



Does Inducing Students to Schedule Lecture Watching in Online Classes Improve Their Academic Performance? An Experimental Analysis of a Time Management Intervention

Rachel Baker¹ · Brent Evans² · Qijie Li¹ · Bianca Cung¹

Received: 13 September 2017 / Published online: 28 July 2018
© Springer Nature B.V. 2018

Abstract

Time management skills are an essential component of college student success, especially in online classes. Through a randomized control trial of students in a for-credit online course at a public 4-year university, we test the efficacy of a scheduling intervention aimed at improving students' time management. Results indicate the intervention had positive effects on initial achievement scores; students who were given the opportunity to schedule their lecture watching in advance scored about a third of a standard deviation better on the first quiz than students who were not given that opportunity. These effects are concentrated in students with the lowest self-reported time management skills. However, these effects diminish over time such that we see a marginally significant negative effect of treatment on the last week's quiz grade and no difference in overall course scores. We examine the effect of the intervention on plausible mechanisms to explain the observed achievement effects. We find no evidence that the intervention affected cramming, procrastination, or the time at which students did work.

Keywords Scheduling · Procrastination · Online courses · Persistence · Performance · Achievement

Introduction

Concerns about student learning and poor academic performance are prevalent among parents, faculty, administrators, and policymakers in postsecondary education. Empirical evidence suggests that these concerns are not unfounded. Research indicates that student learning and academic growth is limited in some sectors of modern higher education (Arum and Roksa 2011) and academic performance, as measured by grades, can affect a student's

✉ Rachel Baker
rachelbb@uci.edu

¹ University of California, Irvine – School of Education, 3000 Education, Irvine, CA 92697, USA

² Peabody College, Vanderbilt University, 230 Appleton Place, PMB 414, Nashville, TN 37203, USA

eligibility for financial aid, choice of major, and probability of dropping out (Griffith 2010; Rask and Tiefenthaler 2008; Schudde and Scott-Clayton 2016; Stratton et al. 2008). Recently, new modalities, such as fully online courses, have posed novel challenges to students and instructors. These modalities are associated with even lower levels of learning and persistence (Angelino et al. 2007; Cochran et al. 2014; Evans et al. 2016; Figlio et al. 2013; Leeds et al. 2013; Moody 2004; Perna et al. 2014; Rovai 2003; Xu and Jaggars 2013).

Institutions and scholars are searching for causes of and solutions to low academic performance. One promising area of focus is the inherent need for advanced time management skills in college classes and individual student characteristics associated with poor time management. Babcock and Marks (2011) have documented the decline of study time over the last several decades and have suggested that an increase in leisure activities might be a cause. Relatedly, a long line of empirical evidence suggests that poor time management and relatively few study hours are predictors of poor academic performance (Beattie et al. 2017; Macan et al. 1990; Trueman and Hartley 1996).

While most of the extant research has focused on traditional, in-person higher education, there are at least two reasons to believe time management is even more critical in online higher education. First, the asynchronous setting, with few, if any, set times at which a student must do work or participate in class does not generate an environment with a consistent schedule. This greater flexibility forces students to make more decisions about when to do work and puts greater demands on the self-regulatory skills necessary for making plans for learning (Broadbent and Poon 2015; Schwartz and Ward 2004; Tuckman 2005). Second, the lack of a face-to-face connection and joint social presence with instructors and classmates creates fewer opportunities for extrinsic accountability, which can negatively affect student motivation and the extent to which students carry out their plans (Black and Deci 2000; Bowers and Kumar 2015; Miltiadou and Savenye 2003; Mullen and Tallent-Runnels 2006; Zhan and Mei 2013).

These structural issues may disproportionately harm students with worse time management skills. Several studies have identified student characteristics, such as student motivation, student interaction in the course, proclivity towards self-regulated learning, and time management skills, that predict success in online courses (Cochran et al. 2014; Hart 2012; Rostaminezhad et al. 2013). Additionally, a meta-analysis of self-regulated learning in online higher education identified that effective time management is positively related to academic outcomes (Broadbent and Poon 2015).

Continued focus on these issues is especially important as higher education is rapidly expanding into the online world. Both traditional and for-profit providers are offering a greater share of credit bearing courses in certificate and degree programs online and increasing numbers of college students enrolled in traditional postsecondary programs are taking at least one class online. About 28.5% of college students take at least one online course and 14% of college students take all of their classes online (National Center for Education Statistics (NCES) 2015).

This trend is unlikely to stop; over 60% of chief academic officers across institutions of higher education say online education is part of their long-term institutional strategy (Allen et al. 2016). This growth is motivated by the opportunity to expand enrollments with minimal investments in infrastructure and the view that providing postsecondary content online may slow the rapidly growing costs of delivering higher education (Deming et al. 2015).

Given the increasing importance of online higher education and the link between time management and students' academic performance, many institutions, hoping to promote positive academic outcomes in online courses, are seeking simple, low-cost, scalable interventions aimed at improving time management. Because several studies

have suggested time management skills are critical to success in online courses (Elvers et al. 2003; Michinov et al. 2011; Nawrot and Doucet 2014; Roper 2007), and providing a scheduling structure for completing tasks has been identified as a principle in the successful design of Massive Open Online Courses (MOOCs) (Guàrdia et al. 2013), the goal of this paper is to test the efficacy of a scheduling intervention in an online post-secondary course. We seek to answer one primary research question: what is the effect of encouraging students to schedule their coursework on academic performance in an online, for-credit postsecondary course?

To answer this question, we employ a randomized control trial testing the effects of a low-cost, scalable scheduling intervention on course achievement in an online, for-credit course for degree seeking students in a 4-year selective public college. The intervention we examine is a suggestion by the course instructor that students schedule when to watch the lecture videos and an online survey in which they could set their lecture watching schedule. The suggestion was delivered to a randomly selected group of students in each of the first 2 weeks of the course. Treated students were asked to state when (day of the week and hour of the day) they intended to watch the daily lectures in the first 2 weeks of the course. In order to ensure that the control students had an equal number of contacts from the instructor in an effort to isolate the effect of the scheduling survey, control students received an email from the course instructor with a link to a survey asking them which web browser they used to access the course (week one) and whether they listened using speakers or headphones (week two). Screenshots of both surveys are provided in Fig. 1.

Although many studies have previously explored the correlation between time management, scheduling, and student outcomes in higher education, there is a dearth of evidence on the causal effect of scheduling on student performance in online education. The causal evidence that does exist is limited to evidence from three studies that took place in MOOCs (Baker et al. 2016; Kizilcec et al. 2016; Patterson 2014). The MOOCs in these three papers present a very different context, and it is questionable whether or not these results can be generalized to students in online classes offered by traditional colleges. In each of the prior studies, the MOOCs were free, open access, and students could not earn college credits for successfully completing them. As such, student motivation to take and complete these courses is likely very different from motivation to pay for and enroll in a credit-bearing online course. This difference in motivation for enrolling leads to very different student populations and likely leads to substantially different behaviors and outcomes among students. Therefore, it is important to test scheduling interventions in for-credit online classes as well as in open access non-credit bearing courses such as MOOCs.

In addition to the different context, the current study also extends the previous literature by examining heterogeneity of treatment across student characteristics. We measure students' perceived time management skills using a pre-course survey. This enables us to assess whether our intervention, which targets improving time management through scheduling, has differential effects on students with high and low self-reported time management skills. We hypothesize that the treatment effects will be stronger for students who report having poor time management abilities.

A final contribution of our analysis is a consideration of potential mechanisms. We have access to many student behaviors within the course (e.g. timing of watching lectures and completing assignments). These detailed clickstream data allow us to observe micro interactions and examine if the scheduling intervention has effects on student behaviors such as procrastination and cramming. This level of detail enables us to examine potential mechanisms through which the scheduling intervention may yield effects on course performance.

A Scheduling Survey

Introduction

Happy first day of class! I hope you, like me, are looking forward to the next five weeks. This is a reminder that the Week 1 lecture videos for the course are available.

Because it is offered online, this class does not have scheduled class meeting times. Research finds that students who treat online classes like in-person classes and schedule their work time in advance get more out of their courses. Toward that end, please complete this short questionnaire to schedule when you intend to watch the videos for this week.

*: Required question

1.

Please select what day and time you will watch the **Lecture 1** video.

If you have already watched Lecture 1, select '*I already watched this video*' when selecting the day.

I will start watching the **Lecture 1** video on:

-- Select an answer... --

at

-- Select an answer... --

2.

Please select what day and time you will watch the **Lecture 2** video.

If you have already watched Lecture 2, select '*I already watched this video*' when selecting the day.

I will start watching the **Lecture 2** video on:

-- Select an answer... --

at

-- Select an answer... --

B Control survey

Introduction

Happy first day of class! I hope you, like me, are looking forward to the next five weeks. This is a reminder that the Week 1 lecture videos for the course are available.

Please help us make your experience taking online classes better by clicking on the following link and answering the survey question about which browser you will use to watch the first lecture video.

1. I will use (or used) the following browser to watch the first lecture video:

- Internet Explorer
- Firefox
- Chrome
- Safari
- Other

Fig. 1 a Scheduling survey. b Control survey

Our employment of a randomized control trial enables us to causally assess the effect of encouraging scheduling on a variety of achievement outcomes. To summarize our results, we find high levels of take-up of the treatment (92–93%) and positive treatment effects of the scheduling intervention on initial quiz grades. However, these effects diminish over time such that we see a marginally significant negative effect of treatment on the last week's quiz grade and no difference in overall course scores. In support of our hypothesis, we find that the positive effects of the first quiz score are concentrated among students who report poor time management skills. We do not find evidence that

the treatment affected student lecture watching behaviors such as procrastination and cramming.

Prior Literature & Theory

Academic Performance in Online Higher Education

A great deal of effort has gone into identifying the effects of online postsecondary coursework on student performance. Lack (2013) cites numerous observational studies that examined the difference in academic performance between online and face-to-face courses. These studies found inconsistent results, but all were plagued by selection bias. Xu and Jaggars (2013) improved on observational studies by comparing the same students when they take face-to-face versus online coursework and found students generally performed worse in online delivery. The best identified studies have shown mixed results. Bowen et al. (2014) randomly assigned students to a hybrid versus face-to-face statistics course at six 4-year public institutions and observed no difference in learning outcomes between the two delivery formats. However, in another randomized control trial, Figlio et al. (2013) found male, Hispanic, and lower achieving students performed worse when assigned to an online economics course relative to their peers assigned to the in-person version.

Although these studies investigated different interventions in different contexts, the balance of evidence suggests that students perform no better, and oftentimes worse, in for-credit online postsecondary coursework relative to traditional face-to-face coursework. However, from an institutional cost–benefit perspective, and in the context of limited resources, it is possible to argue that potential performance loss is worth the substantial savings, to schools and to students, of delivering content online. However, that debate obscures what we consider to be the more relevant question: can we improve learning outcomes in the online context? To accomplish this goal, we must identify the critical challenges students encounter in online education and help students overcome those challenges.

Time Management & Scheduling Study Time

One critical challenge in online classes is time management. Prior research has repeatedly demonstrated that time management is an important skill related to college performance in both face-to-face and online postsecondary classes. Poor time management and fewer study hours are the leading predictors of poor academic performance in a traditional 4-year college education setting (Beattie et al. 2017). Specifically, studying course materials throughout the term, as opposed to cramming right before a deadline, is positively correlated with a higher college GPA (Hartwig and Dunlosky 2012). Similarly, Macan et al. (1990) found that scores on a robust time management scale were positively related not only to higher college GPA but also to higher students' self-perceptions of performance and general satisfaction with life. College students with better time management skills both scored higher on cognitive tests and were more efficient students, spending less total time studying (Van Den Hurk 2006). There is not a large literature focusing explicitly on the scheduling component of time management, but short range planning, including scheduling study time, has been found to be more predictive of college grades than SAT scores (Britton and Tesser 1991).

Important for our specific context, these results have been shown to extend to online learning settings. In a study of online learners who completed degrees, students identified that developing a time management strategy was critical to their success (Roper 2007). Guàrdia et al. (2013) argue that providing a scheduling structure with clear tasks is one of ten critical design principles for designing successful MOOCs.

Hypotheses to explain why scheduling is so important are thin, but we propose several potential mechanisms. Scheduling could simply encourage students to spend more time on their coursework. By scheduling when to work, it is more likely students will spend that scheduled time on their classes rather than on alternate activities.

It is also possible that planning induces more efficient studying by reducing the probability that students will do work at non-ideal times of day. Prior research on adolescents links improved performance on intelligence assessments with working during students' preferred time of day (Goldstein et al. 2007), and a study of college students demonstrated that starting classes later in the morning improves academic performance (Carrell et al. 2011). It is likely that if students establish and stick with a schedule, they are more likely to complete work during ideal times.

Another possible explanation of the importance of scheduling is that it reduces the likelihood of students cramming a lot of work into a short period of time or putting off work until just before the deadline. Cramming and procrastination have both been found to be negatively related to success in online classes (Elvers et al. 2003; Michinov et al. 2011).

A final potential mechanism is that time management is an effective strategy to reduce academic stress and anxiety, which in turn may increase performance (Misra and McKean 2000). We are able to explore the time of day, cramming, and procrastination mechanisms with empirical data in our analysis below.

Theory of Action of the Scheduling Intervention

The goal of our study, unlike much of the prior literature, is not to survey students about their study strategies to look for a relationship between study skills and academic outcomes. Rather, given the consistent evidence that good time management practices are associated with positive outcomes, we attempted to manipulate students' time management practices by encouraging students to schedule their study times. Such an intervention is particularly important in asynchronous online contexts which lack structure, as there are no scheduled class meeting times. Most online courses have weekly deadlines for submitting assignments but lack any sort of meeting schedule like those traditionally found in face-to-face courses. Our goal is to induce students to improve their time management by scheduling when they will watch the lecture videos. Our hypothesis is that the scheduling will result in improved academic performance.

Our intervention encourages creating a structure and timeline in an otherwise unstructured course environment. Self-defined course schedules continue to provide flexibility, a notable advantage of online education, by allowing students to choose when they will watch the lecture videos throughout the week. However, by committing to the days and times they choose, students should be more likely to hold themselves to that schedule instead of putting off the online coursework in favor of other more immediate demands.

Our intervention functions similar to a precommitment device, which has been tested in the economics literature in a variety of contexts. People's preferences tend to change over time, such that intentions to engage in a particular activity in the future are often revised when the future arrives and a different preference takes precedence (Frederick et al. 2002).

A precommitment device works by binding a person's future behavior to reduce the risk of succumbing to immediate desires. For example, at the beginning of the week, a student may intend to do work for their online class every evening, but when each evening approaches, the student will choose another more appealing activity, thereby delaying or forgoing their coursework. By committing in advance and formalizing their schedule, students may be more likely to adhere to their coursework commitment.

Precommitment devices have been found to be effective in diverse contexts such as employee effort (Kaur et al. 2013), smoking (Gine et al. 2010), and savings (Ashraf et al. 2006). One study from higher education (Ariely and Wertenbroch 2002) has demonstrated a demand for, and a positive effect of, precommitment devices that aim to affect student effort and behaviors on out-of-class assignments.

The intervention we study in this paper functions as a type of precommitment device wherein students are given the opportunity to shape their own future behavior. This intervention is in line with Song et al. (2004) and Nawrot and Doucet (2014), who directly called for interventions targeting the development of time management strategies in online classes. To our knowledge, there are only a few tests of similar interventions in online higher education, all of which were conducted in free, open-access MOOCs that did not confer college credits. One study examined the efficacy of a similar intervention in a MOOC (Baker et al. 2016), and found a small and surprisingly negative effect of scheduling on completing the course and on course performance as measured by final grade. A work in progress paper presented at the Learning @ Scale conference provided a similar test by randomly informing a sample of MOOC users that a set of study skills have been reported as effective (Kizilcec et al. 2016). That study found no effect on engagement and persistence outcomes. Finally, Patterson (2014) randomly assigned MOOC students to a more costly commitment: installing software on their computers that limits access to distracting websites (news, Facebook, etc.). The experiment found large effects of the treatment, including an 11 percentage point increase in course completion and more than a quarter of a standard deviation increase in course performance. These larger effects are likely driven by the intrusive commitment device and only apply to the small subset of course registrants (18%) who volunteered for the commitment device after being offered a financial incentive.

Our study improves upon the prior literature by examining the scheduling commitment in the context of a credit-bearing traditional online higher education course. The prior studies have limited external validity to for-credit coursework due to the different student expectations, motivations, and behaviors in a free MOOC that does not confer college credits. Additionally, the prior studies report mixed results with positive, negative, and null effects represented across the three MOOC experiments, which necessitates further evidence of the utility of this type of intervention. Furthermore, we are the first study able to examine heterogeneous effects of this intervention across reported time management skills and to be able to investigate plausible mechanisms of the scheduling intervention.

Setting, Data, and Methods

Experimental Setting

Course

We conducted our study in an online undergraduate STEM course lasting 5 weeks in a selective, public 4-year university. The course was in the summer term, was offered online (though the final exam was given in person), and conferred full credit as if the enrollees had taken the same 10-week course in person during the academic year. The course was lower division, required for the major, and had calculus as a pre- or co-requisite. Although the prior courses in which a similar intervention has been evaluated are also STEM courses, we note that our setting differs in that it is for credit, not open access, not a MOOC, and slightly more advanced given that calculus is a co-requisite.

Taking courses over the summer is common among undergraduate students at this school; 35.7% of students take at least one summer class over the course of their career and the school promotes summer courses as a way to graduate on time. There are two summer terms offered by the university. Each is only 5 weeks long, so many students take one term of summer classes and work or have an internship for the remaining 2 months of the summer.

Students had to meet frequent deadlines throughout the course. Each week required students to watch five lecture videos with in-video quizzes, complete daily homework assignments, complete weekly “challenge problem sets,” and take weekly quizzes. Lecture videos were each approximately 40–50 min in length. To receive credit for watching the lecture videos and completing the in-video quizzes, students had to complete all videos by midnight on Friday each week. Daily problem sets were due every day from Monday to Saturday, challenge problems sets were due once or twice a week, and weekly quizzes were only available on Sundays. In addition to these assignments, the course also had a final exam that was held on campus on the last day of the course. Students’ final grades were determined by participation in watching lectures (15%), performance on daily problem sets (15%), scores on the weekly quizzes (30%), and the final exam score (40%).

Participants

A total of 176 students enrolled in this course and more than half of these students had taken at least one online class before. Nineteen of these students did not participate in our experiment because they enrolled in the course after randomization, hence 157 students were randomly assigned.¹ Prior to random assignment, and before the course officially started, a pre-course survey, which was worth 1% of extra credit toward the course grade, was sent out. It included questions about student characteristics (e.g., major, first language, and summer housing), self-assessments of self-regulation skills, expected time commitment for the course (expected hours of work per week), and self-assessments of time-management skills. Among the 157 students in the potential sample, 145 of them completed the pre-course survey. We use the pre-course survey responses to examine heterogeneity

¹ Due to an administrative error, two students who were assigned to the control condition were given the treatment scheduling surveys. We consider this a compliance issue and leave them in our ITT estimates.

of treatment effects across key student groups. For example, we identify students with low self-reported time management skills, as defined by responses to pre-course survey questions, and analyze the heterogeneity of the treatment effect by this characteristic. In an effort to maintain the same sample throughout the analysis, we use the 145 students who completed the pre-course survey as our analytic sample.

Although we observe the outcomes for the full 157 students randomized, we could consider the reduction in sample to 145 to be a form of pseudo-attrition at a rate of 7.6%. There is also a differential response rate to the pre-course survey across treatment and control assignments of 7.7%. Because the pre-course survey was conducted prior to random assignment, it is exogenous to treatment; therefore, these rates fall into the bounds of acceptable attrition as outlined by the What Works Clearinghouse (2017).²

Data

Demographic Variables

The institution provided administrative data on demographics and prior academic achievement including standardized test scores, prior course enrollment, and prior course performance for all students enrolled in the course who were also enrolled as degree-seeking students in the institution. Of the 145 students who completed the pre-course survey, full demographic information was available for 111 students (we had information on between 117 and 128 students for each demographic variable provided by the institution).³ We use demographic data to check for balance between our treatment and control groups (described below), but because of the reduction in sample size, do not include these covariates in our main analyses. Given the randomized nature of our study, the inclusion of these covariates should not affect our results, and indeed, we find similar results with this reduced sample and the addition of covariates as demonstrated in Appendix Table 7.

Table 1 presents summary statistics on student characteristics for the analytic sample, the treatment group, and the control group.⁴ Students in our samples were, on average, 20.3 years old (SD 1.2). There are slightly more female students than male students (55.5% compared to 44.5%), and our sample consisted of 67.5% Asian American, 12.8% Hispanic, 8.5% White, and 1.7% Black with 9.5% not indicating their race/ethnicity. The majority of the students were sophomores (70.9%), followed by freshman (15.7%), juniors (11.8%), and seniors (1.6%). Half of the students were first generation students (defined as neither parent having a college degree). Individual t-tests were conducted on the treatment and control group mean difference for each variable to assess experimental balance of the

² To further alleviate concerns related to this sample reduction, we present results from ITT and TOT models without covariates using the fully randomized sample of 157 students in the first two columns of Appendix Table 7. Reassuringly, results are consistent with the reported results from the pre-course survey sample of 145 students.

³ While most of the students enrolled in the course were degree-seeking students at the university, some (fewer than ten) students were enrolled in another university. We do not have demographic information available for these students, but we do have full data (including survey responses) for these students. They are included in our analyses.

⁴ Much of the data presented in this table come from institutional data provided to us by the school's institutional research office. Sample sizes vary based on which data were available for which students. The second set of variables were asked on the pre-course survey for this specific class and are available for all students in our analytic sample.

Table 1 Demographic characteristics by experimental condition

	Analytic sample (N = 145)			Control group (N = 69)			Treatment group (N = 76)			P
	M	SD	N	M	SD	N	M	SD	N	
	<i>Institutional administrative data</i>									
Age in years	20.295	(1.193)	128	20.283	(1.019)	60	20.306	(1.337)	68	0.916
SAT score										
Math	637.258	(74.357)	124	637.069	(77.437)	58	637.424	(72.140)	66	0.979
Verbal	577.258	(88.734)	124	567.931	(87.032)	58	585.454	(90.063)	66	0.274
Writing	590.484	(83.402)	124	588.103	(84.427)	58	592.576	(83.081)	66	0.767
Female	0.555		128	0.533		60	0.574		68	0.648
Race										
White	0.085		117	0.091		55	0.081		62	0.843
Hispanic	0.128		117	0.127		55	0.129		62	0.977
Black	0.017		117	0.036		55	0.000		62	0.13
Asian/Pacific Islander	0.675		117	0.673		55	0.677		62	0.957
Other/did not indicate	0.095		117	0.073		55	0.113		62	0.457
Year in college										
Freshmen	0.157		127	0.150		60	0.164		67	0.827
Sophomores	0.709		127	0.667		60	0.746		67	0.324
Juniors	0.118		127	0.150		60	0.090		67	0.292
Seniors	0.016		127	0.033		60	0.000		67	0.132
Low income (yes/no)	0.346		127	0.367		60	0.328		67	0.651
First generation college student (yes/no)	0.500		122	0.526		57	0.477		65	0.586
Full time status for summer term (yes/no)	0.578		128	0.617		60	0.544		68	0.407
<i>Pre-course survey data</i>										
On campus housing (yes/no)	0.235		145	0.217		69	0.250		76	0.939
English as first language (yes/no)	0.559		145	0.551		69	0.566		76	0.855
Self-confidence	3.779	(0.893)	145	3.957	(0.865)	69	3.618	(0.894)	76	0.022

Table 1 (continued)

	Analytic sample (N=145)			Control group (N=69)			Treatment group (N=76)			p
	M	SD	N	M	SD	N	M	SD	N	
Avoiding distraction	3.993	(1.140)	145	4.087	(1.067)	69	3.908	(1.202)	76	0.346
Ignoring distraction	3.117	(1.064)	145	3.116	(1.051)	69	3.118	(1.083)	76	0.989
Help seeking	3.952	(1.114)	145	4.058	(1.069)	69	3.855	(1.151)	76	0.276
Giving up b/c lazy or bored (reverse coded)	3.393	(0.952)	145	3.275	(1.013)	69	3.500	(0.887)	76	0.157
Working hard to do well	3.952	(0.793)	145	4.043	(0.775)	69	3.868	(0.806)	76	0.186
Giving up b/c of difficulty (reverse coded)	3.876	(0.832)	145	3.957	(0.812)	69	3.803	(0.849)	76	0.268
Keeping working when not interested	3.848	(0.853)	145	3.884	(0.777)	69	3.816	(0.920)	76	0.632
Keeping record of deadlines	3.903	(1.069)	145	3.913	(0.981)	69	3.895	(1.150)	76	0.918
Planning work in advance	3.931	(1.039)	145	3.957	(0.992)	69	3.908	(1.085)	76	0.779
Expected time commitment	11.000	(4.805)	145	10.493	(5.240)	69	11.461	(4.356)	76	0.227

Expected time commitment is measured in hours. Self-regulation and self-management variables from the pre-course survey are measured on a scale from 1 (strongly disagree) to 5 (strongly agree). The last three variables are those that we used to form sub-groups to test heterogeneous effects (presented in Table 4). The standard deviation is provided for non-binary variables. The p-value in the last column tests the difference of treatment and control group means of each variable

randomization. None of the variables exhibited any significant difference between the two groups, suggesting that randomization produced treatment and control groups that were equal in expectation.

Time Management and Self-regulation Skills

In the pre-course survey, questions adapted from a widely used and validated measure of students' self-reported self-regulation and self-management skills (Pintrich 1991) were used to measure students' perceptions of their own self-regulation and time management abilities. Except for the time commitment question, which was measured in hours, students responded to the questions on a Likert scale from 1 (i.e., strongly disagree) to 5 (i.e., strongly agree).

It is important to note that these measures represent students' *self-reported* time management skills. Students' perceptions of their own time management abilities might be weakly, or even negatively, correlated to their actual abilities (i.e. Dunning–Kruger effects, Kruger and Dunning 1999). In our data, correlational evidence suggests that these self-reports are generally unrelated to behavioral measures of time-management skills and prior academic performance, though we do have some evidence that students who reported very poor planning and scheduling did actually procrastinate more.

We use self-reported self-regulation skill measures as an additional test of balance across the treatment and control groups. Table 1 presents summary statistics of measures from the pre-course survey for the analytic sample and treatment and control groups. We observe that, on average, students report fairly high levels of self-regulation abilities. All variables have means above three (closer to *agree* than *disagree*), with most average scores close to four. Students expect to spend about 11 hours on the course per week. We test the balance on each of these survey responses between the treatment and control group using *t* tests and report the *p* value of that test in the last column of Table 1. We observe only one variable having a significant *p*-value at conventional levels (students in the control group are more confident than students in the treatment group about their ability to succeed in online classes), which provides additional confidence that randomization produced equivalent groups.

We use these self-reported self-regulation measures in two other ways: we include them as controls in our treatment effect regressions to increase precision as they predict course performance, and we use the three time-related pre-course survey measures to examine heterogeneous effects of the intervention. We divide the students into meaningful categories that we theorize might be differentially affected by scheduling structure. The three measures that we use to divide students in our heterogeneity checks are the last three variables in Table 1 (expected time commitment for this course, propensity to keep a record of deadlines, and propensity to plan work in advance). Given that these measures are all self-reported, our subgroup analyses examine differences in estimated treatment effects by students' perceptions, and not by actual time management abilities.

Course Performance

We observe achievement outcomes (weekly quiz scores, daily homework scores, and final course grade) as well as video clickstream data from the course management platform. We focus our analyses on the weekly quiz scores and final course grade as opposed to the daily homework scores for two reasons. First, the weekly quizzes are most closely aligned

Table 2 Descriptive statistics for weekly quiz and final course grade outcomes

	Analytic sample (N = 145)		Control group (N = 69)		Treatment group (N = 76)	
	M	SD	M	SD	M	SD
Week 1 quiz (max = 15)	9.910	(3.595)	9.326	(3.763)	10.441	(3.373)
Week 2 quiz (max = 6)	3.945	(1.423)	3.855	(1.375)	4.026	(1.469)
Week 3 quiz (max = 6)	4.186	(1.434)	4.130	(1.371)	4.237	(1.495)
Week 4 quiz (max = 6)	3.979	(1.496)	3.930	(1.384)	4.024	(1.599)
Week 5 quiz (max = 6)	4.890	(1.507)	5.072	(1.304)	4.724	(1.662)
Course grade (max = 100)	81.525	(12.553)	80.454	(13.991)	82.498	(11.094)

Quiz and course grades were collected from the course platform at the end of the course

with the content presented in the lectures. Second, our intervention aimed to affect weekly, not daily, scheduling, so any effects should be most present in weekly assignments (Koch and Nafziger 2017). We provide a check of treatment effects on daily homework scores in Appendix Table 8; we find weaker effects than we see for the weekly quizzes, but a similar pattern among student subgroups.

Table 2 presents summary statistics describing students' scores on their weekly quizzes and their overall course grade. Except for the week one quiz, which had a maximum score of 15, all of the other quizzes had a maximum score of 6. On average, students scored 9.9 on the first quiz and 3.9–4.8 on the following quizzes. The overall course grade, measured on a 100-point scale, was based on course participation (measured by watching the lecture videos and completing the in-video quizzes), five weekly quizzes, performance on the daily problem sets, and final exam score. On average, students scored 81.5 out of 100 available points at the end of the course.

Course Engagement

Clickstream data collected from the course management platform allow us to assess student engagement and time use by observing when students clicked on the lecture videos in the course platform (but not the total time they spent engaged with any particular component of the course). We use these data to examine the mechanisms by which the treatment could have an effect, specifically by investigating the time at which students watched lectures and students' procrastination and cramming behavior.

Experimental Design

Students were randomly assigned into treatment (N = 79) and control (N = 78) groups on the first day of the course. The 76 students from the treatment group and 69 students from the control group who completed the pre-course survey were used in our analyses to estimate the treatment effect. On the first day of week one and week two, the treatment students received an email with a link to an online scheduling survey (which was separate from the pre-course survey which all students received) from the course instructor asking them to schedule on which day and at what time of the day they would watch each of the

five video lectures for that week. The email and survey that were shown to students in the treatment group are presented in Fig. 1a.

The control students received emails from the course instructor asking them to respond to an online survey about how they watched course lectures (in week one students were asked which web browser they used to access the course and in week two they were asked if they listened to the lecture using the computer's speakers or using headphones—the email and survey from week one are presented in Fig. 1b).

It is possible these control emails could affect student outcomes in two ways.⁵ First, the questions could have induced students to think about their online coursework, and thus could have increased their likelihood of watching lecture videos. Second, there is some evidence that increased contact from instructors is positively related to students' academic satisfaction and self-reported learning gains (Bjorklund et al. 2004; Kang and Im 2013; Heiman 2008). Therefore, the treatment contrast is not as strong as it would be if the control students received no contact from the course instructor. Notably, this reduces the treatment contrast and likely provides conservative estimates of the treatment effect relative to an implementation without our control condition.

We provided control students with an emailed survey for two reasons. First, it enabled us to provide extra credit for the control students as we did for the treatment students. Not only was this equitable, but it ensured pursuing extra credit was not driving any differences in academic performance we observed. Second, it ensured that students received an equal number of contacts from the instructor. This allows us to isolate the effect of the scheduling intervention instead of estimating a combined effect of scheduling and additional contact from the course instructor.

The take-up rate of the intervention was very high. Among the 76 students in the treatment group, 71 of them (93.4%) completed the week one scheduling survey and 70 of them (92.1%) completed the week two scheduling survey. Combining the 2 weeks, 74 (97.4%) students in the treatment group completed at least one scheduling survey. Take-up rates among the control students were a bit lower: 84.1% answered the survey the first week, 85.5% answered the survey the second week, and 91.3% completed at least one survey. Although we cannot distinguish between never opening the survey from opening the survey but not responding to the survey questions, the slightly higher take-up rate for the treatment survey might be due to students finding value in the scheduling survey questions.

Methods of Analysis

We employ linear regression to estimate the effect of treatment assignment on several course performance outcomes: standardized measures of quiz performance in weeks one through five and standardized final course grade. We include a vector of student-level covariates, Θ_i , which includes 13 covariates from the pre-course survey.⁶ Equation (1) was used to estimate the intent-to-treat effect (ITT) for student i :

$$Y_i = a + \beta \text{Treatment}_i + \Theta_i \gamma + \varepsilon_i \quad (1)$$

⁵ We note that we could appropriately call our control group a placebo group, although we preference the former term.

⁶ As noted earlier, we did not include student demographic characteristics in our analyses because we were missing these measures for about one quarter of our sample. Appendix Table 7 Columns 7–10 include estimates of the treatment effect from this reduced sample with and without controls. Those estimates are a bit larger, but remain qualitatively similar (same sign and significance).

The ITT estimate is of interest if we conceive of the treatment as being asked to schedule when to watch the lecture video via an online survey in an email. Alternatively, we may be interested in the effect of actually completing the survey and scheduling when to watch each lecture video. This second effect is the treatment on the treated (TOT) estimate, which we estimate using a two stage least squares instrumental variable approach in which random assignment to treatment is the instrument for actually completing the scheduling survey. Because we asked treatment group students to make schedules in both week one and week two, we define the treatment for the TOT estimate in two different ways. For outcomes in the first week, we estimate the effect of whether a student completed the scheduling survey in week one on the week one outcomes. For outcomes after week one, we consider whether a student completed *either* of the surveys in week one or week two to estimate TOT on outcomes of week two to week five as well as the final course grade.

There are three important assumptions to consider in the application of the two stage least squares instrumental variables approach to estimating the TOT effects. The first is a strong first-stage demonstrating treatment assignment is related to treatment received, for which we provided evidence given the high take-up rates of treatment. Second, the instrument must satisfy the exclusion restriction, which states the instrument can only affect the outcome through treatment. Given the random nature of the treatment assignment instrument, this assumption is satisfied. Finally, in order for the instrumental variable estimate to provide a TOT effect (also known as a local average treatment effect or a treatment effect for compliers), we must assume there are no defiers. We think it unlikely that there are students who would have completed the scheduling survey if assigned control and would not have completed it if assigned treatment.

We also estimate treatment effect heterogeneity across three student characteristics measured on the pre-course survey related to time management: how many hours students planned to work on the course, how likely the student said he or she was to keep a record of deadlines, and how likely the student said he or she was to plan work for the course in advance. These three measures directly relate to existing literature on the relationship between time management and course performance. Beattie et al. (2017) identified the numbers of hours spent studying as a critical factor related to academic performance in higher education. We measure students' expected time commitment on our pre-course survey. Admittedly, expected time commitment could vary from actual time spent studying, but we believe our expectations question is a reasonable proxy measure for actual time spent on the course. The two time management measures captured on the pre-course survey, keeping record of deadlines and planning work in advance, are the most direct measures of time management available. They are also the two measures most directly related to our scheduling intervention which aims to induce students to schedule in advance and record their schedule.

We divided students into high and low categories for each of these three measures. Specifically, students who expected to work less than 7 hours per week (the 25th percentile in the class) on this course were categorized into the low expected effort group, while the others were put into the high expected effort group. Students who responded with values of three or lower on their ability to keep a record of deadlines and plan in advance were put into the low self-regulation skills groups for these two items, while those who responded with values of four and five were categorized into the high self-reported self-regulation groups. For each of these two types of self-regulation, the low category represents roughly the bottom third of the class. We estimated the ITT and TOT for each subgroup separately.

Table 3 Treatment effect estimates on course performance

Outcomes	ITT	TOT
Week 1 quiz	0.341* (0.168)	0.363* (0.169)
Week 2 quiz	0.195 (0.173)	0.201 (0.170)
Week 3 quiz	0.100 (0.176)	0.103 (0.172)
Week 4 quiz	0.054 (0.167)	0.056 (0.148)
Week 5 quiz	−0.294 ⁺ (0.172)	−0.303 ⁺ (0.160)
Course grade	0.173 (0.160)	0.191 (0.154)

Each row reports the coefficient on the treatment variable from a regression of that row's outcome (measured in standard deviation units) on a dummy for treatment and thirteen student-level covariates from the pre-course survey. Heteroskedastic robust standard errors are included in parentheses. N = 145

⁺p < 0.10, *p < 0.05, **p < 0.01

Results

Did Students Watch When They Said They Would?

We first examine if students who were assigned to and complied with treatment followed the plans they set. That is, we examine if these students watched the lecture videos at the day and time they scheduled. We define *watching at scheduled time* in two ways: if the time that students watched the video was *within 1 hour* of the time they scheduled and if the watch time was *within 3 hours* of when they scheduled. For example, a student who scheduled to watch the video at 4:00 p.m. on Tuesday would be said to have watched it within one (three) hour if their recorded watch time was between 3:00 and 5:00 p.m. (1:00 and 7:00 p.m.) on Tuesday. In the first week, 32% of treated students watched lecture video one within 1 hour of their scheduled time and 53% within 3 hours. This declined by the end of the week; 6% of students had watched lecture video five within 1 hour of their scheduled time and 11% within 3 hours. Rates were similar in the second week, with 34% of students watching the first lecture within an hour of the scheduled time (49% within 3 hours) and 8% watching the fifth lecture video within an hour of the scheduled time (21% within 3 hours).

Although this appears to be a low rate, many students watched the videos *before* they had planned to do so. We formalize this by relaxing the definition above to include all students who watched each lecture *by 1 hour after* their scheduled time. For example, a student who scheduled to watch the first video at 4:00 p.m. on Tuesday would be counted as having watched it before their planned time if they watched it any time before 5:00 p.m. on Tuesday. In the first week, 79% (36%) of students had watched the first (fifth) lecture video *by 1 hour after their scheduled time*. In the second week, the rates were 56% (30%). On the whole, most treatment students watched the lecture videos before or near the time that they schedule.

Treatment Effect on Course Performance

We provide results addressing our first research question in Table 3, in which we report the treatment effect on two measures of course performance: weekly quiz score and final

course grade for our analytic sample of pre-course survey completers who were randomly assigned treatment and control. The estimates in column 1 are ITT estimates, and the estimates in column 2 are TOT estimates. For the TOT estimates, we consider completing the week one survey as the treatment for week one outcomes and completing either survey as the treatment for all other outcomes.

Relative to control students, students assigned to the treatment group scored 0.341 standard deviations (a little more than one point out of fifteen) higher on the week one quiz than students assigned to the control group, and this is statistically significant at the 5% level.⁷ Given that not all students assigned to the treatment group completed the scheduling survey, the estimated TOT is slightly larger. This finding demonstrates that suggesting students schedule their work in advance in an online course improves their initial course performance.

The magnitude of this effect is also notable; a simple scheduling survey induced more than a third of a standard deviation increase in first week course performance, which is a large effect relative to many low-touch educational interventions. Given that the first quiz accounted for 6% of the total grade, and that the standard deviation on this quiz was about 3.6 points, this amounts to about 0.5 point increase on the final course grade. While this is clearly not a huge increase, half a percentage point could meaningfully affect course grades for students who are on the margin between two grades (indeed, 9% of students in the control group would have received a higher letter grade had they received an additional 0.5 points on the course grade).⁸

In subsequent weeks, however, the difference in quiz scores between treatment and control groups attenuates and becomes insignificant. In week two (the second and final week that students received the scheduling survey), the treatment effect is still positive, but the smaller effect size is not statistically significant. The effect attenuates further in weeks three and four, likely because the scheduling treatment was removed. In the final week of quiz performance, the treatment group actually experiences statistically significant (at the ten percent level) *lower* scores by three tenths of a standard deviation relative to the control group. Although the point estimates are positive, we observe no statistically significant effect on overall course final grade.

Treatment Heterogeneity

We next turn to whether the treatment has the same effect on outcomes across students that vary in their expected time commitment and self-reported time management skills. We conducted treatment heterogeneity checks for three variables: whether students expected to work more than 7 hours on this course, whether students perceived themselves to be good at keeping record of deadlines, and whether students perceived themselves to be good at planning in advance. Table 4 presents estimates of ITT and TOT for low and high levels of each of those three characteristics.

⁷ In Appendix Table 7, we present results from all students who were randomized into the treatment and control groups, and results from our pre-course survey sample but without the 13 controls. Results from all models are similar in magnitude and significance to the main results we present here.

⁸ Half a percentage point on the course grade is the same amount of extra credit that we offered to students for answering the weekly surveys (both treatment and control students). While many things, including wanting to respond to the professor, liking filling out surveys, etc. could induce student to fill out the survey, the very high take up rates indicate that an increase of 0.5 points is likely meaningful to students.

Table 4 Heterogeneity of treatment effects by expectations and time management skills

Dependent variable	Expected working hour		Keeping record of deadlines		Planning in advance	
	Low	High	Low	High	Low	High
Week 1 quiz						
ITT	0.437 (0.417)	0.338 ⁺ (0.192)	0.850* (0.382)	0.148 (0.202)	0.612 ⁺ (0.353)	0.348 ⁺ (0.201)
TOT	0.520 (0.350)	0.356 ⁺ (0.183)	1.130** (0.434)	0.149 (0.186)	0.692* (0.340)	0.361 ⁺ (0.186)
Week 2 quiz						
ITT	-0.272 (0.400)	0.441* (0.202)	0.124 (0.415)	0.177 (0.194)	0.875* (0.377)	-0.004 (0.187)
TOT	-0.272 (0.316)	0.454* (0.194)	0.151 (0.419)	0.177 (0.180)	0.927** (0.333)	-0.004 (0.177)
Week 3 quiz						
ITT	-0.453 (0.503)	0.133 (0.197)	-0.377 (0.370)	0.196 (0.219)	0.016 (0.355)	0.191 (0.216)
TOT	-0.453 (0.397)	0.137 (0.188)	-0.458 (0.386)	0.196 (0.203)	0.019 (0.270)	0.195 (0.205)
Week 4 quiz						
ITT	0.520 (0.334)	-0.010 (0.207)	0.290 (0.362)	-0.116 (0.206)	0.216 (0.337)	0.066 (0.199)
TOT	0.520* (0.263)	-0.011 (0.198)	0.352 (0.368)	-0.116 (0.191)	0.019 (0.270)	0.068 (0.189)
Week 5 Quiz						
ITT	-0.846 ⁺ (0.418)	-0.143 (0.199)	-0.569 (0.415)	-0.298 (0.202)	-0.483 (0.365)	-0.289 (0.205)
TOT	-0.846* (0.330)	-0.147 (0.191)	-0.691 (0.428)	-0.298 (0.187)	0.019 (0.270)	-0.295 (0.194)
Course grade						
ITT	0.012 (0.289)	0.234 (0.196)	-0.160 (0.248)	0.192 (0.205)	0.414 (0.269)	0.167 (0.196)
TOT	0.012 (0.228)	0.241 (0.186)	-0.195 (0.257)	0.192 (0.191)	0.019 (0.270)	0.171 (0.186)
N	37	108	45	100	46	99

Each cell reports the coefficient on the treatment variable from a regression of that row's outcome (in standard deviation units) on a dummy for treatment and thirteen student-level covariates from the pre-course survey. Heteroskedastic robust standard errors are included in parentheses

⁺p < 0.10, *p < 0.05, **p < 0.01

We first examine treatment effect differences for students who had low and high expected hours of work on the course. For a student who has low expected hours of studying for the course, our scheduling intervention might induce them to spend more time on the course than they otherwise would have. Absent our intervention, students were not told how much time this course will take each week. Our scheduling intervention, which encourages students to think about when they will watch each of five roughly hour-long lectures each week, might change students expected work time. Given evidence on the relationship between time spent studying for a course and course outcomes (e.g. Beattie et al. 2017), we expect our intervention to have a larger positive effect for students who expect to spend little time on the course each week.

For the week one quiz outcome, where we see positive effects in the full sample, we estimate positive point estimates for both the low and high expected working hour groups (ITT=0.437, $p>0.10$ for the low-expectation group, and 0.338, $p<0.10$ for the high expectation group). There is no significant difference between the two groups when we estimate the two together using a pooled regression and include an interaction term for low-expectation with treatment. We do, however, see a statistically and practically significant difference in week two quiz scores in which the high expected working group has a positive and significant treatment effect of (0.441, $p<0.05$) while the low time expectation students had a negative point estimate that is not statistically different from zero (-0.272 , $p>0.10$). We observe negative point estimates for both groups in week five (-0.846 , $p<0.10$ for the low expectation group, -0.143 , $p>0.10$ for the high expectation group), and there is no statistically significant difference between the two groups. Overall, students with high and low expected hours of work for the class responded relatively similarly to the scheduling treatment.

In contrast, we observe differential effects when examining treatment heterogeneity across self-reported time management characteristics. For students who state that they do not keep records of deadlines, the treatment has a very large effect: an increase of 0.850 standard deviations ($p<0.05$) for week one quiz performance, which is statistically different from the small and statistically insignificant treatment effect for students who state that they keep a record of deadlines (0.148, $p>0.10$). Similarly, the positive treatment effect on week one quiz score is larger (though not statistically significantly larger) for students who report that they do not tend to plan in advance than those who report planning in advance (0.612, $p<0.10$ for the low planning group, 0.348, $p<0.10$ for the high planning group). These results align with our hypothesis that being encouraged to make a schedule for video watching would most help students with poor time management skills.

Although we did not observe any treatment effect on the week two quiz score in the full sample, we see a strong positive treatment effect of the scheduling intervention in week two among students who report poor skills at planning in advance (0.875, $p<0.05$). Similar to the overall sample, we do not observe any statistically significant results for the week three and four quizzes, ostensibly because the scheduling treatment did not encourage advance scheduling for those weeks. We find generally negative effects of assignment to treatment for the week five quiz scores, but we do not observe a consistent pattern of differential results across the low and high levels of our time management variables in the final week quiz scores. The negative effects are apparent across all student groups and are generally not or only marginally significant. We also see no statistically significant impacts of the treatment on final course grade for any subgroups.

These heterogeneous effects merit further consideration. As we examine heterogeneous effects by *self-reported* time management skills, it is possible that students who report poor time management and self-regulation skills are not actually worse at these abilities than

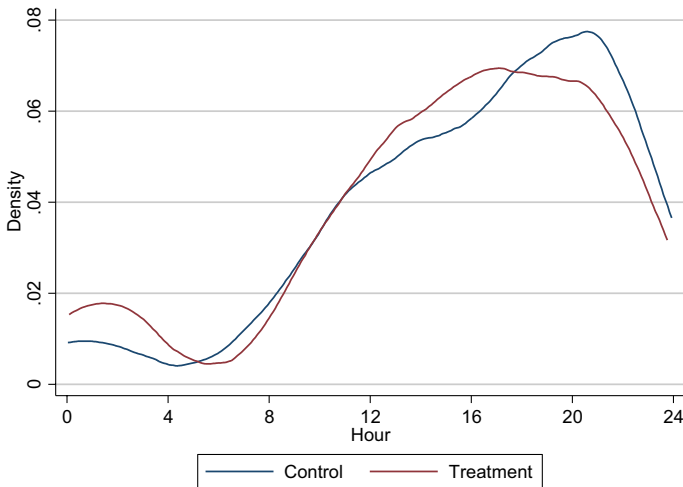


Fig. 2 Distribution of time of watching lecture videos in week 1, treatment and control students

other students. In a set of auxiliary analyses (available upon request) we find that this is mostly true—students who report inferior planning and commitment skills do not, in general, show higher rates of procrastination and cramming. These students are also not more likely to have lower prior GPAs. Thus, it appears as if this type of intervention is most helpful for students who believe that their time management skills are poor, regardless of their actual skills.

Potential Mechanisms

We hypothesize two potential mechanisms to explain our intervention’s effect on achievement that we can explore empirically: reducing the proportion of work done at inopportune times of day and reducing procrastinating work and cramming at the end of the week. We consider these analyses to be exploratory and think that they could provide stepping stones for future work. First, the scheduling encouragement could affect whether students watch lecture videos at times of day that are not conducive to work. Past research has shown that the time of day when students do work is associated with learning outcomes (Goldstein et al. 2007) and that early morning classes are associated with worse outcomes for adolescents (Carrell et al. 2011). Although not directly related to our experiment, we find results consistent with these prior findings in our sample: students who watched lectures in the early morning hours had worse final grades, on average, than students who did work later in the day (results available upon request).

By exploiting the randomization in our experiment, we can explore if the encouragement to schedule altered the time of day in which students watched lecture videos. In our study, students tended to watch lecture videos in the evening and at night—over 50% of lectures were watched between 4:00 p.m. and midnight (and an additional 14% between midnight and 4:00 a.m.). We examine if our treatment was associated with the time that students watched lectures in Fig. 2 and Table 5. Figure 2 provides a kernel density plot, separately for treatment and control students, of the distribution of the timing of lecturing watching throughout the day (0 = midnight, 12 = noon) summed over all

Table 5 Treatment effect estimates on timing of student's lecture watching throughout the day

Proportion of lectures watched	ITT	TOT
Post midnight (midnight 4:00 a.m.)	−0.003 (0.027)	−0.003 (0.029)
Very early morning (4:00–8:00 a.m.)	−0.008 (0.009)	−0.008 (0.009)
Morning (8:00 a.m.–noon)	0.015 (0.027)	0.015 (0.027)
Afternoon (noon–4:00 p.m.)	−0.002 (0.026)	−0.003 (0.026)
Evening (4:00–8:00 p.m.)	−0.032 (0.023)	−0.033 (0.024)
Night (8:00 p.m.–midnight)	0.031 (0.033)	0.032 (0.031)

Heteroskedastic robust standard errors are included in parentheses. The outcome is the proportion of lecture videos that were watched by each student in each of the six 4-h periods during the day. Thirteen covariates from the pre-course survey are included. $N = 145$

⁺ $p < 0.10$, ^{*} $p < 0.05$, ^{**} $p < 0.01$

days in the first week of the course. While the distributions appear to be slightly different, with students in the treatment group watching more videos in the early afternoon and students in the control group watching more videos late at night (after 8:00 p.m.), they are not statistically significantly different from each other (a two-sample Kolmogorov–Smirnov test yields no significant difference).

Table 5 extends this analysis to all 5 weeks of the course. In this table we present regression analyses that test the effect of scheduling on time of day of watching. In the models presented in this table, we examine if being assigned to treatment (ITT) or scheduling lecture watching (TOT) was significantly associated with the proportion of lectures watched in each of the 4 hours periods of time throughout the day. Again, it does not appear that our treatment had any effect on the time of day that students watched the lecture videos; the coefficient on treatment is not significant (and very close to zero) in each model.

We turn to our second proposed mechanism, that treatment could affect whether students wait until the weekly deadline to complete coursework (procrastination) or if they watch all of the lecture videos together in a short period of time (cramming). Using data from the course platform that records when students watched lecture videos (descriptive statistics are provided in Appendix Table 9), we can empirically investigate whether the scheduling treatment altered procrastination and cramming behaviors.⁹ Existing literature has shown that the spacing and timing of work can affect course outcomes (Elvers et al. 2003; Hartwig & Dunlosky 2012; Michinov et al. 2011), and we also find that this is true for the students in our course. Appendix Table 10 presents the relationships between procrastination and cramming and course outcomes. We clearly observe that higher levels of cramming and higher levels of procrastination predict lower final course grades.

To evaluate the possibility of the scheduling intervention affecting cramming and procrastination as a mechanism for the treatment's effect on academic outcomes, we investigate the treatment effect using cramming and procrastination as the outcome variables. Results are reported in Table 6. We neither observe any statistically significant coefficients nor any consistent pattern of results. This suggests that our low-cost and scalable intervention does not have measurable direct effect on altering the cramming

⁹ We describe how we define and operationalize procrastination and cramming in depth in Appendix 2.

Table 6 Treatment effect on cramming and procrastination

	Cramming			Procrastination		
	ITT	TOT	N	ITT	TOT	N
Week 1	-0.142 (0.215)	-0.151 (0.217)	139	0.123 (0.29)	0.131 (0.283)	141
Week 2	0.129 (0.147)	0.140 (0.165)	140	-0.030 (0.266)	-0.033 (0.268)	141
Week 3	-0.093 (0.214)	-0.096 (0.210)	138	0.141 (0.324)	0.145 (0.310)	138
Week 4	-0.059 (0.133)	-0.060 (0.130)	138	-0.158 (0.383)	-0.163 (0.334)	139
Week 5	-0.071 (0.116)	-0.073 (0.114)	137	0.124 (0.258)	0.128 (0.238)	139
Overall grade	-0.080 (0.089)	-0.082 (0.088)	132	0.052 (0.244)	0.054 (0.230)	135

Heteroskedastic robust standard errors are included in parentheses. Spacing is defined as the standard deviation of the first time (in days) students watched each of the five videos for the week. Procrastination is defined as the average time (in days) between watching each lecture and the Friday midnight deadline across all five weekly videos, where negative numbers indicate an earlier time point. Sample size varied for each week's cramming and procrastination because students needed to watch at least one video to get a quantitative measure for procrastination and at least two videos to get a quantitative measure for cramming. Thirteen covariates from the pre-course survey are included

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

and procrastination of video watching behavior, although we acknowledge that our sample size limits the ability to detect small effects.

Discussion

This study finds that encouraging students in an online class to schedule when to watch lecture videos improves achievement early in the course by a third of a standard deviation. The positive effects are concentrated among students who have self-reported poor time management skills. That advantage fades in subsequent weeks of the course.

The week two results are not as consistent as the week one results as they appear for only students with high expected working hours and for students who state that they do not plan their work in advance. We examine six possible explanations for the fading treatment effects between weeks one and two. We do not find evidence for three of these explanations. We do not see differential take-up of the scheduling survey between weeks one and two, we do not find evidence that students scheduled to watch the lectures earlier in the first week than the second, and we do not find that treatment students' time management habits decreased more than the control students' habits between weeks one and two. We discuss the three explanations for which we do have some evidence below.

First, it could be an artifact of the difference in quiz scoring between weeks one and two. The quiz in week one was out of 15 points and had a wider distribution of scores, while the quiz in week two was out of six points, and scores were much more tightly clustered.

Second, we see some evidence that the treatment was less effective in the second week. Even though treatment students planned to watch videos in the second week an average of 17 hours earlier than they did in the first week, they followed their schedules with less

fidelity in the second week; on average, students watched about half a video more as scheduled (within one hour) in week 1 compared to week 2. This suggests that perhaps students' habits are most malleable in the first week of the term, before other engagements and time commitments are firmly set, and a scheduling intervention is most effective early in the term.

Another explanation for the observed difference between week one and week two effects is that the control students took longer to establish a good schedule for studying but eventually caught up (that is, the observed shrinking difference between treatment and control students is due to control students catching up rather than treatment students doing worse). Although we are unable to fully test this possibility, we find weak evidence that this might be part of the story. Treatment students' quiz scores decreased, on average, between week one and week two while control students' quiz scores increased between the 2 weeks. Neither change is statistically significant, but the two trends do support the hypothesis that the treatment effect fades for the treated students because the control students learn how to succeed in the course on their own.

We observe no statistically significant difference between the control and treatment groups in weeks three and four (on average and in any subgroup) and negative impacts in week five. One potential explanation for the attenuation of effects over the length of the course is that we removed the scheduling intervention after the first 2 weeks—instead of teaching students time management skills that persist, the scheduling intervention might have induced students to schedule and improve academic performance only in the weeks when they were directed to schedule. The decline in treatment effect after the intervention was removed could suggest that the encouragement to schedule, while initially effective, did not induce students to internalize a change in time management behavior. It is also possible that students began to rely on the intervention and that its removal harmed their long term performance resulting in negative effects in the final week. However, we do not have robust support for this hypothesis, as we might expect that harm to appear immediately in week three as opposed to week five. The results suggest that future research should seek to implement and test interventions that teach time management tools that will persist over time and in other contexts.

Collectively, these findings deepen our understanding of results from previous studies. Our experimental findings confirm a long line of observational research on the importance of time management in academic performance. Additionally, they extend that confirmation into an important and growing context: online, for-credit postsecondary courses. Prior studies, such as Patterson (2014), who found positive effects using a more invasive intervention, have only been conducted in open-access, free MOOCs. Using a comparable treatment to Baker et al. (2016), who also studied a scheduling intervention in an open-access, free MOOC, we similarly find weak evidence of negative effects on distal outcomes, but, in contrast to that prior work, we observe positive effects on immediate achievement outcomes. The difference in these findings emphasizes the value of studying similar interventions in different contexts.

Students taking a credit-bearing class, as opposed to enrolling in an open-access MOOC that does not confer credit, are likely to be more motivated to successfully complete a course, as they have paid tuition to take the course and are typically enrolled in degree seeking program. We see evidence of this increased motivation in the high take-up of the

scheduling intervention among the treatment group (over 90%) relative to the take-up observed in Baker et al.'s prior MOOC study (about 13%).

While the course in this study is very similar to the modal online course offered at this university (between 2015 and 2017, about 50% of the online classes at this school were in STEM fields, 70% were lower division, and 55% had a pre- or co-requisite), this course is likely somewhat advanced compared to the average online course given calculus is a co-requisite. It is possible the intervention would have stronger effects for less academically accomplished students who may think that they require more time management support.

It is also notable that the focal course in this study was offered over the summer. While many students at this university enroll in courses over the summer, there are socio-economic and academic reasons why a student might enroll in courses over the summer. When examining the group of students who choose to enroll in the summer, as compared to those students who take the same course in the fall, winter or spring terms, we see that the two groups are very similar in most respects. Students who took the class in the summer, however, had stronger prior academic backgrounds (in terms of SAT scores and prior GPA) and were more likely to be male. Our findings suggest replicating analyses among different populations and contexts in online higher education could reveal important heterogeneity and add to our understanding of the mechanisms through which these interventions work.

Although our study reflects data from only one online course and the sample size is small relative to MOOCs, the analysis serves as an important extension to the extant literature. It is exactly because we have a smaller sample size of more motivated students that we can collect detailed data on students' self-reported expectations and time management skills, which enables us to examine how individual characteristics moderate the scheduling treatment. We hope that future analyses of similar interventions targeting time management skills can widen the sample to include additional types of online learning programs in higher education.

The data that result from our sample of more motivated students also allows us to examine some potential mechanisms that could explain our results, but still leaves some hypothesized mechanisms that we are unable to explore. We did not find any evidence to suggest that the treatment affected the time of day at which students completed coursework or their propensity to procrastinate or cram. There are two additional mechanisms through which the treatment could affect student outcomes that we hypothesize but that we do not have the empirical data to test. The first is that the scheduling prompt reduces student anxiety. If students are anxious about finding time to complete the course, inducing them to consciously schedule may reduce that anxiety and enhance performance. It is plausible that this mechanism would work best for students who are aware of their poor time management abilities thereby explaining the heterogeneous results we observe. We encourage future research to include a measure of anxiety to test this hypothesis.

Second, it is possible the scheduling intervention prompted students to spend more overall time in the course. Although the course platform provides data on when students started their lecture videos, it does not provide data on the amount of time students spent in the course platform, so we cannot test whether differences exist in total time spent on the course across treatment and control groups. This also seems a worthwhile inclusion in future studies.

Our findings provide encouraging news for institutions seeking to increase academic performance in online coursework. By implementing a scheduling intervention, whose cost is close to zero, instructors can induce students to improve academic performance in the initial week of the course. Because the treatment is effective among students with lower time management skills, it would be beneficial to assess those skills at the beginning of the course and target interventions to those students. Although we are concerned that the intervention may lead to poorer performance at the end of the course, this negative effect may be driven by the removal of the encouragement to schedule, and we encourage researchers to test an intervention that lasts throughout the full length of the course. Given the rapid expansion of online courses in postsecondary education and the time management challenges faced by students in these courses, expanding cost effective strategies to mitigate time management concerns is an important endeavor, and our study demonstrates that even small improvements in course design may lead to improved academic performance that may provide a large return on investment.

Acknowledgements The authors are grateful for feedback and advice from the Investigating Virtual Learning Environments and Digital Learning Lab groups at UCI's School of Education, particularly Di Xu, Fernando Rodriguez, and Mark Warschauer; seminar participants at AAFP; and the instructor with whom we partnered to implement this intervention. This work was supported by grant number 1535300 from the National Science Foundation.

Appendix 1

See Tables 7, 8, 9, and 10.

Table 7 Robustness of treatment effect results to sample selection and covariates

Outcomes	All students who were randomized		Pre-course survey completers		Pre-course survey completers		Demographic variable sample		Demographic variable sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	ITT	TOT	ITT	TOT	ITT	TOT	ITT	TOT	ITT	TOT
Week 1 quiz	0.321* (0.158)	0.347* (0.169)	0.302+ (0.161)	0.324+ (0.172)	0.341* (0.168)	0.363* (0.169)	0.414* (0.187)	0.436* (0.197)	0.423* (0.201)	0.442** (0.170)
Week 2 quiz	0.105 (0.160)	0.109 (0.165)	0.117 (0.162)	0.120 (0.164)	0.195 (0.173)	0.201 (0.170)	0.117 (0.162)	0.225 (0.190)	0.256 (0.204)	0.259 (0.195)
Week 3 quiz	0.087 (0.160)	0.090 (0.165)	0.073 (0.165)	0.075 (0.167)	0.100 (0.176)	0.103 (0.172)	0.217 (0.189)	0.220 (0.189)	0.301 (0.200)	0.305+ (0.167)
Week 4 quiz	0.082 (0.160)	0.085 (0.165)	0.060 (0.158)	0.061 (0.160)	0.054 (0.167)	0.056 (0.148)	0.180 (0.183)	0.183 (0.183)	0.084 (0.197)	0.085 (0.163)
Week 5 quiz	-0.273+ (0.159)	-0.284+ (0.165)	-0.221 (0.158)	-0.227 (0.160)	-0.294+ (0.172)	-0.303+ (0.160)	-0.168 (0.187)	-0.171 (0.186)	-0.236 (0.206)	-0.239 (0.168)
Course grade	0.173 (0.160)	0.179 (0.163)	0.143 (0.146)	0.147 (0.150)	0.186 (0.155)	0.191 (0.154)	0.201 (0.165)	0.205 (0.168)	0.079 (0.157)	0.080 (0.141)
Self-regulation covariates					X	X				
Self-regulation and demographic covariates									X	X
Observations	157	157	145	145	145	145	111	111	111	111

Each row reports the coefficient on the treatment variable from a regression of that row's outcome (in standard deviation units) on a dummy for treatment. Models presented in columns 5 and 6 include thirteen student-level covariates from the pre-course survey. The models in columns 9 and 10 include student-level characteristics from the pre-course survey as well as demographic characteristics from the school's institutional data. Heteroskedastic robust standard errors are included in parentheses

+p < 0.10, *p < 0.05, **p < 0.01

Table 8 Intent to treat estimates on homework grades

Dependent variable	Full analytic sample		Expected working hours		Keeping record of deadlines		Planning in advance	
	Low	High	Low	High	Low	High	Low	High
Week 1 homework	0.050 (0.140)	-0.115 (0.184)	0.090 (0.183)	0.054 (0.173)	-0.185 (0.248)	0.054 (0.173)	0.716* (0.326)	-0.064 (0.137)
Week 2 homework	0.240 (0.158)	0.213 (0.290)	0.309 (0.201)	0.243 (0.211)	0.011 (0.235)	0.243 (0.211)	0.465 (0.298)	0.224 (0.178)
Week 3 homework	0.239 (0.155)	0.196 (0.259)	0.267 (0.195)	0.372+ (0.193)	-0.329 (0.324)	0.372+ (0.193)	0.250 (0.345)	0.326+ (0.175)
Week 4 homework	0.003 (0.157)	-0.223 (0.234)	0.107 (0.203)	0.000 (0.191)	-0.271 (0.349)	0.000 (0.191)	-0.095 (0.317)	0.096 (0.187)
Week 5 homework	0.068 (0.155)	-0.145 (0.278)	0.115 (0.197)	0.153 (0.194)	-0.491 (0.323)	0.153 (0.194)	-0.068 (0.334)	0.131 (0.183)
N	145	37	108	100	45	100	46	99

Each cell reports the coefficient on the treatment variable from a regression of that row's outcome on a dummy for treatment and thirteen student-level covariates from the pre-course survey. Heteroskedastic robust standard errors are included in parentheses

+ p < 0.10, *p < 0.05, **p < 0.01

Table 9 Summary statistics for cramming and procrastination

	Analytic sample			Control group			Treatment group		
	M	SD	N	M	SD	N	M	SD	N
Week 1									
Cramming	2.25	(1.22)	139	2.20	(1.07)	67	2.30	(1.35)	72
Procrastination	-3.93	(1.67)	141	-3.92	(1.53)	68	-3.93	(1.79)	73
Week 2									
Cramming	1.58	(0.82)	140	1.69	(0.67)	67	1.49	(0.93)	73
Procrastination	-3.48	(1.54)	141	-3.41	(1.14)	68	-3.54	(1.84)	73
Week 3									
Cramming	1.36	(1.15)	138	1.37	(0.89)	65	1.35	(1.35)	73
Procrastination	-3.12	(1.8)	138	-3.16	(1.74)	65	-3.08	(1.87)	73
Week 4									
Cramming	1.21	(0.75)	138	1.18	(0.71)	67	1.24	(0.79)	71
Procrastination	-3.15	(2.23)	139	-3.13	(2.19)	67	-3.17	(2.29)	72
Week 5									
Cramming	0.97	(0.68)	137	0.95	(0.66)	66	0.99	(0.69)	71
Procrastination	-3.19	(1.46)	139	-3.25	(1.36)	66	-3.13	(1.55)	73
Overall course									
Cramming	1.5	(0.50)	132	1.49	(0.42)	64	1.51	(0.56)	68
Procrastination	-3.41	(1.39)	135	-3.43	(1.21)	64	-3.4	(1.55)	71

Cramming is defined as the standard deviation of the first time (in days) students watched each of the five videos for the week. Thus, smaller numbers indicate more cramming. Procrastination is defined as the average time (in days) between watching each lecture and the Friday midnight deadline across all five weekly videos, where negative numbers indicate an earlier time point. Sample size varied for each week's cramming and procrastination because students needed to watch at least one video to get a quantitative measure for procrastination and at least two videos to get a quantitative measure for cramming. For overall course, cramming and procrastination are averaged across all five weeks, if measures for all five weeks were available

Table 10 Relationship of cramming and procrastination with course performance

Outcomes	Separate models				Combined single model		
	Cramming	N	Procrast.	N	Cramming	Procrast.	N
Week 1 quiz	-0.123 ⁺ (0.071)	139	-0.042 (0.052)	141	-0.116 (0.072)	-0.030 (0.053)	139
Week 2 quiz	0.028 (0.105)	140	-0.042 (0.059)	141	0.024 (0.105)	-0.058 (0.058)	140
Week 3 quiz	0.101 (0.073)	138	-0.12* (0.048)	138	0.061 (0.074)	-0.110* (0.049)	138
Week 4 quiz	-0.175 (0.116)	138	-0.012 (0.040)	139	-0.176 (0.119)	0.002 (0.041)	138
Week 5 quiz	0.011 (0.132)	137	-0.075 (0.058)	139	0.041 (0.132)	-0.100 (0.061)	137
Course grade	-0.250 ⁺ (0.147)	132	-0.219** (0.050)	135	-0.263 ⁺ (0.137)	-0.212** (0.050)	132

Heteroskedastic robust standard errors are included in parentheses. Cramming is defined as the standard deviation of the first time (in days) students watched each of the five videos for the week. Procrastination is defined as the average time (in days) between watching each lecture and the Friday midnight deadline across all five weekly videos, where negative numbers indicate an earlier time point. Sample size varied for each week's cramming and procrastination because students needed to watch at least one video to get a quantitative measure for procrastination and at least two videos to get a quantitative measure for cramming. Weekly measures of cramming and procrastination are used to predict weekly quiz scores while overall cramming and procrastination are used to predict course score. Pre-course survey covariates are included

⁺p < 0.10, *p < 0.05, **p < 0.01

Appendix 2

Measuring Procrastination and Cramming

We examine two potential mechanisms that could explain how the scheduling intervention affects students' performance: procrastination and cramming. We define procrastination as how far the average watch time for the five lecture videos in a given week is from the Friday midnight deadline; a larger negative number indicates more time before the deadline (and thus less procrastination). Cramming is defined as the standard deviation of the watch time for each of the five course videos within a week. A smaller number indicates more cramming. Both of these variables are measured in days. The overall course cramming and procrastinating for each student is the average of each student's average across all five weeks.¹⁰

¹⁰ There are a number of factors that need to be taken into consideration when creating these statistics. Although many students clicked play for the same video multiple times, (the average is 1.29 times across all videos), we only used students' first video-watching record in our analyses. Videos that were watched for the first time after the deadline were still included in the cramming and procrastination calculation. No more than 2% of video watches each week were after the deadline. Some students also do not have a recorded watching time for all videos. This could be because they watched it on a friend's computer or accessed it some other way. All students in the class received credit for watching all of the lectures (meaning they completed the accompanying quiz), so we assume these missing video watching times are noise. For the purposes of calculating procrastination and cramming, we define a video as skipped if a student did not have a record of clicking on the video on the course management platform. Across all students and weeks, about 9% of videos were skipped. For students who skipped at least one video in a week, the cramming and procrastination variables reflects the standard deviation and mean video-watching time for less than five videos. Students who watched one or no videos in a given week do not have cramming or procrastination for that week. Students only had overall course cramming and procrastinating scores if they had corresponding scores for each of the 5 weeks. We acknowledge the possibility of measurement error for these video watching measures but we believe such error is likely to be small given the high percentage of students who we observe watching each lecture.

Appendix Table 9 provides summary statistics of students' video-watching habits in terms of their cramming and procrastination.

We first examined the relationship between course outcomes and our measures of procrastination and cramming to examine if these time management measures predict academic outcomes, as has been found in other studies. These two variables were individually included in a series of linear regression models, with the same student level covariates as used in the main regressions, predicting each week's quiz scores and the final course score. For the weekly quizzes, we used the student's cramming and procrastination estimate for that given week. For the final course score, we used the average cramming and procrastination across all five weeks. We subsequently combined cramming and procrastination into a single model to account for the shared variance between the two predictors.

The left panel of Appendix Table 10 shows that, in the separate model specification, students who procrastinated more tended to have lower quiz and final course scores. This negative trend is consistent across all five weeks, and it is statistically significant for the third week and the overall course score. On average across students and weeks, watching the weekly videos a day closer to the Friday deadline is associated with over a fifth of a standard deviation worse final course score. We believe the effect appears strongest in the final course score because it reflects the sum total of students' online problem set scores, participation in lecture-video quizzes, weekly quiz scores, and final exam score. If students are procrastinating on lecture watching, they are also likely to procrastinate on the completion of assignments further reducing their final course score. Furthermore, students who spaced their lecture-video watching out over the entire week instead of watching multiple videos in quick succession tended to have better first week quiz and final course scores; the coefficients on cramming are negative in most weeks. These results for both cramming and procrastinating hold for the final course score even in the combined model which includes cramming and procrastination simultaneously as independent variables. Appendix Table 10 results suggest students who do not cram or procrastinate have higher scores, on average.

References

- Allen, E., Seaman, J., Poulin, R., & Straut, T. (2016). Online report card: Tracking online education in the United States. Babson Survey Research Group and Quahog Research Group, LLC.
- Angelino, L., Williams, F., & Natvig, D. (2007). Strategies to engage online students and reduce attrition rates. *The Journal of Educators Online*, 4, 1–14.
- Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological Science*, 13, 219–224.
- Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago, IL: University of Chicago Press.
- Ashraf, N., Karlan, D., & Yin, W. (2006). Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines. *Quarterly Journal of Economics*, 121, 635–672.
- Babcock, P., & Marks, M. (2011). The falling time cost of college: Evidence from half a century of time use data. *The Review of Economics and Statistics*, 93, 468–478.
- Baker, R., Evans, B., & Dee, T. (2016). A randomized experiment testing the efficacy of a scheduling nudge in a Massive Open Online Course (MOOC). *AERA Open*, 2, 1–18.
- Beattie, G., Laliberté, J. P., Michaud-Leclerc, C., & Oreopoulos, P. (2017). What sets college thrivers and divers apart? A contrast in study habits, attitudes, and mental health. National Bureau of Economic Research Working Paper No. 23588.
- Bjorklund, S. A., Parente, J. M., & Sathianathan, D. (2004). Effects of faculty interaction and feedback on gains in student skills. *Journal of Engineering Education*, 93(2), 153–160.
- Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education*, 84(6), 740–756.

- Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2014). Interactive learning online at public universities: Evidence from a six-campus randomized trial. *Journal of Policy Analysis & Management*, 33, 94–111.
- Bowers, J., & Kumar, P. (2015). Students' perceptions of teaching and social presence: A comparative analysis of face-to-face and online learning environments. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT)*, 10(1), 27–44.
- Britton, B. K., & Tesser, A. (1991). Effects of time-management practices on college grades. *Journal of Educational Psychology*, 83, 405–410.
- Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies and academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, 27, 1–13.
- Carrell, S. E., Maghakian, T., & West, J. E. (2011). A's from Zzzz's? The causal effect of school start time on the academic achievement of adolescents. *American Economic Journal: Economic Policy*, 3, 62–81.
- Cochran, J. D., Campbell, S. M., Baker, H. M., & Leeds, E. M. (2014). The role of student characteristics in predicting retention in online courses. *Research in Higher Education*, 55, 27–48.
- Deming, D. J., Goldin, C., Katz, L. F., & Yuchtman, N. (2015). Can online learning bend the higher education cost curve? *American Economic Review*, 105, 496–501.
- Elvers, G. C., Polzella, D. J., & Graetz, K. (2003). Procrastination in online courses: Performance and attitudinal differences. *Teaching of Psychology*, 30, 159–162.
- Evans, B., Baker, R., & Dee, T. (2016). Persistence patterns in Massive Open Online Courses (MOOCs). *Journal of Higher Education*, 87, 206–242.
- Figlio, D., Rush, M., & Yin, L. (2013). Is it live or is it internet? Experimental estimates of the effects of online instruction on student learning. *Journal of Labor Economics*, 31, 763–784.
- Frederick, S., Loewenstein, G., & O'donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of economic literature*, 40(2), 351–401.
- Gine, Xavier, Karlan, Dean, & Zinman, Jonathan. (2010). Put your money where your butt is: A commitment contract for smoking cessation. *American Economic Journal: Applied Economics*, 2, 213–225.
- Goldstein, D., Hahn, C. S., Hasher, L., Wiprzycka, U. J., & Zelazo, P. D. (2007). Time of day, intellectual performance, and behavioral problems in morning versus evening type adolescents: Is there a synchrony effect? *Personality and Individual Differences*, 42, 431–440.
- Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters? *Economics of Education Review*, 29, 911–922.
- Guàrdia, L., Maina, M., & Sangrà, A. (2013). MOOC design principles. A pedagogical approach from the learner's perspective. *eLearning Papers*, 33, 1–6.
- Hart, C. (2012). Factors associated with student persistence in an online program of study: A review of the literature. *Journal of Interactive Online Learning*, 11, 19–42.
- Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin & Review*, 19, 126–134.
- Heiman, T. (2008). The effects of e-mail messages in a distance learning university on perceived academic and social support, academic satisfaction, and coping. *Quarterly Review of Distance Education*, 9(3), 237.
- Kang, M., & Im, T. (2013). Factors of learner–instructor interaction which predict perceived learning outcomes in online learning environment. *Journal of Computer Assisted Learning*, 29(3), 292–301.
- Kaur, S., Kremer, M., & Mullainathan, S. (2013). Self-control at work. Duke University working paper.
- Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2016). Recommending self-regulated learning strategies does not improve performance in a MOOC. Learning @ Scale Work in Progress.
- Koch, A. K., & Nafziger, J. (2017). *Motivational goal bracketing: An experiment* (No. 10955). Institute for the Study of Labor (IZA) Discussion Paper.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121.
- Lack, K. A. (2013). *Current status of research on online learning in postsecondary education*. ITHAKA S+R.
- Leeds, E. M., Campbell, S. M., Baker, H., Ali, R., & Brawley, D. (2013). The impact of student retention strategies: An empirical study. *International Journal of Management in Education*, 7, 22–43.
- Macan, T. H., Shahani, C., Dipboye, R. L., & Phillips, A. P. (1990). College students' time management: Correlations with academic performance and stress. *Journal of Educational Psychology*, 82, 760–768.
- Michinov, N., Brunot, S., Le Bohec, O., Juhel, J., & Delaval, M. (2011). Procrastination, participation, and performance in online learning environments. *Computers & Education*, 56, 243–252.

- Miltiadou, M., & Savenye, W. C. (2003). Applying social cognitive constructs of motivation to enhance student success in online distance education. *AACE Journal*, *11*(1), 78–95.
- Misra, R., & McKean, M. (2000). College students' academic stress and its relation to their anxiety, time management, and leisure satisfaction. *American Journal of Health Studies*, *16*, 41–51.
- Moody, J. (2004). Distance education: Why are the attrition rates so high? *Quarterly Review of Distance Education*, *5*, 205–210.
- Mullen, G. E., & Tallent-Runnels, M. K. (2006). Student outcomes and perceptions of instructors' demands and support in online and traditional classrooms. *The Internet and Higher Education*, *9*(4), 257–266.
- National Center for Education Statistics. (NCES). (2015). *Digest of Education Statistics*, Table 311.15.
- Nawrot, I., & Doucet, A. (2014). Building engagement for MOOC students: Introduction support for time management on online learning platforms. In *Proceedings of the 23rd international conference on the World Wide Web* (pp. 1077–1082). New York, NY: ACM.
- Patterson, R. W. (2014). Can behavioral tools improve online student outcomes? Experimental evidence from a Massive Open Online Course. Working paper.
- Perna, L. W., Ruby, A., Boruch, R. F., Wang, N., Scull, J., Ahmad, S., et al. (2014). Moving through MOOCs: Understanding the progression of users in Massive Open Online Courses. *Educational Researcher*, *43*, 421–432.
- Pintrich, P. R. (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).
- Rask, K., & Tiefenthaler, J. (2008). The role of grade selectivity in explaining the gender imbalance in undergraduate economics. *Economics of Education Review*, *27*, 676–687.
- Roper, A. R. (2007). How students develop online learning skills. *Educause Quarterly*, *30*, 62–65.
- Rostaminezhad, M. A., Mozayani, N., Norozi, D., & Iziy, M. (2013). Factors related to e-learner dropout: Case study of IUST elearning center. *Procedia*, *83*, 522–527.
- Rovai, A. P. (2003). In search of higher persistence rates in distance education online programs. *The Internet and Higher Education*, *6*, 1–16.
- Schudde, L., & Scott-Clayton, J. (2016). Pell Grants as performance-based scholarships? An examination of satisfactory academic progress requirements in the nation's largest need-based aid program. *Research in Higher Education*, *57*, 943–967.
- Schwartz, B., & Ward, A. (2004). Doing better but feeling worse: The paradox of choice. In P. A. Linley & S. Joseph (Eds.), *Positive Psychology in Practice* (pp. 86–104). Hoboken, NJ: Wiley.
- Song, L., Singleton, E. S., Hill, J. R., & Koh, M. H. (2004). Improving online learning: Student perceptions of useful and challenging characteristics. *Internet and Higher Education*, *7*, 59–70.
- Stratton, L. S., O'Toole, D. M., & Wetzel, J. N. (2008). A multinomial logit model of college stopout and dropout behavior. *Economics of Education Review*, *27*, 319–331.
- Trueman, M., & Hartley, J. (1996). A comparison between the time-management skills and academic performance of mature and traditional-entry university students. *Higher Education*, *32*, 199–215.
- Tuckman, B. W. (2005). Relations of academic procrastination, rationalizations, and performance in a web course with deadlines. *Psychological Reports*, *96*(3_suppl), 1015–1021.
- Van Den Hurk, M. (2006). The relation between self-regulated strategies and individual study time, prepared participation and achievement in a problem-based curriculum. *Active Learning in Higher Education*, *7*, 155–169.
- What Works Clearinghouse. (2017). Standard handbook version 4.0. Washington, DC: Institute of Education Sciences.
- Xu, D., & Jaggars, S. S. (2013). The impact of online learning on students' course outcomes: Evidence from a large community and technical college system. *Economics of Education Review*, *37*, 46–57.
- Zhan, Z., & Mei, H. (2013). Academic self-concept and social presence in face-to-face and online learning: Perceptions and effects on students' learning achievement and satisfaction across environments. *Computers & Education*, *69*, 131–138.