Identifying the STEM Learning Technology: Evidence from Online Learning

Hammad Shaikh*
University of Stavanger

March 28, 2023

Abstract

The number of annual graduates with STEM degrees has grown dramatically over the past decade. Despite this growth, many students enrolled in STEM struggle to complete their program. An inherent feature of STEM education is that it has a natural cumulative learning structure which makes learning advanced skills in STEM quite challenging. This paper is the first to credibly estimate the cumulative learning technology in a foundational STEM course. Doing so is incredibly challenging as precise measures of effort inputs are typically unavailable and are also dynamic endogenous choice variables. To overcome such challenges, I first gather rich panel data covering nearly 3,700 undergraduates at a large public university taking an online introductory programming course that has a cumulative structure. The online learning environment serves to monitor students' effort allocation and knowledge accumulation at each stage of the learning process. Then I carry out a field experiment which generates period-by-period exogenous variation in effort allocation, enabling me to identify dynamic interactions across effort inputs in the learning technology. I find evidence of dynamic learning complementarities as the marginal benefit to studying in each learning period is increasing in prior knowledge accumulated. This result suggests that students studying effectively should front-load their effort allocation, but the opposite behaviour is documented in the data. The findings in this study have implications for effective learning strategies and approaches to course design.

*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, Fernando Saltiel, 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, Empirical Microeconomics Seminar, Canadian Economic Association Conference, Norwegian Economic Association Conference, and the University of Stavanger's 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. The study is registered at the AEA RCT Registry (ID 0008787). A previous version of this paper was circulated under the title "Improving Online Learning Through Course Design: A Microeconomic Approach". Financial support from an Ontario Graduate Scholarship is greatly appreciated. All remaining errors are my own. Contact: Department of Economics, University of Stavanger Business School, Kjell Arholms gate 35, 4021 Stavanger. Please send comments to hammad.shaikh@uis.no.

1 Introduction

STEM education has expanded rapidly over the past decade as technical skills are fundamental to the development of any modern technological society (Hanushek and Woessmann, 2015). The number of students conferring STEM degrees in the US, for example, has increased from around 472,000 in 2009 to around 733,000 in 2018 (US National Centre for Education Statistics). Policy makers continue to call for a dramatic increase in the supply of STEM majors, and substantial government funding has been allocated towards STEM education with a focus on promoting computer science.[1]

An inherent feature of STEM is that the learning process is typically cumulative. Consequently, acquiring advanced skills in STEM is challenging as students must first obtain mastery in preceding foundational concepts.[2] The cumulative learning structure of STEM can cause students to permanently fall behind if they do not exert sufficient effort learning fundamental skills. Additionally, it is challenging to design STEM courses as their large class sizes give rise to heterogeneity in students prior preparation and organizational skills to learn effectively. Given these challenges, many students struggle to complete their STEM majors and dropout during their first two years of university.[3] Then understanding the effective learning process is especially important in foundational STEM courses as they are mandatory prerequisites for all other advanced STEM coursework.

This paper is the first to credibly estimate the cumulative learning technology which maps dynamic effort inputs into learning outcomes in a foundational STEM course. The specified technology is cumulative with multiple learning periods and allows for past knowledge to persist into the future. The learning technology also incorporates dynamic learning complementarities as the productivity of effort in a learning period can depend on previous effort exertion. Estimating the learning production function helps inform effective learning strategies and approaches to course design.

Identifying such a cumulative learning process is incredibly challenging for several reasons. First, repeated learning measures to observe the knowledge accumulation process are not typically available. Second, precise students' effort inputs such as study time allocation are typically unobserved. Although data on student performance on various cognitive assessments are read-

[1] The US Department of Education spends around \$400 − \$500 million on promoting STEM education annually. Around 25% of this funding is typically allocated towards computer science education.

[2] Introduction to programming, for example, has a standard cumulative structure where students learn numerical operations, functions, and then algorithms.

[3] Stinebrickner and Stinebrickner (2012) and Stinebrickner and Stinebrickner (2014) find that drop-out decisions can be largely explained by students learning about their academic abilities and being overoptimistic about succeeding in a science program at college entry.

ily available, administrative data on students' precise study time allocation is nearly nonexistent. Third, the effort inputs across the cumulative learning process are dynamic endogenous choice variables. Then period-by-period exogenous variation in effort is required to identify the dynamic learning complementarities, even though just the availability of a single exogenous shock in effort is a rare occurrence.

I develop a unique approach to address all theses challenges for the first time. I first gather unique administrative data from a prominent online STEM course which enables me to precisely measures effort allocation and corresponding knowledge accumulation throughout the entire course. I then carry out a field experiment which generates exogenous variation in effort allocation, credibly identifying each element of the cumulative learning process. The estimates of the learning technology help inform effective learning strategies for students in STEM. Lastly, I use rich survey data to study heterogeneity in effort allocation across different types of students.

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 a voluntary 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.

I collect data on nearly 3,700 students who consented to participate in the research.[4] The rich administrative data include time-stamped student interactions with the online learning environments throughout the entire semester. 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. For example, each week around $15-20\%$ of students spend no time whatsoever doing the homework. 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. I find that more attentive students have a higher propensity to start the low-stakes homework assignments and have higher awareness of course resources.

To identify the cumulative learning STEM process, I conduct a field experiment to generate

[4]The student consent rate is around 87% and over 40,000 student-week observations are available.

exogenous variation in students' effort allocation. The interventions considered in this study can 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 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 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.

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. Reflective of the actual course structure, the learning process is estimated separately across three learning stages: basic, intermediate, and advanced. 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.

The estimated model has implications for effective learning strategies for students and approaches to course design for instructors. Given the cumulative nature of the STEM learning process, students learning effectively should sufficiently front-load their effort allocation to achieve proficiency in the basic concepts before moving onto learning more advanced topics. Contrary to this learning strategy, I find that students on average, however, increase their effort supply as the course progresses. The evidence suggests that many students do not internalize the cumulative learning process when allocating effort. In a follow-up paper, I study the extent to which front-loading the grading incentives can help improve learning outcomes for students who learn ineffectively.

Although the analysis in this paper focuses on a single online STEM course, there are good reasons to think the results can inform the learning process in 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 guidance on how to learn effectively. First-year university courses also typically have a cumulative structure, especially in STEM, with considerable heterogeneity in the student body. Then students who do not internalize the cumulative learning process and are at risk of falling behind early on will be present in any large foundational university course. Most popular learning management systems are equipped with online homework hosting capabilities and student activity analytics. Then these online facilities can be used monitor student effort allocation and knowledge accumulation to track whether they are learning effectively. In these ways, understanding the learning process using the approach considered in this paper can be applied to 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 dynamic effort supply is discussed in Section 7, and the cumulative learning process is estimated in Section 8. Finally, Section 9 concludes and discusses avenues of future research.

2 Literature Review

This paper builds on several prior literatures. These include research that estimates cumulative education production functions and those that 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 paper contributes to each of them.

First, my paper contributes to a body of research estimating the marginal learning returns to student effort, an essential input to the education production function. Stinebrickner and Stinebrickner (2008) estimates the returns to effort using self-reported diary data on time use from college students, and also collect information on whether their roommate owns video games. The authors use roommate ownership of videos games as an instrument for study time and find that increasing study time by 1 hour a day increases the first semester GPA by around 0.3 - 0.4. More recently, Ersoy (2021) uses administrative data from a popular language learning platform, Duolingo, to measure causal learning returns to effort. In the study, students learning Spanish are randomly assigned to complete a different number of lessons. The author finds that spending around a hour completing 9 lessons results in increasing achievement by 0.057 SD on tests that are external to the online learning environment. My paper contributes to this literature by using administrative data from a large prominent STEM course together with a field experiment to measure the causal learning returns to study time and assignment completion.

Second, my paper also 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.1 describes the education production function literature in more depth. Although the education production function literature includes parental and school inputs, time varying students inputs are not incorporated into the learning technology. Consequently, policies designed to increase educational attainment in these papers target either parental or school inputs, and implicitly assume that students' effort is held constant. The prior literature has yet to identify the cumulative education production as a function of student effort, important for designing dynamic policies to encourage students to learn effectively. 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.

Finally, the 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 build-

[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.

ing 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 did not result in a significant increase in learning. Appendix A.2 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 programming fundamentals using Python (see Appendix B.1 for the course outline). It is offered every semester and typically enrols around 1,000-1,500 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 stages: 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. 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

[6] For example, learning nested lists and nested loops in week 6 requires an understanding of basics of loops and lists, covered in weeks 4 and 5, respectively.

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 and final exams are typically written in weeks 5 and 13, respectively. Additionally, 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 the 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 of 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. Students 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 awarded any additional course credit for participation.

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 that is separate from their institutional email. Instructors can send student messages through Canvas, and such messages

[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.

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 at 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 6.

Student Surveys. The baseline survey collects information about students' demographics and elicits information about their behavioral 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 environment. Overall the consent rate is around 87%. The sample of total consenting students who completed the course consists of 3,686 students.

Student Activity on the Learning Management System. I collect student activity data from student interaction reports captured on 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

[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 two days before the deadline received a reminder to do so.

over the study period.[11]

Student Achievement. Student achievement data are collected from the weekly online homework, the midterm, and the final exam. This high-frequency achievement data allows me to assess student learning throughout the course. The primary measure of student learning is their cumulative final exam grade, which I standardized to have a mean of 0 and standard deviation of 1.[12]

Student Discussion Board Activity Data. Students' 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 types of student interactions with the online platforms are observed in the administrative data, recorded 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 log in or out, 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.

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

[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.

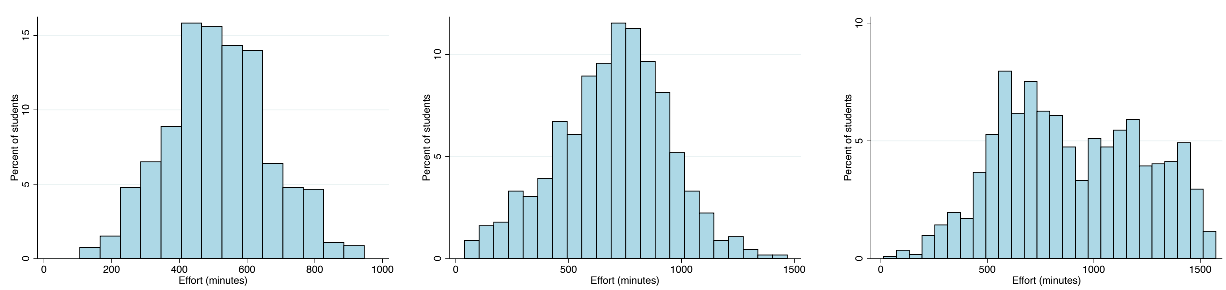
with the online learning platform for at least 5 minutes, and concludes after 5 minutes of inactivity. Video watching time is computed based on when students play or pause the instructional videos. Students' time spent attempting homework problems is measured using information when students submit problems 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 data set 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 average time they spend on average writing and reading a post in minutes (see Appendix C.3 for survey questions). The administrative data on student engagement, and corresponding student-level survey data on average time use are 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 1 below shows the distribution of online study time at each learning stage of the course. This figure shows that, on average, student 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.

Figure 1: Distribution of Study Time by Learning Stage

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



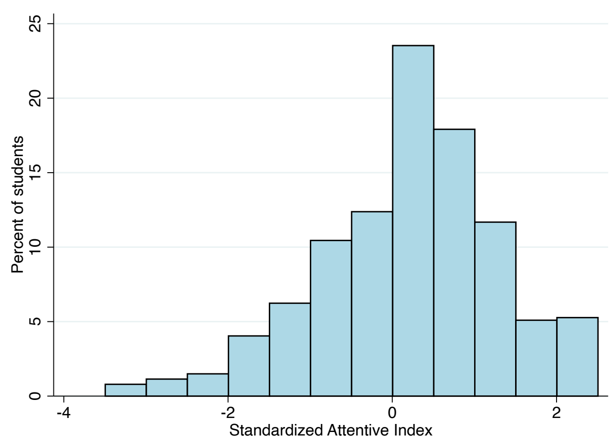
Notes: The figure presents the distribution of total online study time for each stage of the course: basic, intermediate, and advanced. All histograms use a bin width of 60-minutes.

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

4.2 Student Attentiveness

The baseline survey elicits student attentiveness 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). To measure attentiveness, one question students are asked whether "I tend to read all the instructor announcements for this course." All questions relating to students' attentiveness are included in Appendix C.2. To construct a continuous index of the behavioral responses, replies to all relevant questions are aggregated together so that they are increasing in attentiveness. The distribution of student attentiveness is shown in the following figure:

Figure 2: Distribution of Attentiveness



Notes: The figure presents the distribution of the standardized attentiveness index. All histograms use a bin width of 0.5 SD.

The apparent left skew of the attentiveness distribution suggests that most students self-report themselves to be attentive.

4.3 Summary Statistics

Table 1 presents a rich set of summary statistics related to student demographics, characteristics, behavioral information, homework participation activity, and overall study time. 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. Panel A shows 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. Panel C shows that around 76% of students are attentive. Panel D indicates that 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 each homework assignment. Students who registered for the discussion board spend around 28 minutes each week on making posts or viewing content.

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 interventions presented in this section.

5.2 Description of the Intervention

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, 2) they provide information on how to clearly execute the action, and 3) they serve as a reminder for the specified task. The design of the nudge is inspired by insights from psychology and behavioural economics research (Damgaard and Nielsen, 2018). In particular, the intervention is designed to nudge inattentive students who may have a tendency to forget homework deadlines.

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.1 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 7 illustrates the assignment of students to the number of homework reminders.[16]

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.

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 a control of treatment group each week. I investigate the validity of the random assignment by testing whether the pre-treatment student demographics and characteristics are balanced across the experimental conditions. I do so by standardizing each pre-treatment control and regressing these on the number of reminders received. Figure 8 shows that students who are assigned to receive

[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.

[16] Alternatively, students could have been uniformly randomized to receive between 0 to 10 reminders at the start of the course. This design was not implemented due to caveats that were discovered while piloting the reminder messages in the Summer 2019 and Fall 2019 cohorts. Since a small portion of students complete the homework soon after its release, these students should not receive a reminder. Additionally, it was important to check with the instructor each week that the reminder message would not crowd-out any important announcement that they may of wanted to make.

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 2 examines whether the number of reminders received impacts the propensity to drop out. The analysis suggests that the reminder messages did not cause 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 homework reminders is homogenous across the treatment groups as all students in the treatment condition receive the same reminder. 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 group who received the reminder will interact with the control group who did not receive any reminder. Assuming the reminder 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.[17]

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 reminder messages

[17] Upwards biased estimates due to information spillover are possible, but unlikely. For example, suppose students in the control group discover that their peer in the treatment group received a reminder. Then the control group student may feel discouraged and exert less effort because they are not being supported by the instructional team.

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.4 includes the survey questions about student peer interactions.

6 Experimental Results

This section discusses the results from the field experiment described in Section 5, and outlines the corresponding empirical methodology. For simplicity of exposition and interpretation, the analysis is carried out by aggregating the data so that the parameters of interest are estimated by a cross-sectional regression with cohort fixed effects.

6.1 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:

$$D_{ic} = \delta_0 + \delta_1 RemindersFreq_{ic} + \pi_c + X_{ic}'\Delta + \epsilon_{ic},$$

where D_{ic} is either the number of homework assessments completed or the total hours a student spends studying; π_c is cohort fixed effects; $RemindersFreq_{ic}$ is total number of homework reminders a student receives. Control variables X_{ic} include student demographics and other pre-treatment characteristics listed in Panels A and B of Table 1.

Table 3 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 the online homework platform, 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. Consequently, the frequency of reminders received will provide a strong first stage for inducing exogenous variation in study time.[18]

Figure 9 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

[18] The F-statistic exceeds the threshold of 104.6 stated in Lee et al. (2022) for assessing a strong first stage.

statistically significant.[19]

Mechanisms Underlying the Homework Reminders. Next, 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 10 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. The evidence suggests that the homework reminders are most effective at encouraging inattentive students to exert more effort.

6.2 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 innate 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.

I argue that 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).[20]

Relevance. A theoretical model presented in Ericson (2017) shows that reminders can be helpful in task completion when individuals have a limited memory. I find empirical evidence consistent with the models implication as Figure 10 shows that the reminder messages are perceived to be most beneficial by more inattentive students. Additionally, Table 3 shows that on average the reminder messages are successful in significantly increase effect exertion (as previous discussed).

Exclusion. Exclusion is violated if receiving homework reminders affects learning through channels aside from homework participation. This is plausible if the reminder message induces students to also participate further, for example, on the online discussion board. Receiving reminders over

[19] 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.

[20] I assume a homogeneous treatment effects framework for simplicity. Under a heterogeneous effects framework, monotonicity of the instrument and the absence of defiers can also be reasonably argued.

time may also help students build better time management and organization skills, enabling them to learn more by more efficiently using a given amount of study time. As the reminder message directly targets the homework assessment, it is unlikely to affect participation on the discussion board. Study habits and organizations ability will be well established prior to enrolling in the course. It is unlikely a low intensity nudge such as receiving a few reminder messages will have persistent effects on long term study habits.

Independence. Since students are randomly assigned to whether they receive a homework reminder for each homework assessment, then the number of reminders they receive is independent of any observed or unobserved determinants of learning. Concerns around finite sample imbalances across the type students who receive many reminders versus few reminders are alleviated by the large sample of students involved in this study. Evidence to support the independence assumption is presented in 8 (as previously discussed).

I now employ the frequency of reminders received as an instrument to estimate the causal effects of homework participation on homework performance using the following 2SLS model:

$$\begin{cases} ExamGrade_{ic} = \lambda_0 + \lambda_1 D_{ic} + \pi_c + X'_{ic}\Pi + \epsilon_{ic}, \\ D_{ic} = \phi_0 + \phi_1 RemindersFreq_{ic} + \pi_c + X'_{ic}\Gamma + \epsilon_{ic} \end{cases}$$

where $ExamGrade_i$ denotes the final exam grade. Table 4 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 the fact that the homework is the primary source of learning the course material in this online STEM course.

6.3 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 9% of students in the final survey attested to discussing contents of the reminder messages with their peers at least once. 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 the reminder messages vary according to whether students are in a study group at baseline. The analysis is presented in Table 5. The results suggest that the efficacy of the reminder messages does not vary according to whether a student is in a study group.

7 Theoretical Framework

In this section, a learning environment with a cumulative structure is conceptualized, where students acquire knowledge by exerting effort over multiple learning stages. The framework formalizes an effective learning strategy in courses with a cumulative structure.

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). All students are assumed to be forward-looking who internalize the cumulative learning process when allocating effort.

7.2 The Student Effort Choice Problem

Students allocate their effort to maximize their total knowledge net of effort costs as follows:

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

where $C(\cdot)$ is an convex function that is increasing in effort, representing the cost of effort exertion. The learning technology is concave in each effort input, increasing in effort and baseline human capital.

For simplicity of the exposition, we assume total learning can be separated as follows:

$$L_i(e_i^{basic}, e_i^{int}, e_i^{adv}; h_i) = L_i^{basic}(e_i^{basic}, h_i) + L_i^{int}(e_i^{int}, L_i^{basic}) + L_i^{adv}(e_i^{adv}, L_i^{int}).$$

That is, 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}.[21] Similarly, I also assume the cost of effort is separable across learning stages:

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

imposing that student fatigue does not spillover across learning periods, a reasonable assumptions if student 'burn out' from exerting too much effort is not a major concern in the learning setting.

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

Students exert effort until the marginal benefit of effort exceeds the marginal cost of effort. The first order condition that characterizes effort exerted at the basic stage is:

$$\frac{dL_i}{de_i^{basic}} = \frac{dL_i^{basic}}{de_i^{basic}}\left[1 + \frac{dL_i^{int}}{dL_i^{basic}}\left(1 + \frac{dL_i^{adv}}{dL_i^{int}}\right)\right] = \frac{dC}{de_i^{basic}}.$$

Assuming a cumulative learning structure with previous knowledge persisting into future learning periods, then $\frac{dL_i^{int}}{dL_i^{basic}} > 0$ and $\frac{dL_i^{adv}}{dL_i^{int}} > 0$. That is, exerting effort in the basic stage has spillover learning benefits in the intermediate and advanced learning stages. Intermediate learning stage effort is determined by the following equation:

$$\frac{dL_i}{de_i^{int}} = \frac{dL_i^{int}}{de_i^{int}} \times \left[1 + \frac{dL_i^{adv}}{dL_i^{int}}\right] = \frac{dC}{de_i^{int}}.$$

Finally, effort in the advance learning stage is characterized by:

$$\frac{dL_i}{de_i^{adv}} = \frac{dL_i^{adv}}{de_i^{adv}} = \frac{dC}{de_i^{adv}}.$$

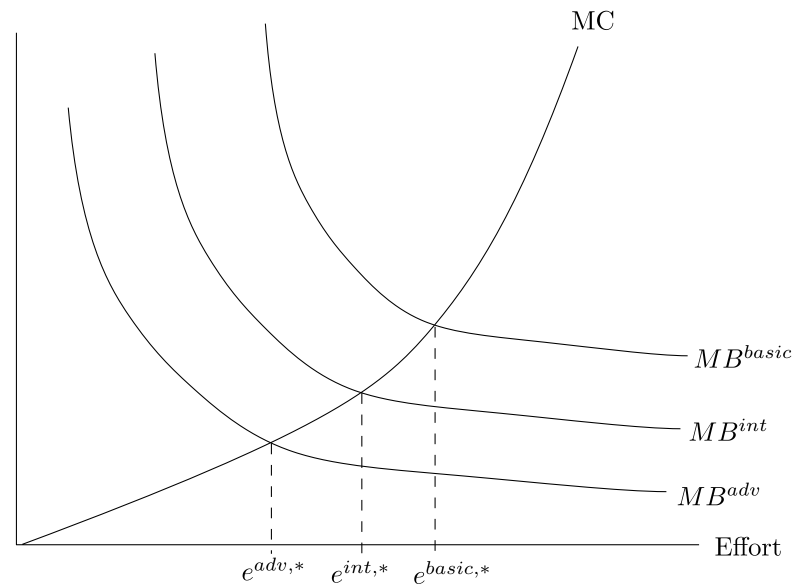
As the advanced stage is the last learning period, there is no future spillover learning benefit from exerting effort in the advanced stage. I will show that under reasonable assumptions, the effective learning strategy in a course with a cumulative structure is to front-load effort allocation so that $e_i^{basic,*} > e_i^{int,*} > e_i^{adv,*}$.

As course material typically becomes more challenging as the course progresses, the period specific marginal benefit of effort decreases across learning stages: $\frac{dL_i^{basic}}{de_i^{basic}} > \frac{dL_i^{int}}{de_i^{int}} > \frac{dL_i^{adv}}{de_i^{adv}}$. The persistence of learning from one period to the next will likely be less than 100%: $\frac{dL_i^{int}}{dL_i^{basic}} < 1$ and $\frac{dL_i^{adv}}{dL_i^{int}} < 1$. Suppose that the persistence of knowledge from basic to intermediate exceeds the persistence from intermediate to advanced (i.e., $\frac{dL_i^{int}}{dL_i^{basic}} > \frac{dL_i^{adv}}{dL_i^{int}}$) then marginal total learning benefit of effort in the basic stage exceeds the intermediate stage (i.e., $\frac{dL_i}{de_i^{basic}} > \frac{dL_i}{de_i^{int}}$). Although such an assumption may be strong, $\frac{dL_i}{de_i^{basic}} > \frac{dL_i}{de_i^{int}}$ holds as long as $\frac{dL_i^{int}}{dL_i^{basic}}$ is sufficiently large. Under these assumptions, the marginal total learning benefit of effort is decreasing as the course progresses

$$\frac{dL_i}{de_i^{basic}} > \frac{dL_i}{de_i^{int}} > \frac{dL_i}{de_i^{adv}}.$$

Then clearly the optimal effort exerted decreases across the learning stages as illustrated by the following figure:

Figure 3: Effective Effort Allocation When Learning is Cumulative



Notes: The figure illustrates the optimal effort allocation across the basic, intermediate, and advanced learning stage as the intersection of the respective marginal benefit and marginal cost of effort curves. The marginal benefit of effort shifts downwards across learning stages under a reasnoable cumulative learning structure.

7.3 Stylized Example

To intuitively illustrate the implications of the model, consider a course with a cumulative structure and two learning stages. Suppose that students learn basic concepts in the first half of the course and advanced concepts in the remaining half. That is, $t \in \{basic, adv\}$.

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

$$L_i^{basic} = \alpha_0 + \alpha_1 e_i^{basic} + \alpha_2 h_i,$$

$$L_i^{adv} = \beta_0 + \beta_1 e_i^{adv} + \beta_2 L_i^{basic} + \beta_3 e_i^{adv} \times L_i^{basic}.$$

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:

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

The Students' Optimal Effort Choice. The effective learning allocation acrosss the basic and advanced stage to maximize learning net of effort costs is given by:

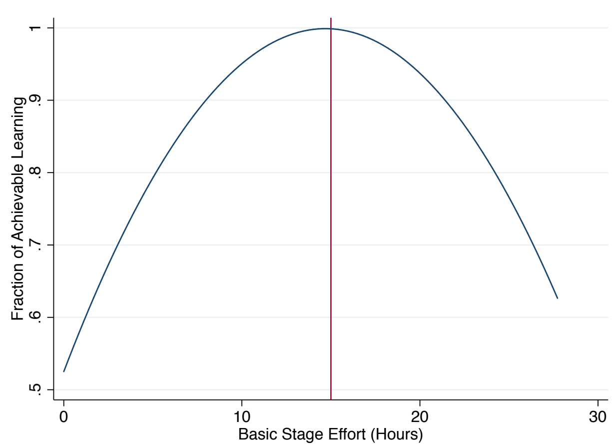
$$e_i^{basic,*} = \frac{\alpha_1\left[(1+\beta_2)(1-\beta_3\alpha_1)(1+\beta_3\alpha_1)+\beta_3(\beta_1+\alpha_0\beta_3+\alpha_1\beta_3(1+\beta_2)+\alpha_2 h_i)\right]}{(1-\beta_3\alpha_1)(1+\beta_3\alpha_1)},$$

$$e_i^{adv,*} = \frac{\beta_1+\beta_3\left[\alpha_0+\alpha_2 h_i+\alpha_1^2(1+\beta_2)\right]}{(1-\beta_3\alpha_1)(1+\beta_3\alpha_1)}.$$

Note that if there are no dynamic learning complimentarities (i.e., $\beta_3 = 0$), then the optimal effort allocation simplifies to $e_i^{basic,*} = \alpha_1(1+\beta_2)$, and $e_i^{adv,*} = \beta_1$. Additionally, suppose the course structure was not cumulative, and completely distinct content was covered in both learning stages (i.e., $\beta_2 = 0$). Then effort allocation at each stage would be the corresponding marginal effort benefit: $e_i^{basic,*} = \alpha_1$ and $e_i^{adv,*} = \beta_1$.

Simulating Optimal Effort Choice. Next, I will simulate an example effort allocation using the stylized model. Assuming learning performance is measured on a scale from 0 to 100 and effort is measured is hours of study time, I will consider the following reasonable parameter values: $\alpha_0 = 5, \alpha_1 = 3, \alpha_2 = 0.1, \beta_0 = 0, \beta_1 = 2, \beta_2 = 0.8, \beta_3 = 0.2$, and $h_i = 70$. Then the optimal effort allocation is 15.7 hours in the basic stage, and 11.9 hours in the advanced stage. That is, the student studies around 28 hours in this course, and slightly front-loads there effort allocation to the basic stage.

Figure 4: Learning and Effort Allocaton



Notes: The figure shows learning as a function of basic stage effort with total study time fixed. The vertical line denotes optimal effort allocation to maximize learning.

Figure 4 illustrates the fraction of achievable learning achieved as a function of basic stage effort when 28 hours in total area available for studying. As learning is cumulative in this setup, it is important to sufficiently exert effort in both learning periods, rather than focusing entirely on a single learning period.

8 Estimating a Cumulative Education Production Function

In this section, I describe the estimation of a multi-stage education production function. The parameterization of the production function 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 estimation 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.

8.1 Specifying 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 build upon previously attained knowledge. For example, I allow learning in the intermediate stage to be increasing in the knowledge accumulated in the basic stage. The cumulative technology reflects the cumulative course structure as programming topics build on each other.

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. That is, dynamic interactions across effort inputs may be present in the production function. Putting all three assumptions together, the learning technology in the basic stage is as follows:

$$L_{ic}^{basic} = \alpha_0 + \alpha_1 e_{ic}^{basic} + \alpha_2 h_{ic} + \delta_c + \alpha_3 e_{ic}^{basic} \times h_{ic} + \epsilon_{ic}^{basic}, \tag{2}$$

where L_{ic}^{basic} is the basic stage homework performance; e_{ic}^{basic} is the total online study time at the basic learning stage; h_{ic} denotes baseline programming experience; δ_c is cohort fixed effects, and ϵ_{ic}^{basic} is a mean 0 stochastic error term. In equation 2, $\alpha_1 > 0$ implies a positive marginal benefit of effort. The extent to which prior programming knowledge persists to the basic stage is captured by $\alpha_2 > 0$ (Assumption 2). For $\alpha_3 > 0$, the marginal learning gains from basic effort exertion are increasing in baseline knowledge (Assumption 3). The learning technology at the intermediate and advanced learning stages are analogously defined as:

$$L_{ic}^{int} = \beta_0 + \beta_1 e_{ic}^{int} + \beta_2 L_{ic}^{basic} + \delta_c + \beta_3 e_{ic}^{int} \times L_{ic}^{basic} + \epsilon_{ic}^{int}, \tag{3}$$

$$L_{ic}^{adv} = \lambda_0 + \lambda_1 e_{ic}^{adv} + \lambda_2 L_{ic}^{int} + \delta_c + \lambda_3 e_{ic}^{adv} \times L_{ic}^{int} + \epsilon_{ic}^{adv}, \tag{4}$$

where L_{ic}^{int} is a sufficient statistic for previously accumulated knowledge in equation 4, 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 2. Similarly the technology that maps effort inputs to the cumulative final exam performance is defined analogously to equation 4.

Other Parameterization of the Learning Technology. The main specification of the cumulative learning process is represented by equations 2, 3, and 4. These are designed to be easily interpretable and minimalistic representation of the cumulative STEM learning process. I now discuss more complex representations of the cumulative learning process that may better fit the data.[22] For example, to allow for diminishing marginal returns to effort exertion, I consider the following specification in the basic stage:

$$L_i^{basic} = \alpha_0 + \alpha_1 e_i^{basic} + \alpha_2 (e_i^{basic})^2 + \alpha_3 h_i + \alpha_4 e_i^{basic} \times h_i + \epsilon_i^{basic}, \tag{5}$$

with $\alpha_2 < 0$ representing diminishing returns to effort in the basic stage. Adapting similar specifications to the intermediate and advanced stage will also allow for diminishing marginal returns

[22] I omit the cohort index for simplicity of the exposition.

to effort, but with a more complex characterization of concavity due to the dynamic complementarities.

Instead of using the prior stage grade as a sufficient statistics for previous knowledge, cumulative learning can be modelled entirely using historical effort inputs. For example, for learning in the intermediate stage could be represented by:

$$L_i^{int} = \beta_0 + \beta_1 e_i^{int} + \beta_2 e_i^{basic} + \beta_3 h_i + \beta_4 e_i^{basic} \times e_i^{int} + \beta_5 e_i^{int} \times h_i + \epsilon_i^{int}. \tag{6}$$

Now we have two separate dynamic interaction parameters, β_4 and β_5.

Instead of imposing linearity, a non-linear function as follows could be considered instead:

$$L_i^{int} = \lambda(\theta_1 e_i^{int} + \theta_2 e_i^{basic} + \theta_3 h_i)^{\gamma} + \epsilon_i^{int}, \tag{7}$$

where $\lambda > 0$ governs the curvature, and $\gamma < 0$ represents the concavity of the learning production function. Alternatively, the non-linear accumulation of knowledge can be represented by a CES production function as follows:

$$L_i^{int} = \lambda(\theta_1 (e_i^{int})^{\sigma} + \theta_2 (L_i^{basic})^{\sigma})^{\frac{1}{\sigma}} + \epsilon_i^{int}, \tag{8}$$

where $\sigma > 0$ governs the elasticity of substitution between present effort and previous knowledge. The CES production function can be linearized as a translog production function for estimation:

$$\begin{split} ln(L_i^{int}) = ln(\lambda) + \theta_1 ln(e_i^{int}) + \theta_2 ln(L_i^{basic}) + \theta_{11} ln^2(e_i^{int}) + \\ \theta_{22} ln^2(e_i^{int}) + \theta_{22} ln^2(L_i^{basic}) + \theta_{12} ln(e_i^{int}) ln(L_i^{basic}) + \epsilon_i. \end{split} \tag{9}$$

When the coefficients associated with the quadratic term are 0, note that equation 3 and equation 9 have an analogous structure. That is, the education production function being estimated by equations 2, 3, and 4 are sufficiently flexible to capture features of a CES production function.

Identification of the Learning Technology. Identifying the cumulative technology requires exogenous variation 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.

8.2 Estimation

The marginal benefit of effort parameters are estimated using 2SLS by 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. Following Angrist (2006), I estimate the learning technology with two endogenous variable as follows:

$$\begin{cases} L_{ic}^{int} = \beta_0 + \beta_1 e_{ic}^{int} + \beta_2 L_{ic}^{basic} + \delta_c + \beta_3 e_{ic}^{int} \times L_{ic}^{basic} + \epsilon_{ic}^{int}, \\ e_{ic}^{int} = \rho_0 + \rho_1 RemindersFreqBasic_{ic} + \rho_2 RemindersFreqInt_{ic} + \delta_c + \epsilon_{ic}^{int} \\ L_{ic}^{basic} = \gamma_0 + \gamma_1 RemindersFreqBasic_{ic} + \gamma_2 RemindersFreqInt_{ic} + \delta_c + \epsilon_{ic}^{int} \end{cases}$$

where $RemindersFreqBasic_{ic}$ is the number of reminders received in the basic stage, and $RemindersFreqInt_{ic}$ is the number of reminders received in the intermediate stage. If reminders are effective in inducing effort exertion at each learning stage, we expect $\rho_2 > 0$ and $\gamma_1 > 0$. The advanced learning technology is estimated analogously.

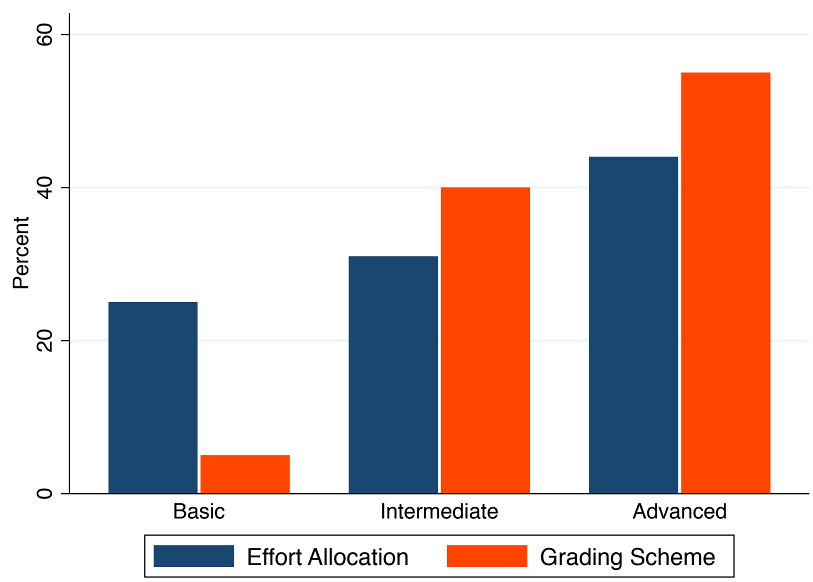
8.3 Discussion of Parameter Estimates

The marginal benefit parameter estimates are shown in Table 6. The estimates show a positive marginal benefit of effort at each learning stage. For example, at the basic stage, an additional hour spent studying increases students basic stage homework grade by around 0.14 SD for the average student. The estimates also show that exerting effort on the low-stakes homework assessments results in higher midterm and final exam performance. 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 homework performance. The evidence is consistent with the online homework assessments being a primary source of learning in this course.

The estimates also indicate evidence of dynamic interactions in effort inputs across the learning stages. Complementarities in present effort exertion and previous knowledge are present in both the intermediate and advanced learning stages as β_3 and λ_3 are both statistically significant from 0. For example, the marginal benefit of studying for an hour in the advanced learning stage is increasing by 0.08 SD for each every 1 SD increase in intermediate homework performance. The results are consistent with most students having no prior programming experience, and also reflect the cumulative learning structure of programming. That is, students are accumulating programming knowledge over time through their effort exertion across varying programming topics that naturally build on each other.

Overall, the results suggest that students learning effectively should appropriately front-load their effort allocation, thereby becoming proficient in foundational skills that serve as the building blocks for rest of the course. Figure 5 illustrates the proportion of total effort, on average, allocated over the three learning stages and the corresponding grading weight.

Figure 5: Average Effort Allocation and Grading Scheme Across Learning Periods



Notes: The figure presents the mean proportion of total effort allocated across the basic, intermediate, and advanced learning stages. The corresponding grading weights are also presented.

Figure 5 shows that contrary to the effective learning strategy implied by the results, the average student's efforts are loaded towards the later learning stages. Such an effort allocation is consistent with the course grading scheme which increases as the course progress, as is typical in most courses. In other research, I use survey data further to characterize students who are at risk of falling behind early on, and those that study effectively as implied by the learning technology. Additionally I also explores the role of grading scheme design in encouraging students to study effectively.

9 Conclusion

The share of students pursuing education in STEM has been increasing rapidly over the past decade. Although STEM graduates are a vital input to any modern society, learning advanced skills in STEM can be challenging given its cumulative nature. Identifying the cumulative learning process is beneficial for understanding effective learning strategies and approaches to course design. In light of the benefits, this paper presented an approach to credibly estimating the learning process while circumventing the typical issues of measurement error and endogenity of effort inputs.

The empirical approach I developed involves several components. I employed rich administrative data from a large pre-existing foundational online STEM course that has a cumulative structure. I then conducted randomized interventions that were successful in nudging students to spend more time learning and complete more online assignments. The administrative dataset, including precise measures of student study time allocation, together with the field experiment were then used to estimate a cumulative education production function. Exogenous experimental variation arising from the field experiment served to credibly identify the marginal benefit each stage of the cumulative learning process built into the STEM course.

The findings presented help to inform the design of large foundational courses that apply to both online and traditional in-person setups. First, completing low-stakes online assignments throughout the course is essential for student learning: spending an extra hour on online assignments increases final exam grades by 0.09 SD (noting that online homework is the key means of learning in the course). Second, I find evidence of strong dynamic learning complementarities in the cumulative education production function. That is, the productivity of effort learning advanced skills is increasing in prior knowledge. Contrary to this evidence, however, I find that students effort allocation increases as the course progresses. Then, given a cumulative course structure, an instructor can consider setting the assignment grading weights so that students exert sufficient effort early on learning the fundamentals. In a follow-up paper, I investigate the design of grading schemes that best encourage students to allocate effort to learn effectively in courses with a cumulative structure and heterogeneous students.

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.

Angrist, Joshua D (2006) "Instrumental variables methods in experimental criminological research: what, why and how," *Journal of Experimental Criminology*, 2 (1), 23–44.

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.

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.

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

Ericson, Keith Marzilli (2017) "On the interaction of memory and procrastination: Implications for reminders, deadlines, and empirical estimation," *Journal of the European Economic Association*, 15 (3), 692–719.

Ersoy, Fulya (2021) "Returns to effort: experimental evidence from an online language platform," *Experimental Economics*, 24 (3), 1047–1073.

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

Hanushek, Eric A and Ludger Woessmann (2015) *The knowledge capital of nations: Education and the economics of growth*: MIT press.

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.

Lee, David S, Justin McCrary, Marcelo J Moreira, and Jack Porter (2022) "Valid t-ratio Inference for IV," *American Economic Review*, 112 (10), 3260–90.

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.

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.

Stinebrickner, Ralph and Todd R Stinebrickner (2008) "The causal effect of studying on academic performance," *The BE Journal of Economic Analysis & Policy*, 8 (1).

——— (2014) "A major in science? Initial beliefs and final outcomes for college major and dropout," *Review of Economic Studies*, 81 (1), 426–472.

Stinebrickner, Todd and Ralph Stinebrickner (2012) "Learning about academic ability and the college dropout decision," *Journal of Labor Economics*, 30 (4), 707–748.

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
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 (see Appendix C.2). Panel D statistics are formulated from the administrative data of the online homework platform (C.3). Finally, the statistics shown in Panel E are computed using data gathered from the discussion board.

Table 2: Student Attrition and Treatment Allocation

	(1) I(Dropped course)	(2) I(Dropped course)
No. of reminder messages received	0.0116	-0.0061
	(0.0898)	(0.0192)
Controls	No	Yes
No. of Students	4091	4091
R-squared	0.0015	0.13

Notes: Table shows differential attrition rate by intensity of the treatment condition. Controls include pre-treatment student demographics 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 3: 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 4: 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 5: Efficacy of Reminders by Study Group Involvement

	(1) Homework Completed[b]	(2) Homework Completed
I(Study group) $\times$ No. of reminders	0.061	0.043
	(0.0537)	(0.0317)
Controls	No	Yes
Adjusted R-square	0.142	0.371
No. of Students	3686	3686

Notes: Indicator for whether a student is in a study group and the number of reminders received 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 $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

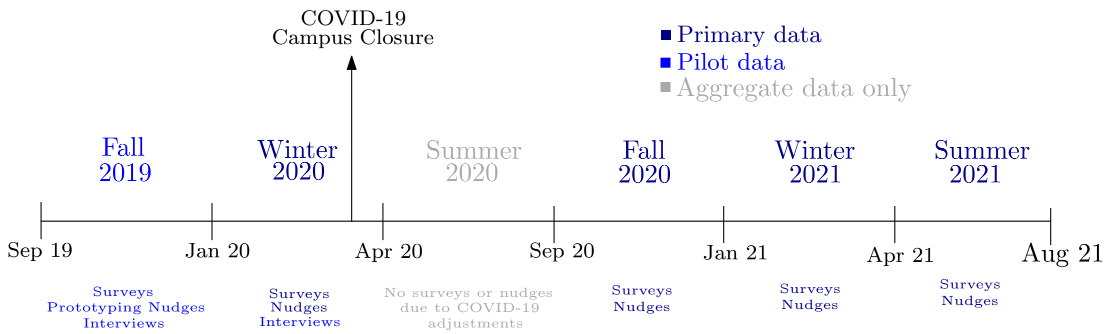
Table 6: Benefit of Effort Parameter Estimates (2SLS)

Parameter	Estimate (SE)
Panel A: Basic Learning Stage Performance	
$\widehat{\alpha_1}$ (basic 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	
$\widehat{\beta_1}$ (intermediate 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	
$\widehat{\lambda_1}$ (intermediate 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 $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

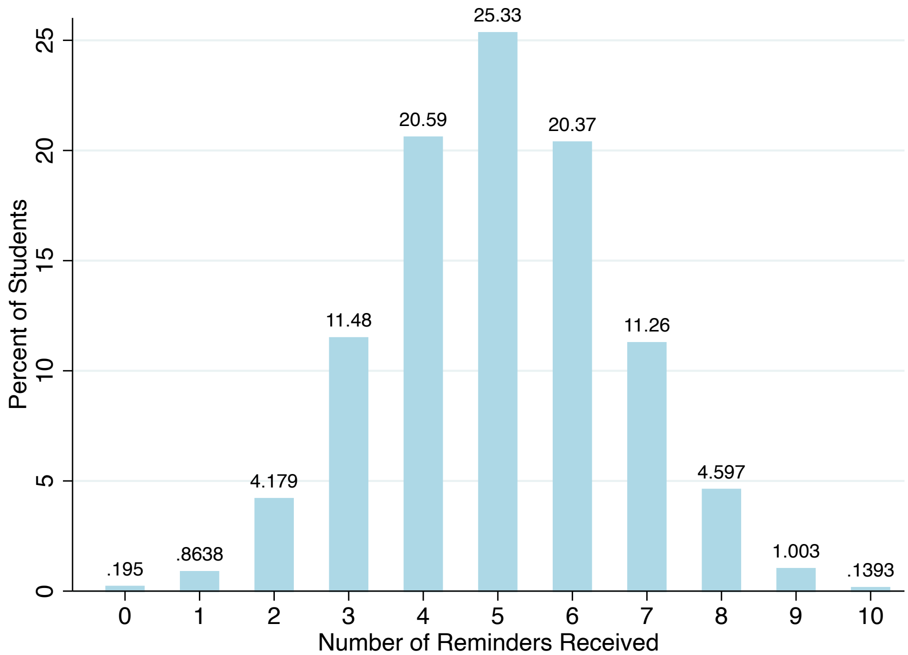
Figures

Figure 6: Timeline of Data Collection



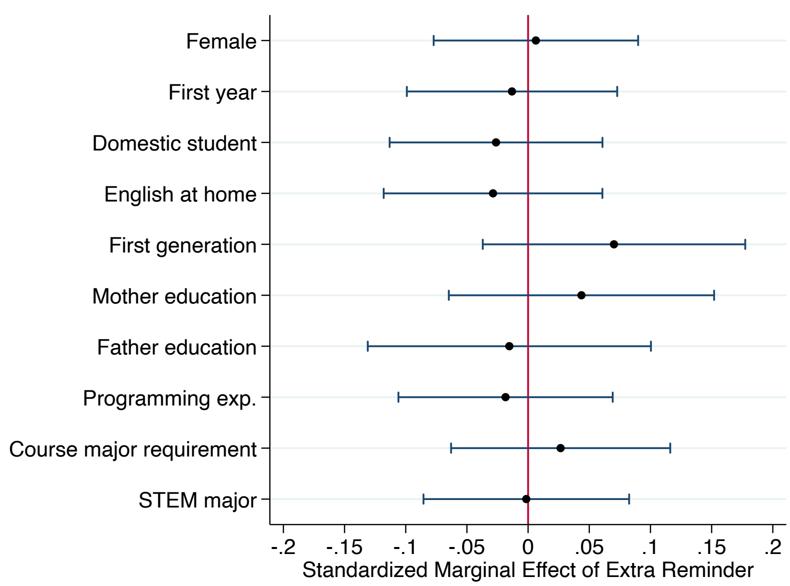
Notes: The figure illustrates the timeline of the data collection. The primary data collection is gathered from the Winter 2020, Fall 2020, Winter 2021, and Summer 2021 cohorts. Pilot data is gathered from the Fall 2019 cohort and involved conducting interviews with students and instructors, surveying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

Figure 7: Distribution of Homework Reminders Received



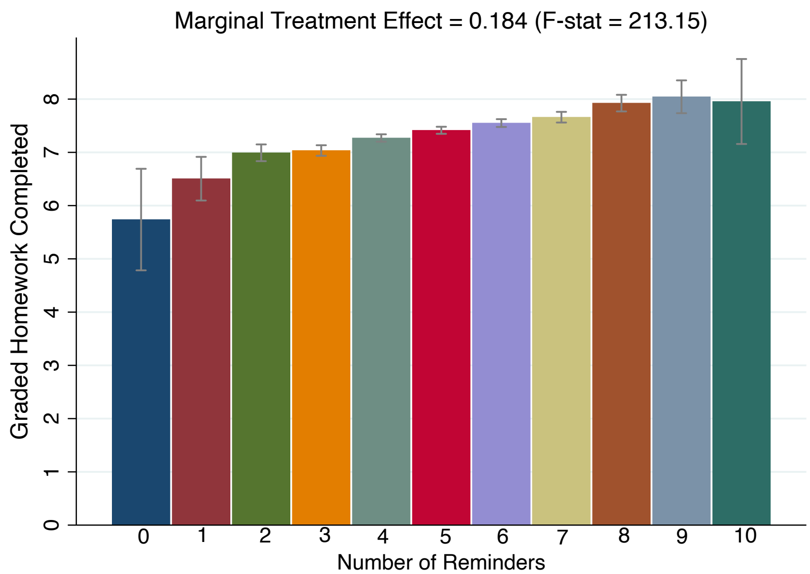
Notes: The figure shows the distribution of reminders received by students. For each of the 10 graded homework assessments, half of the students are randomly selected to receive a reminder. Then each student is eligible to receive between 0 and 10 homework reminders in total.

Figure 8: Student Demographic and Characteristics Balance Check for Homework Reminders



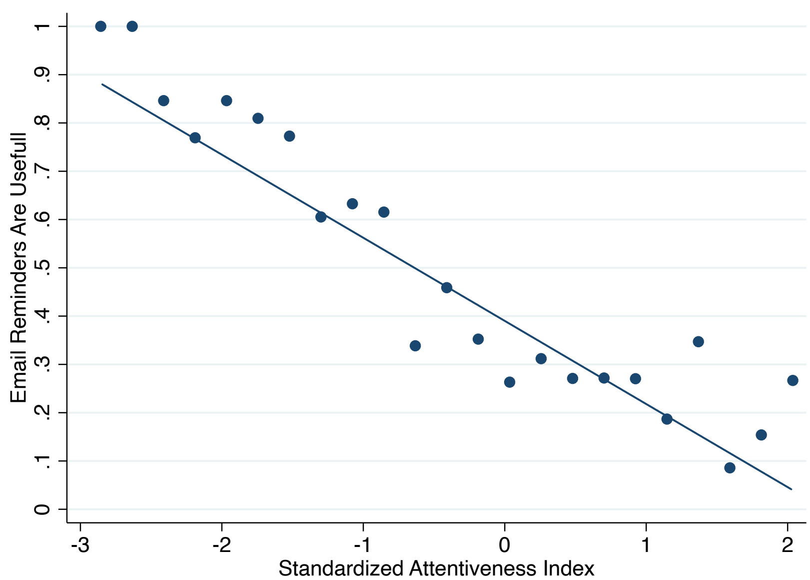
Notes: The estimates displayed are computed by regressing each standardized variable presented on the vertical axis on the number of reminders received. The error bars represent 95% confidence intervals.

Figure 9: 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 10: 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 **??**).

A Appendix: Summary of Related Literature

A.1 Related Education Production Function Literature

Table 7: 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.2 Related Student Effort Literature

Table 8: 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.

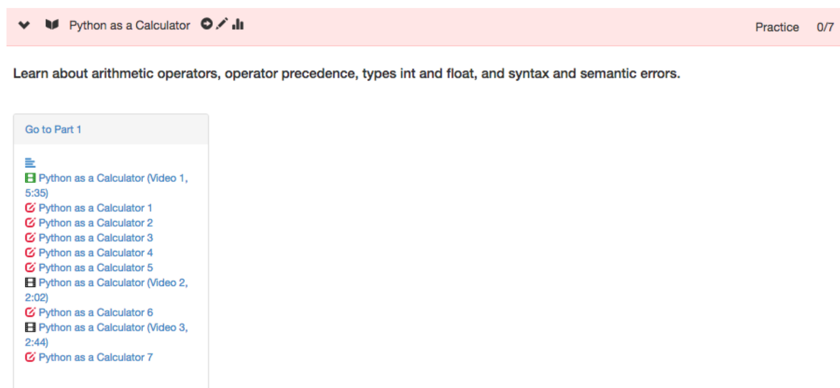
B.2 Online Homework Environment

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	

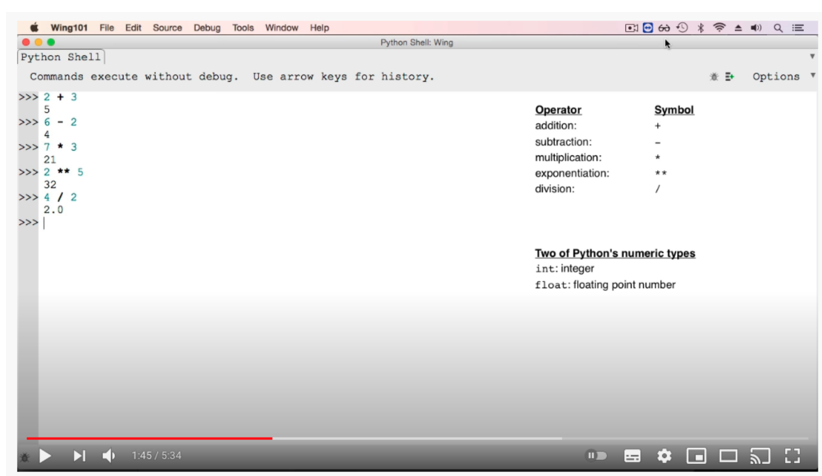
Students provided an outline for how to learn a topic:

Figure 1: Outline for Learning Numerical Operations



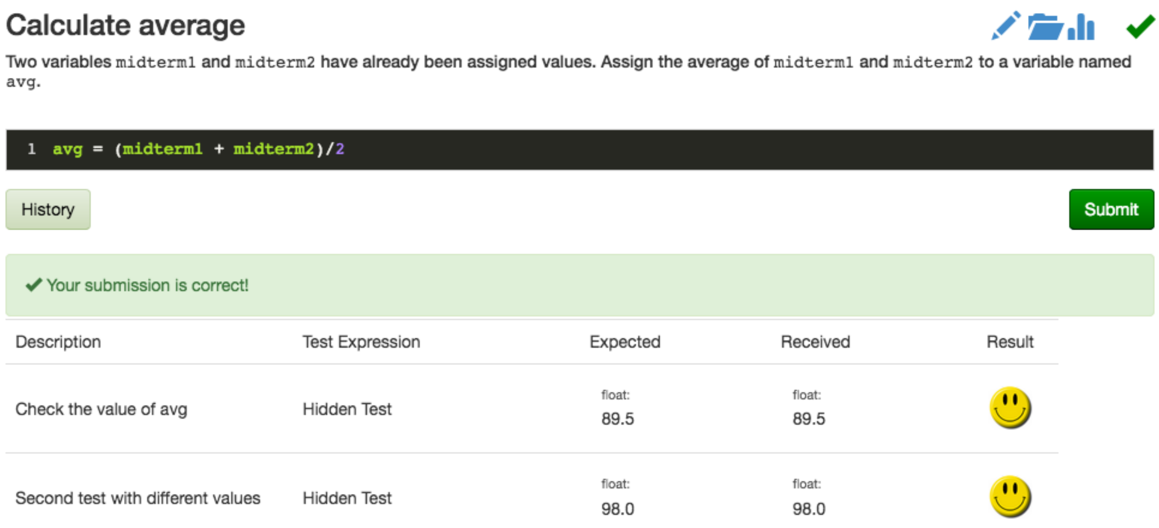
Students begin the course by watching a video about numerical operations in Python:

Figure 2: Video on Numerical Operations



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

Figure 3: Sample Coding Exercise



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 9: 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

Table 10: 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

C.2 Measuring attentiveness

The baseline survey elicits a student's attentiveness 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 attentiveness.

Students are classified as attentive if they responded with at least 5 to the first three questions in Table 10, and at most 3 to the last question.

C.3 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.3.1 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.3.2 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.3.3 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.4 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 contents of the homework reminder messages with other students in the course [Likert Scale]

D Appendix: Nudges

The reminder messages 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 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