Incentives and Educational Investments: The Impact of Performance-Based Scholarships on Student Time Use	Lisa Barrow and Cecilia Elena Rouse (2018, EFP).	N = 5160 high school seniors in California.	Students randomized to performance based (obtain a C average) post-secondary scholarships of $1000 − $4000.	Financial incentives induce more time usage on educational activities and allocate less time on work and leisure.
What sets college thrivers and divers apart? A contrast in study habits, attitudes, and mental health	Beattie et al. (2019, EL)	N = 3849 students enrolled in introductory economics in 2017 at University of Toronto.	Compare student characteristics and habits across thrivers and divers.	Thrivers study around 15 hours per week, seven more hours per week than divers (8 hours per week).
When Study and Nudge Don't Go As Planned: Unsuccessful Attempts to Help College Students	Oreopoulos et al. (2018, NBER WP).	N = 9503 students from University of Toronto (N = 3438) and Western Governors University (N= 6065) in the 2017-18 academic year.	Students randomly assigned to 1) personality test (control) or 2) planning module (build weekly calendar + assigned coach).	Despite marginal increase in study time for those in treatment group, null effects on course grades and retention.
Using Goals to Motivate College Students: Theory and Evidence from Field Experiments	Clark et al. (2020, ReStat)	N = 2004 students for task-based experiment, and N = 1967 for performance based experiment. First year introductory course.	Students randomly assigned to control or goals treatment. Fall 2013 cohort for performance-based goals, and Fall 2014 for task-based goals.	Task-based goals increased task completion and resulted in significant performance gains. Although, performance-based goals are not as effective.

B Appendix: Institutional Details

B.1 Course Outline

The course is taught over 12 weeks. Learning the principles of programming can be broken down into the following three stages: 1) basic concepts (e.g., variables and loops), 2) intermediate concepts (e.g., nested loops and parallel lists), and 3) advanced higher order concepts (e.g., algorithms and object oriented programming). That is, the course have a cumulative structure where topics build on each other. The following table includes the syllabus for the foundation programming course.

Week	Topics Coverage
1	Numerical operations, variable assignment, and common coding errors
2	Defining functions and string variables
3	Conditional statements (if, elif, and else) and boolean variables
4	Loops (for and while)
5	Properties of lists (e.g., aliasing and mutability)
6	Nested lists and nested loops
7	Tuples, dictionary, and parallel lists
8	Palindromes classification algorithm and more about lists, tuples, and dictionaries
9	Good programming practices for testing and debugging code (e.g., unit tests)
10	Search and sorting algorithms (e.g., binary search and bubble sort)
11	Writing classes and methods
12	More object oriented programming (classes and methods)

The course employs two online learning platforms: an online homework environment and an online peer discussion board.

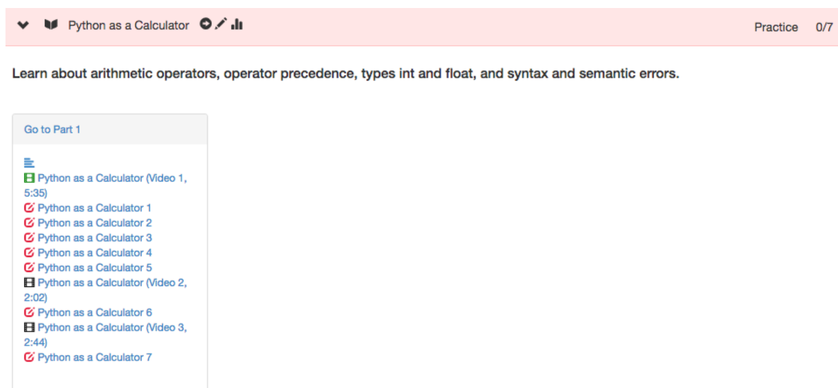
B.2 Online Homework Environment

Each week students receive an online homework module where students watch videos and then subsequently solve homework problems. Students login to the platform, and are given an outline for the videos they should watch and are presented with the follow-up coding problems. The online learning platform hosts a total of 133 videos (7.1 hours) and 401 follow-up homework problems. All homework problems are graded through an automatic artificial intelligent system. The following table presents summary statistics for the weekly content available on the platform.

Variable	Mean	SD
No. of videos assigned per week	11.1	4.4
Minutes of video lectures assigned per week	35.4	14.402
No. of questions assigned per week	33.3	13.614
Proportion of coding questions per week	0.22	0.121
No. of weeks	12	

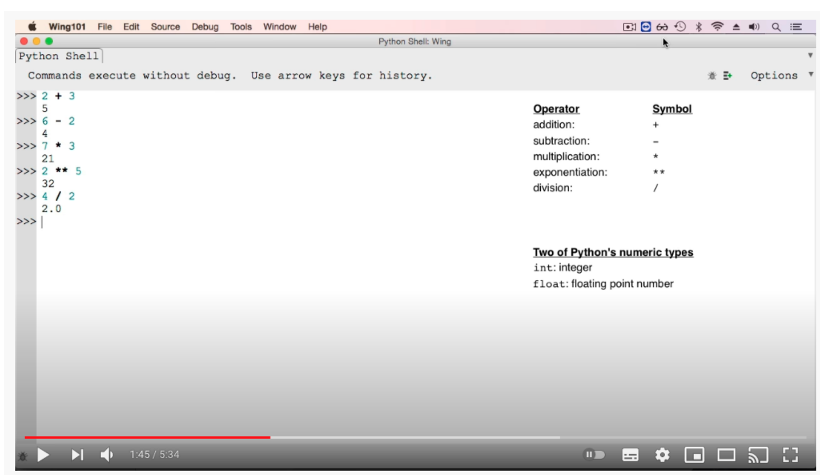
Students provided an outline for how to learn a topic:

Figure 21: Outline for Learning Numerical Operations



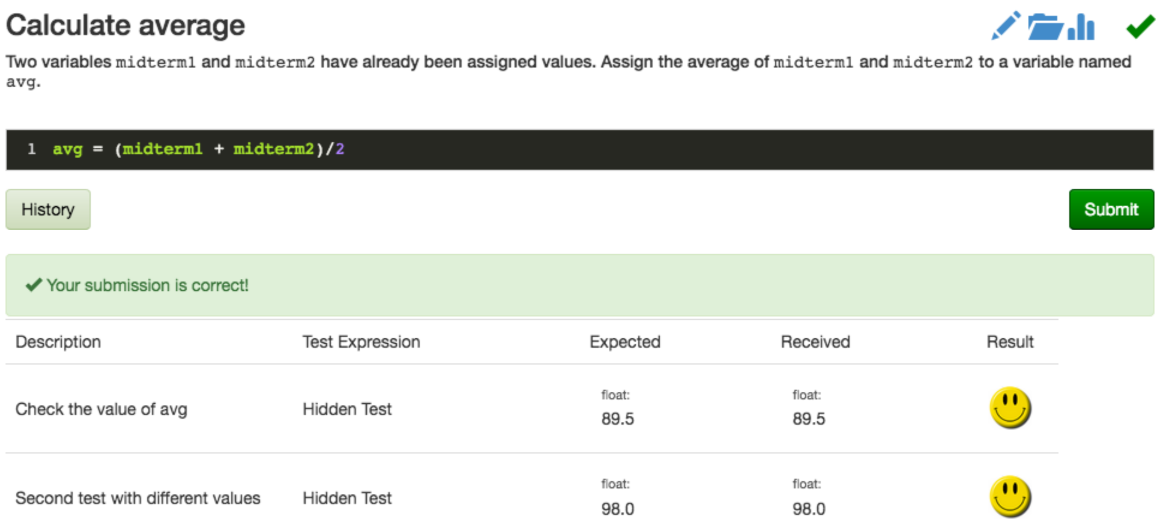
Students begin the course by watching a video about numerical operations in Python:

Figure 22: Video on Numerical Operations



The following figure shows an example of a follow-up coding problem:

Figure 23: Video on Numerical Operations



B.3 Online Peer Discussion Board

Students can use the online peer discussion board to get help with course material through asking questions. The questions are answered by peers, and answers can be validated by TAs or instructors. Students can also comment on either questions and answers. Comments can be used to further clarify the question, or give ideas on how to start solving the problems. The following table shows an example of student interactions on the discussion board.

Table 15: Example of Student Peer Interaction on Discussion Board

Interaction Type	Response
Question	How do we write a new line in a file using python?
Answer	Similar to how you would create a new line in a print function: file = open("somefile.txt", "w") file.write('\n') file.close() I hope that helps.
Comment	The code in the answer works, but note that opening a file in write mode will delete the contents of the file. Use append mode if you want to add to the file.

C Appendix: Data

C.1 Measuring Student Demographics and Characteristics

Student demographics and other characteristics shown in Panels A and B of Table 1 are used as pre-treatment controls for most regression specifications. Several of the student controls are constructed using the following questions on the baseline survey. Aside from the options listed below, students could also opt-out from answering the question by selection "Prefer not to answer".

- What is your gender identification?
  - Male; Female; Other

- Are you the first one in your immediate family to attend university?
  - Yes; No

- What is your mother's highest level of education?
  - Did not finish high school; high school graduate; some college; college graduate; graduate degree (e.g., masters or doctorate)

- What is your father's highest level of education?
  - Did not finish high school; high school graduate; some college; college graduate; graduate degree (e.g., masters or doctorate)

- Is English your native/first language?

  - Yes; No

- What language do you speak at home? *(open response)*

- How would you describe your prior experience with programming?
  - I have never programmed before; I have written a few lines of code; I have written basic programs before; I have extensive experience programming

- Which of these is closest to your (intended) program of study?
  - Computer Science; Commerce; Humanities; Life Sciences; Physical and Mathematical Sciences; Social Sciences; Other

C.2 Measuring attentiveness and forward-looking Perspective

The baseline survey elicits a student's attentiveness and forward-looking perspective through a series of questions. Responses to the relevant questions are aggregated so that they are increasing in the attribute of interest. The tables below include the 7-point Likert scale survey questions used to measure forward-looking perspective and attentiveness.

Table 16: Survey Components for Forward-looking Perspective

Survey question (7-point scale)	Relationship with Forward-looking
I consider myself to be a forward-looking person who has clear plans about the future	Increasing
I tend to think about how working hard on the present homework will make doing future homework easier	Increasing
I tend to think about how working hard on homework each week will help me do better on the exam	Increasing
When I have multiple deadlines, I think ahead and plan how to split my time before I begin working	Increasing
I have a good sense of my expected career trajectory after completing my current degree	Increasing

Table 17: Survey Components for Attentiveness

Survey question (7-point scale)	Relationship with attentiveness
I tend to read all the instructor announcements for this course each week	Increasing
I have read the course syllabus in detail	Increasing
I know how to access office hours	Increasing
I know when office hours are held	Increasing
I tend to forget about my assignment deadlines	Decreasing

I characterize students as being forward-looking if they responded with at least 5 to each question in Table 16. Similarly, students are classified as attentive if they responded with at least 5 to the first three questions in Table 16, and at most 3 to the last question.

C.3 Measuring English proficiency

The tables below include the 7-point Likert scale survey questions used to measure a student's English proficiency. Responses are aggregated so the resulting index variable is increasing in English proficiency.

Table 18: Survey Components for English Proficiency

Survey question	Relationship with English proficiency
I am comfortable writing in English	Increasing
I am comfortable speaking in English	Increasing
I have difficulty reading and listening in English	Decreasing
I have difficulty learning by watching instructional videos in English	Decreasing

C.4 Measuring Study Time

I use the time-stamped online interactions to construct a measure of total study time for each learning stage. Students primarily spend their time on the online homework platform. Additionally, students participate in the online peer discussion board by writing and reading posts.

C.4.1 Study time on online homework

The administrative data includes time-stamps for when students log-in, log-out, click to play/pause videos, submit a solution to a problem, and various other interactions with the platform. I develop a simple algorithm that uses the time-stamped data to measure the number of minutes of videos watched (v) and minutes spent doing homework problems (h). The algorithm leverages the fact that students tend to study in around 30-minute blocks throughout the week. The blocks of study time are identified to the nearest 5-minute of inactivity and aggregated together. Then, for each learning stage t, the time spent on the online homework is:

$$e_{i,t}^{H} = v_{i,t} + h_{i,t}.$$

C.4.2 Study time on peer discussion board

Although the administrative data includes the number of posts written (w) and unique posts read (r), the time spent on these activities is not included. To fill this gap, the final survey asks students how many minutes on average they spend writing (m^w) and reading a post (m^r). The survey questions eliciting time costs to discussion board participation are as follows:

- Roughly how much time (in minutes) do you believe it takes you to write an average quality discussion board post (i.e. make new question or answer peer question)?
  - Minutes it takes to write a question [numeric response]
  - Minutes it takes to write a answer [numeric response]
- In a hypothetical scenario, suppose you were given 10 minutes to browse the discussion board and read posts (question or answers). How many posts do you think you could read in detail in that time period?
  - Number of questions carefully read in 10 minutes [numeric response]
  - Number of answers carefully read in 10 minutes [numeric response]

Then, for each learning stage t, the time spent on the discussion board is:

$$e_{i,t}^{D} = m_i^w w_{i,t} + m_i^w r_{i,t}.$$

C.4.3 Total study time

Time spent across the online homework and discussion boards aggregated at each learning stage to construct study time:

$$e_{i,t} = e_{i,t}^{H} + e_{i,t}^{D}$$

C.5 Survey questions eliciting student peer interactions

The surveys include the following questions to measure the extent to which students interact with other peers in the course.

- Are you in a study group for [CourseCode]?
  - I am in a study group officially recognized by [institution name]
  - I am in another study group with students from this course
  - No
- Around how many students in the course do you study with per week? [Numerical Entry]
- Around how many hours per week do you study with other students in this course? [Numerical Entry]
- I discussed the discussion board information received during the baseline survey with other students in the course [Likert Scale]
- I discussed the contents of the homework reminder messages with other students in the course [Likert Scale]

D Appendix: Nudges

The nudges are designed using various behavioural insights such as implementation intentions, utility value, and self-reflection. Kizilcec et al. (2020), Harackiewicz and Priniski (2018) and Damgaard and Nielsen (2018) provides excellent reviews on the behavioural nudging literature in education.

D.1 Sign-up activity

The sign-up activity is designed to help students internalize the information provided through self-reflection. Students were given the following questions as a part of the activity.

- (Open Response) Based on the instructions presented, approximately how long does it take to sign up for the discussion board?

- (Disagree-Agree) Are you aware that $X\%$ of questions asked have received a response? *[X varies by cohort and is always above* 85%*]*

- (Disagree-Agree) I am aware students can use the discussion board to learn through ...
  - asking their own questions
  - answering questions of other students
  - engaging in discussion with peers by commenting on posts

- (Open Response) How could you potentially benefit from the discussion board?

- (Multiple Choice) Would it be worthwhile for you to sign-up to the discussion board? Here is the link to sign-up page: [Link]
  - Yes, I have just registered for the discussion board
  - Yes, I will sign-up to the discussion board this week *[Follow-up prompt to schedule day and time]*
  - No, I will not sign-up to the discussion board *[Follow-up prompt to request reasoning for not using discussion board]*

D.2 Homework reminder messages

The homework reminders are sent through the learning management system. Students receive the reminder in their personal university email inbox and a notification of the message on the learning management system. The template for the homework reminder is as follows.

Hi [Student Name],
The homework is due by [Deadline]. Please take a moment to think about the following prompts:

When will you next work on this week's homework? Can you set aside time on your schedule to progress on the homework?

Some students find it valuable to just open up the online homework system and spend a minute on a problem. Here is the link to the homework: [Link to Homework]

[Course Code] Learning Support Team