How a data-driven course planning tool affects college students' GPA: Evidence from two field experiments

Sorathan Chaturapruek, Thomas S. Dee, Ramesh Johari, René F. Kizilcec and Mitchell L. Stevens
Stanford University
Stanford, CA
{sorathan, tdee, rjohari, kizilcec, mitchell.stevens}@stanford.edu

ABSTRACT
College students rely on increasingly data-rich environments when making learning-relevant decisions about the courses they take and their expected time commitments. However, we know little about how their exposure to such data may influence student course choice, effort regulation, and performance. We conducted a large-scale field experiment in which all the undergraduates at a large, selective university were randomized to an encouragement to use a course-planning web application that integrates information from official transcripts from the past fifteen years with detailed end-of-course evaluation surveys. We found that use of the platform lowered students' GPA by 0.28 standard deviations on average. In a subsequent field experiment, we varied access to information about course grades and time commitment on the platform and found that access to grade information in particular lowered students' overall GPA. Our exploratory analysis suggests these effects are not due to changes in the portfolio of courses that students choose, but rather by changes to their behavior within courses.

ACM Classification Keywords
K.3.m. Computers and Education: Miscellaneous.

Author Keywords
Information platform; higher education; dashboard; randomized field experiment; GPA.

INTRODUCTION
Enabling informed decisions is the fundamental incentive for transparency in the face of consequential choices. Yet college students routinely face the task of choosing courses and investing effort in their studies under conditions of informational opacity. Universities are famously anarchic organizations [7], whose myriad intramural units are typically only loosely coordinated with one another [18], creating chronic information problems for students who seek to thoughtfully plan their investments of attention, effort and time. Since schools rarely provide accessible information describing prior cohorts' academic choices and outcomes, current students source what they know from peers and form opinions in light of their beliefs about those peers–an inherently biased sampling procedure [10, 20, 5]. Official records might provide representative summary of selected course outcomes that can be distributed across all students independent of their social network ties. Yet while the publication of performance data may yield benefits of efficiency and equity of information access, it may also produce deleterious consequences, such as myopic choices and attempts to game the system [16].

Only a handful of studies offer empirical insight into how students respond to summaries of official course outcomes. Cornell University began publishing median course grades online in 1998, which yielded a natural experiment for studying the effects of exposure to course information. The findings indicate that students, especially lower-performing ones, were subsequently drawn to leniently graded courses [4, 3]. This move toward grade-driven course selection contributed to a steady rise in average GPA at the school. Grade inflation is a concerning trend for universities if it is the result of reward-driven behavior by students or faculty, rather than a reflection of enhanced student learning [11]. Nevertheless, the extent to which students are influenced by grade information remains contested in the literature. In a recent quasi-experiment, Main and Ost [13] found no evidence that letter grades influence undergraduate students' course-taking behavior or decision to major in economics, only a positive effect on student effort within courses.

Thus the current empirical evidence paints an ambiguous picture of the effect of making course outcome data available on subsequent student performance. The research presented here offers new evidence on the consequences of publishing course outcome data based on a pair of randomized experiments at a large, private research university. Our research utilized a course information platform we created that provides aggregate statistics from official university enrollment, course descriptions, course evaluations, and grade distributions derived from official transcripts (for details, see Section *Carta: A course exploration platform*). The platform is maintained both to serve students and enable research. It launched officially in August of 2016 and is available to all students enrolled at the

L@S 2018, June 26–28, 2018, London, United Kingdom

ISBN 978-1-4503-5886-6/18/06. . . $15.00

DOI: https://doi.org/10.1145/3231644.3231668

university. Although use of the platform is fully voluntary, it has enjoyed widespread adoption by the undergraduate student body. Over 80% of undergraduate students interact with the platform during each academic quarter.

We leverage this use to investigate three research questions about the impact of exposure to course outcome information. First, we ask about impact on student performance:

RQ1. Does exposure to course outcome information influence students' subsequent overall GPA?

Presuming there is some effect of course outcome information on GPA, we are led to consider potential mechanisms. Course outcome information may affect students' grades by altering the *portfolio* of courses they choose. In fact, prior research found anticipated grades to be the most important factor in undergraduate course selection [14]. There are myriad ways in which such effects might manifest: for example, a system with rich course outcome information might enable students to engage in strategic planning over a much larger choice set, encouraging exploration into domains of knowledge that are new to them. If these choices challenge students to venture into courses outside their comfort zones, we might expect that students' course performance will deteriorate with the availability of richer course information. Alternatively, students may forgo "challenging" courses for which they are less well prepared in favor of "safe" courses that historically have conferred high marks. They may receive higher grades by choosing a less demanding portfolio of courses than they would have otherwise, consistent with the findings at Cornell University [4]. Students may also use information about course outcomes to strategically maximize their grades by selecting courses that others report as "easy." Taken together, it is unclear whether and how the revelation of course outcome data affects a student's course portfolio, and ultimately their subsequent course performance. Hence our second research question:

RQ2. Does exposure to course outcome information influence the portfolio of courses students choose?

Next, consider how course outcome information may affect students' grades *given a particular course portfolio*. Exposure to course outcome information may cause students to modulate their effort within individual courses. If a student finds that prior outcome data suggests a particular course is relatively leniently graded or requires a low time commitment, then she may decrease the effort devoted to the course. It is therefore possible for grades to be affected in the absence of any course portfolio changes, due to these adjustments in students' effort allocation: grades may go up if students correct planned misallocations of effort, or they may go down if students become overconfident and exert too little effort in classes. (Nevertheless, in return for lower grades, students may experience less stress and more well-being—a likely desirable trade-off in a high-pressure academic environment.) We therefore pose a third research question about *within-course* effects on grades:

RQ3. Does exposure to course outcome information influence students' GPA within a particular course?

In the first experiment (Study 1), we explicitly encouraged a randomly chosen group of students to use the Carta system while most of them were unaware of its existence at launch. We found that platform usage caused a significant reduction in GPA during the subsequent academic term (RQ1). We did not find evidence of changes in course portfolios (RQ2), but did find evidence of within-course effects (RQ3). In a follow-up experiment (Study 2), we manipulated the ease of access to specific course outcome information: grade information and self-reported time commitment data. We found that access to grade information caused a significant reduction in GPA (RQ1), and again did not find evidence of changes in course portfolios (RQ2) but did find evidence of within-course effects (RQ3). In this study we also found that there are subtle interactions between grade and time commitment information that influence students' subsequent performance.

Our field studies reveal that the public provision of course information has a significant impact on student course performance, and thus raise questions for future research into the effect's causal pathways and best practices for the design of course information systems. We offer an interpretation of our findings through through the lens of an established learning theory [15], and conclude by specifying key normative and pedagogical questions that accompany any systematic display of academic information.

CONTEXT

This section describes the context of the university in which this study is set and the *Carta* platform which serves as the substrate for the inquiries described below.

The University

The study site is a highly selective private research university in the US. It offers courses during three ten-week quarters of each academic year and during a summer quarter. During the period of our investigation, the university enrolled roughly 7,000 undergraduate students, with approximately equal numbers of women and men. Like most selective institutions in the US, this university's curriculum is largely elective. Students are obliged to fulfill certain academic requirements, but they retain discretion over which particular courses they choose to meet those requirements. Students are encouraged to declare majors (one or more) before the beginning of their third year. Official campus literature advocates wide exploration of the curriculum, especially in the first and second years before declaring majors.

The university maintains a fairly elaborate program of academic advising that includes full-time professional advisors serving all students, coupled with department-based faculty advising for those with declared majors, and a program of faculty and staff volunteers who provide consultation to students who have yet to declare majors. The research team made Carta available to students as an additional and fully voluntary

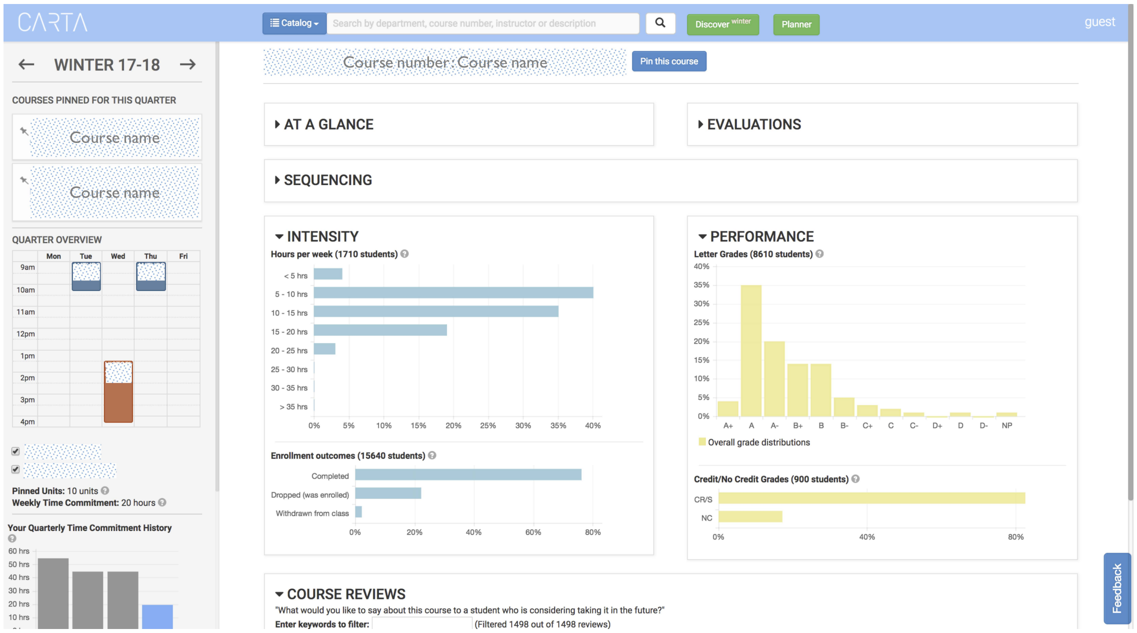

Figure 1. A screenshot of the Carta platform. Examples of information shown are grade and time commitment distributions of prior students

resource. It became one component of a complicated intramural information ecosystem that includes professional advisors, official university documentation about course offerings, and the experiential and reputational information on which college students have long relied to weigh academic options and make decisions. The analysis presented here represents an initial step in an ongoing effort to understand how variation in information supply influences student behavior within that ecosystem.

Carta: A course exploration platform

Carta is a novel web application for course exploration and planning that visualizes data from registrar records and student course evaluations (see Fig. 1). It provides an interface for students to search for classes, as well as to "pin" (or save) classes for further consideration. By making prior students' choices available in aggregated form to current students, Carta expands students' knowledge of course offerings and provides a support tool as students construct their academic schedules. Carta is made available to students under the full knowledge of the university's professional advising staff, some of whom use the platform in their consultations with students.

At the course level, Carta provides students with several types of information: basic descriptive information about courses from the university's course bulletin; formal student evaluations of instructors; sequencing information (i.e., when courses are taken relative to others in students' careers); and textual course reviews from the student course evaluations. Carta provides information on prior students' academic outcomes, grouped under the heading "Performance": the distribution of grades received, and the percentage of students who dropped or withdrew from prior offerings of the same course. Finally, Carta provides information on the number of hours per week students report spending on courses, a measure presented under the heading "Intensity."

The platform has become a popular campus resource. Since its official launch in August 2016, Carta has served approximately 5,500 undergraduates representing over 80% of undergraduate enrollment, with comparable usage rates among students across all academic divisions. Carta is explicitly presented to users as a research project.

STUDY 1: ENCOURAGEMENT DESIGN

This first study investigates whether and how access to information about historical course outcomes affects students' academic performance and behavior. We employed a *randomized encouragement design* because we deemed it both impractical and unethical to compel or prevent use of Carta. Randomized encouragement is a standard tool in econometrics to establish causal effects when a fully randomized controlled trial is not feasible.[1] Our confirmatory analysis tests if the encouragement to use Carta and Carta usage itself has any discernible effect on three primary academic outcome measures: GPA, course drops, and course withdrawals.[2] We test two hypotheses:

H1. Carta usage affects students' GPA at the end of the subsequent quarter.
H2. Carta usage affects students' likelihood of dropping or withdrawing from courses.

[1] It can be viewed as a special case of an instrumental variables (IV) approach to causal inference, where in our case the encouragement to use the platform is the instrument for use.
[2] Students can drop a course until the add/drop deadline (three weeks into the quarter) and it does not appear on their transcript. In contrast, students can withdraw from a course until eight weeks into the quarter but this action is recorded on their transcript.

Rejection of the null hypotheses in favor of H1 or H2 would provide evidence that the platform influences student course performance. After testing for the hypothesized effects, we conduct exploratory analyses to investigate potential mechanisms. Specifically, we attempt to disentangle the extent to which changes in GPA are due to either portfolio effects or within-course effects by examining the degree of diversity and difficulty of classes taken by students.

Method

Participants

Study participants are all undergraduate students at the university who enrolled in classes in Fall 2016-17. Participants include the new cohort of incoming first-year students. The study retains co-terminal Master's students if they remain matriculated in their first four years of study. We exclude any students matriculated at year five or above. We exclude transfer students and those who do not have a GPA in the fall quarter (e.g., those who do not take any letter-grade courses). Accounting for all these features, our total study sample is $N = 6516$.

Procedure

Study participants were randomly assigned to a treatment or control condition with equal probability. We confirmed that randomization was balanced on observable covariates (e.g., students' class year and declared major). Students in the treatment condition were encouraged to use Carta via two mechanisms: an e-mail from the university Registrar (sent on August 1) and a link prominently displayed in students' Fall "to-do list" in the university's official administrative course enrollment system. Both encouragements described the platform as a course exploration tool populated with official historical data from the university. Students in the control condition received neither the Registrar's e-mail nor the link.

Outcome and covariate measures

The primary outcome measure in our confirmatory analysis is the grade point average (GPA) of a student in the courses taken in the Fall quarter. For this measure and all other GPAs, we standardize by subtracting the mean GPA for the population and dividing by the standard deviation (i.e., z-scoring). This outcome measure is labeled FALLGPA. Two additional outcomes in our confirmatory analysis are *drops* of courses, and *withdrawals* from courses.

Predictors in our regression analyses include a binary treatment indicator (ENCOURAGED) and the following covariates:

1. USECARTA: Binary indicator for whether (1) or not (0) the student used Carta before the add/drop deadline at the end of the third week of classes.
2. CLASSYEAR: Categorical variable for student's year (freshman, sophomore, junior, senior).
3. MAJOR: Categorical variable for students' major(s).[3]
4. MAJORCATEGORY: Categorical variable for the school-level categorization of students' major (e.g., engineering, humanities, social sciences).
5. PRIORGPA: Cumulative GPA across all classes taken by a student prior to Fall quarter.[4]

Preregistration

For this study, we followed an internal registration protocol to ensure that we committed to the confirmatory hypotheses above prior to the start of the study, separate from any exploratory analysis conducted post hoc. After conclusion of the first study, we became aware of the Open Science Framework (`https://osf.io`); we subsequently used this platform to preregister our second study (see below).

Results

Descriptive statistics of Carta usage

Table 1 provides descriptive statistics for multiple indicators of Carta usage for students in the treatment and control condition. Indicators include the number of students who sign into Carta, active days on Carta, searching, and pinning. From the descriptive statistics we observe that there were many more students in the treatment condition that used Carta, and that these students also had higher activity levels. This is consistent with the fact that active users in the treatment condition were more likely to have begun using Carta earlier in the treatment period (since they learned of it right at the start). Figure 2 additionally shows the total number of Carta users over time. (For the purposes of our study, we defined USECARTA in terms of usage status at the add/drop deadline.)

			Average number of		
Group	*N*	*N* used Carta	Active days on Carta	Searches	Pins
Freshman Control	865	727	16.0 (14.8)	58 (85)	23.3 (33.7)
Freshman Treatment	870	847	19.6 (13.7)	61 (82)	28.7 (28.9)
Sophomore Control	827	354	7.1 (11.6)	26 (62)	4.5 (10.3)
Sophomore Treatment	860	659	9.9 (11.9)	32 (60)	8.0 (13.7)
Junior Control	802	336	4.9 (8.2)	16 (35)	3.4 (10.1)
Junior Treatment	775	559	7.7 (10.1)	23 (42)	5.9 (11.1)
Senior Control	733	277	3.5 (7.9)	11 (29)	2.3 (7.8)
Senior Treatment	784	554	5.1 (6.5)	14 (25)	5.0 (9.2)
Total Control	3227	1694	8.1 (12.2)	28 (61)	8.8 (21.1)
Total Treatment	3289	2619	10.8 (12.3)	33 (60)	12.2 (20.4)

Table 1. Activity levels of students by the encouragement condition between the official Carta launch on August 1, 2016 and December 16, 2016, the undergraduate housing closing date of the Fall quarter. Students can pin courses to save them on the left-hand panel (see Figure 1). The numbers in parentheses indicate standard deviations.

Confirmatory analysis

We begin the confirmatory analysis by estimating the first-order effect to check if assignment to the encouragement increased the likelihood of Carta usage. Figure 2 plots Carta

[3] Some students had more than one major. To accommodate these cases, we introduced additional values by concatenating the two (or more) major names.

[4] Freshman are missing a PRIORGPA because Carta was introduced before their first quarter. In our analysis, we impute PRIORGPA $= -1$ for Freshmen and use the interaction term PRIORGPA $\times$ CLASSYEAR to account for this imputation.

adoption over time among students in different class years, showing a large increase in adoption among encouraged students. The percentage of Carta users was significantly higher among students in the treatment than in the control condition (80% vs. 52%, $\chi^2 = 534.7, p < 0.001$). This suggests that the encouragement was both effective and notably strong.

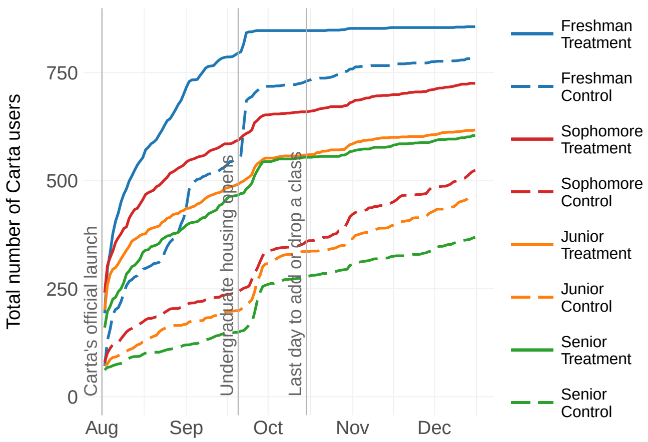

Figure 2. Total number of Carta users over time in 2016 by the encouragement condition and class year.

Next, we estimated the intent-to-treat effect (ITT) of the encouragement on GPA. This analysis focuses on the effect of the encouragement independent of whether students actually used Carta. We find that the encouragement significantly reduced GPA by 0.05 standard deviations ($t = -2.81, p = 0.005$) adjusting for covariates to increase precision (see regression output in Table 2). This provides strong evidence in support of H1. By contrast, we found the ITT effect for course drops and withdrawals not to be significant ($t = 0.25, p = 0.803$). Thus we cannot reject the null hypotheses of no change in drops or withdrawals (H2).

	(1)	(2)
ENCOURAGED	-0.03 (0.02)*	-0.05 (0.02)***
YEAR 3		0.04 (0.03)
YEAR 2		-0.06 (0.05)
YEAR 1		-0.19 (0.14)***
PRIORGPA		0.77 (0.13)***
PRIORGPA $\times$YEAR 3		0.01 (0.06)
PRIORGPA $\times$YEAR 2		0.05 (0.10)
PRIORGPA $\times$YEAR 1		-0.83 (0.08)***
Intercept	0.14 (0.02)***	-0.22 (0.30)

Robust standard errors shown in parentheses; $^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table 2. Regression output for the intention-to-treat analysis estimating the effect of the encouragement on FALLGPA. Model (1) does not adjust for covariates; model (2) adjusts for covariates, and includes fixed effects for MAJOR and MAJORCATEGORY, as well as interactions of those fixed effects with PRIORGPA.

Finally, we estimate the effect of Carta usage itself on GPA using the two-stage least squares (2SLS) estimator. Carta usage is defined by whether a student logs into Carta before the add/drop deadline (USECARTA $= 1$; 0 otherwise). Note that the exclusion restriction[5] appears reasonable in this context, because the encouragement to use Carta (i.e., the email and link in the course enrollment system) is unlikely to influence students' grades except through exposure to information on Carta. Using 2SLS, we find that Carta usage reduced GPA by 0.28 standard deviations ($t = -2.67, p = 0.008$), thus providing strong support for H1. (Again, similar analysis revealed no effect on drops or withdrawals (-0.024 standard deviations, $t = -0.10, p = 0.921$), thus not offering any support for H2.)

Exploratory analysis

We conduct further analyses to explore how Carta usage caused a drop in GPA. We begin by noting that the GPA drop was primarily concentrated among underclassmen. In particular, if we employ the same 2SLS procedure, but subset for underclassmen (freshmen and sophomores) and upperclassmen (juniors and seniors), the GPA drop is significant for underclassmen (-0.47 standard deviations, $t = -2.35$, $p = 0.019$), but not for upperclassmen (-0.16 standard deviations, $t = -1.46$, $p = 0.145$). In other words, the effect of the encouragement was concentrated among students who were relatively "new" to the university. We also test for performance-based heterogeneity in the observed effect of Carta usage on GPA using the same 2SLS procedure for two subgroups.[6] Students with low prior GPA (i.e., below median) showed a significant drop in GPA (-0.35 standard deviations, $t = -2.37$, $p = 0.018$). In contrast, students with high prior GPA (i.e., equal to or above median) had a small and not statistically significant drop in GPA (-0.07 standard deviations, $t = -0.64$, $p = 0.523$).

Next, we investigate two mechanisms described in our RQs above. We first examine if the portfolio of courses selected by students changed as a consequence of Carta usage (i.e. a portfolio effect; *RQ2*). We then examine if students' behavior within the same course changed as a consequence of Carta usage (i.e. a within-course effect; *RQ3*).

Portfolio effect. To check for portfolio effects, we devise multiple metrics to capture the difficulty and familiarity of a given course portfolio (described below). The focus on difficulty and familiarity is motivated by prior research suggesting that students may gravitate towards easier, more familiar courses. Using OLS regression, we estimate the treatment effect on each of these metrics using the same regressors as in the confirmatory analysis (see Table 2) plus corresponding pre-treatment covariates to increase precision. None of the portfolio-level metrics exhibited a statistically significant difference between the experimental conditions (all $p > 0.3$), providing strong evidence that Carta usage caused no substantial changes in students' course portfolios.

The following target metrics served as proxies for course portfolio difficulty:

Prior fraction of A's. The fraction of "A" and "A+" grades given in a course over the past five years averaged over the

[5]The exclusion restriction is a necessary assumption for obtaining credible causal estimates from encouragement designs.

[6]This analysis excludes freshmen for whom prior GPA information is unavailable.

course portfolio of a student. We additionally adjust for the same metric but averaged over prior courses last Fall quarter a student took.[7]

Enrolled course units. The number of course units a student enrolls during Fall quarter. We additionally adjust for the prior number of course units a student took.

Hours per week. The average number of hours per week students report (on course evaluations) spending on a course over the past five years averaged over the course portfolio of a student. We additionally adjust for the same metric but averaged over prior courses a student took.

Class year. The average class year of students in the course (coding freshman = 1, sophomore = 2, junior = 3, senior = 4, others = 5) averaged over the course portfolio of a student. We additionally adjust for the same measure but averaged over prior courses a student took.

Class level. The average class level[8] of a student's course portfolio. We additionally adjust for the same measure but averaged over prior courses a student took.

The following target metrics served as proxies for the degree to which students selected unfamiliar or "exploratory" course portfolios:

Number of departments. The number of distinct departments in which a student enrolled in courses in Fall quarter. We additionally adjust for the same measure but for prior course portfolios a student selected.

Major category. The fraction of courses a student took in each of the major categories (i.e., humanities and arts, social sciences, natural sciences, interdisciplinary, engineering) in Fall quarter. We additionally adjust for the same measure but for prior course portfolios a student selected.

General education requirements. The the number of courses a student took in Fall quarter that satisfy at least one general education requirement (GER).

Within-course effect. To evaluate if students' behavior within the same course changed as a consequence of Carta usage, we check if student performance in the same course differed between the control and treatment conditions. As there was no evidence for portfolio effects on eight different metrics, we assume in the following analyses that any course selection bias is small or negligible. We perform the within-course analysis for freshmen and sophomores, given that underclassmen were most affected by the encouragement. First, we examine treatment effects on GPA among freshmen in the 20 (and 40) most popular courses among freshmen[9] as illustrated in Figure 3 with the treatment t-statistic for each course. In the majority of popular courses, 16 out of 20 (and 24 out of 40) courses, there was a negative effect on GPA, which indicates a within-course effect. Likewise for sophomores, there was a negative effect on GPA in 13 out of 20 (and 27 out of 40) most popular courses among sophomores, as illustrated in Figure 3. We also checked if the within-course effect on GPA varied by course difficulty. While there was no evidence for a trend in effects by difficulty (Spearman Corr $= -0.12$, $S = 23265$, $p = 0.418$), the treatment t-statistic was between 0 and -1 for the vast majority of courses. Overall, these exploratory results strongly suggest that performance within the same course was affected for students in the encouraged group, especially freshmen and sophomores.

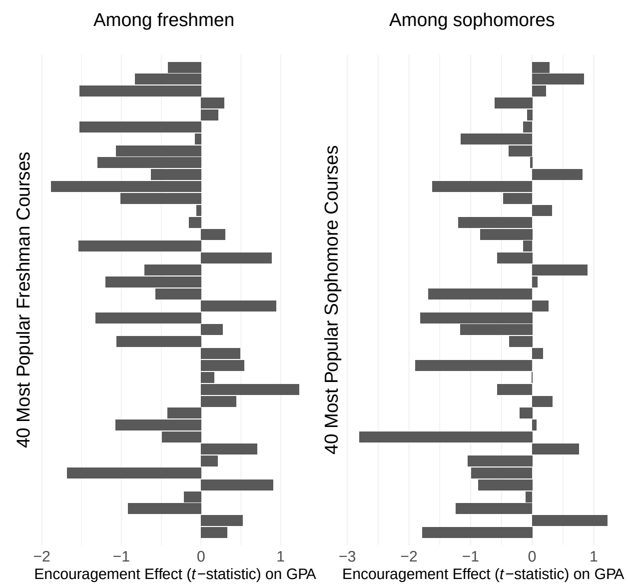

Figure 3. Effect of encouragement (t-statistic) on GPA for the top 40 courses for freshmen (left) and the sophomores (right). The 40 courses are ordered from the most popular at the top to the 40th most popular course at the bottom.

STUDY 2: FRICTION ON INTERFACE ELEMENTS

Study 1 revealed that usage of the Carta platform causes a drop in students' subsequent GPA. However, it does not elucidate the mechanism by which access to course outcome information led to a drop in GPA, as the Carta platform comprises a number of information displays. We therefore conducted a follow-up experiment to investigate how particular course information provided through Carta affects student behavior and performance.

While some information available in Carta, for example course descriptions, is also readily available through other sources, other information is uniquely accessible in the platform. We focus on two kinds of information uniquely

[7]Analogous to PRIORGPA, the prior fraction of A's metric over prior courses for Freshmen is imputed as -1 and accounted for by the interaction term. We use the same methodology across all metrics in this section.

[8]At this university, every course is identified by a department and a three digit numeric code, that is either below 100-level, 100-level, 200-level, 300-level, or above; these encodings broadly indicate the intended audience of the course, with below 100-level and 100-level as undergraduate courses and 200- and 300-level courses or above typically graduate courses.

[9]Popularity was defined based on overall course enrollment among the freshmen.

accessible in Carta that are particularly relevant for students' motivation, ability to self-regulate, and ultimately course outcomes: the *Performance card* containing information on the grade distribution of a course, and the *Intensity card* containing information on the number of hours per week spent on the course that students self-reported on end-of-course evaluations. In our second experiment, we introduce "friction" for students to be able to access the Performance and Intensity cards in Carta. Our goal is to understand the influence that access to information in these panels has on students' subsequent GPA. We therefore registered the following confirmatory hypotheses for this study, consistent with RQ1 as well as H1 in the encouragement design:

H3. Increased friction in accessing prior grade information affects students' GPA at the end of the subsequent quarter.
H4. Increased friction in accessing prior intensity information affects students' GPA at the end of the subsequent quarter.

As in Study 1, we carry out an extensive exploratory analysis after testing these hypotheses. Once again, we attempt to disentangle the extent to which changes in GPA are due to either portfolio effects or within-course effects for either of the two treatments in this study. We are also able to examine the *interaction* effect of providing different combinations of grade and intensity information because we designed the experiment with a 2×2 factorial structure.

Method

Participants

Participants in this study were all undergraduate students at the university who enrolled in classes in Fall 2017-18, applying the same exclusions as in Study 1: for instance, we include incoming first-year students, but exclude students matriculated at year five or above. Additionally, we exclude students who remained unexposed to the experimental manipulation because they never logged into Carta in Fall 2017-2018. The total study sample was $N = 5989$ students.

Procedure

Our study used a 2×2 design, such that participants were randomly assigned to one of the three treatment conditions or a control condition with equal probability. We confirmed that randomization was balanced on observable covariates (e.g., students' class year). The experiment was launched on July 26, 2017, and conducted through the end of the Fall quarter on December 16, 2017.

To enable the experiment we added a feature to the Carta interface to *collapse* and *expand* information cards; see Figure 1. Recall that the Performance card contains information on prior grade outcomes, while the Intensity card contains information on the hours per week students self-reported spending on the course (from end-of-quarter course evaluations). In the control condition, both the Performance card and Intensity card in the Carta interface were collapsed by default on every course page the student visited. We compare the control against three treatment conditions with the following default presentations: (i) Performance card expanded and Intensity card collapsed; (ii) Performance card collapsed and Intensity card expanded; and (iii) both cards expanded. If a card was collapsed by design, a student would need to click anew on every page visit to expand the card. All other cards defaulted to being expanded on page visits.

Throughout our analysis, we refer to an expanded Performance card as *grade visibility*, and an expanded Intensity card as *intensity visibility*. Note that with our definitions, the control condition in this study is a natural analog to the control condition in the encouragement design: in both controls, the default experience is that students have less exposure to prior course outcome information than in the treatment conditions.

Outcome and covariate measures

The measures are identical to those in Study 1, except that our treatment indicators in this study are GRADEVISIBILITY, a binary variable denoting whether grades are visible by default (1) or not (0), and INTENSITYVISIBILITY, a binary variable denoting whether intensity information is visible by default (1) or not (0).

Preregistration

We preregistered our study on OSF (`https://osf.io/shkcn/`). We note that in our preregistration, we registered hypothesis tests with the *binary* outcome measure indicating whether the final Fall quarter GPA of a student is *greater than* versus *less than or equal to* the previous year's median Fall GPA at the university. However, to maintain consistency with our prior study results as well as our exploratory analysis, we report results from linear regression in the main text. Importantly, logistic regression with the preregistered binary outcome variable yields identical significance results to the linear regression results we present here. Results of the logistic regression are reported in a footnote in the confirmatory analysis section.

Results

Descriptive statistics

Analogous to the encouragement design, Table 3 provides descriptive statistics for multiple indicators of Carta usage for students across the treatment condition. Indicators include the number of students who sign into Carta, active days on Carta, searching, and pinning. We observe that all groups were fairly similar on these basic descriptive statistics.

This table also provides information on the rates at which the various cards were expanded by students. In the control condition, students expanded the Performance card more frequently than the Intensity card. The same remains true even if we compare across conditions: students with a collapsed Performance card opened it more frequently than students with a collapsed Intensity card. However, students expanded both cards in substantial numbers in the relevant conditions. This observed behavior suggest that students saw value in the data contained within these cards.

Confirmatory analysis

We use OLS regression with and without covariate-adjustment for observables such as major and class year (same as in Study

Group	*N*	Active days on Carta	Search-es	Pins	Fraction of grade cards opened	Fraction of intensity cards opened
		Average number of				
Control Group	1508	14.0 (12.2)	49.0 (66.2)	11.4 (17.7)	0.47 (0.32)	0.32 (0.27)
Grade Visibility	1495	14.5 (13.6)	52.4 (75.2)	11.7 (20.0)	– –	0.23 (0.25)
Intensity Visibility	1496	13.8 (12.3)	49.2 (65.7)	12.4 (22.0)	0.43 (0.42)	– –
Both Treatments	1490	14.3 (13.1)	51.9 (82.6)	12.3 (27.2)	– –	– –

Table 3. Activity levels of students by the experimental condition between the enrollment opening on August 1, 2017 and December 16, 2017, the undergraduate housing closing date of the Fall quarter. The numbers in parentheses indicate standard deviations.

	(1)	(2)	(3)
GRADEVISIBILITY	−0.02 (0.02)	−0.05 (0.02)**	−0.01 (0.03)
INTENSITYVISIBILITY	0.01 (0.02)	0.03 (0.02)	0.06 (0.03)**
GRADEV×INTENSITYV			−0.07 (0.04)*
YEAR 3		0.09 (0.04)**	0.09 (0.04)**
YEAR 2		0.03 (0.05)	0.03 (0.05)
YEAR 1		0.36 (0.11)	0.36 (0.11)***
PRIORGPA		0.84 (0.07)***	0.84 (0.07)***
PRIORGPA ×YEAR 3		−0.45 (0.08)***	−0.45 (0.08)**
PRIORGPA ×YEAR 2		−0.55 (0.07)***	−0.55 (0.07)***
PRIORGPA ×YEAR 1		−0.82 (0.07)***	−0.82 (0.07)***
Intercept	0.17 (0.02)***	0.33 (0.04)***	0.29 (0.05)***

Robust standard errors shown in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4. Linear regression output for estimating the grade visibility effect and intensity visibility on FALLGPA without adjusting for covariates (1) and with covariate-adjustment and fixed effects for MAJOR (2); added in (3) is the interaction term GRADEVISIBILITY×INTENSITYVISIBILITY.

1). For our confirmatory analysis, we do not include an interaction term between the treatment indicators GRADEVISIBILITY and INTENSITYVISIBILITY, since our hypotheses involve each of these effects on their own (cf. H3 and H4). (In our exploratory analysis, we also study the interaction between these treatment indicators.)

We find that grade visibility caused a significant decrease in GPA ($t = -2.31, p = 0.021$), whereas intensity visibility had no significant effect on GPA ($t = 1.38, p = 0.166$). Thus we find evidence in support of H3 but not H4.[10] Table 4 presents linear regression outputs for each of these effects in detail.

Exploratory analysis

Portfolio effect. We use the same set of eight metrics as in Study 1 to measure “difficulty” and “familiarity” in a student’s course portfolio. We also use the same regression specification, except that we now have two treatment indicators: GRADEVISIBILITY and INTENSITYVISIBILITY. Just as in the encouragement design, none of the metrics exhibited statistically significant differences between the experimental conditions (all $p > 0.3$) with a handful of weak effects ($p > 0.1$) for the INTENSITYVISIBILITY treatment on enrolled course units, hours per week of enrolled courses, the distinct number of departments of courses, and the fraction of courses satisfying general requirements.

Within-course effect. We carry out a parallel analysis to that of the within-course analysis of the encouragement design. As in Figure 3, we examine the grade visibility treatment t-statistic for each course in the most popular 20 (and 40) courses among each class year. In the majority of popular freshman courses, 13 out of 20 (and 28 out of 40), there was a negative effect on GPA among freshmen, which indicates a within-course effect. Likewise for sophomores, there was a negative effect on GPA in 16 out of 20 (and 28 out of 40) courses among sophomores. As for the intensity visibility treatment, analogous within-course analysis of t-statistics among popular courses yields balanced positive and negative values. This result is to be expected given that the intensity visibility treatment did not significantly affect overall GPA.

Treatment interaction. Our 2 × 2 factorial design allows us to study the interaction between grade visibility and intensity visibility. We consider the same regression specification as above, but now include the interaction term GRADEVISIBILITY × INTENSITYVISIBILITY (see Table 4 and Figure 4). Including the interaction term yields several intriguing results. First, relative to the pure control condition (no grade or intensity visibility), adding grade visibility does not affect GPA ($t = -0.35, p = 0.727$), but adding intensity visibility has a *positive* effect on GPA ($t = 2.31, p = 0.021$). However, if both grade and intensity information are made visible, the *negative* interaction term ($t = -2.96, p = 0.003$) effectively cancels out the positive main effect of intensity visibility (see Fig. 4).

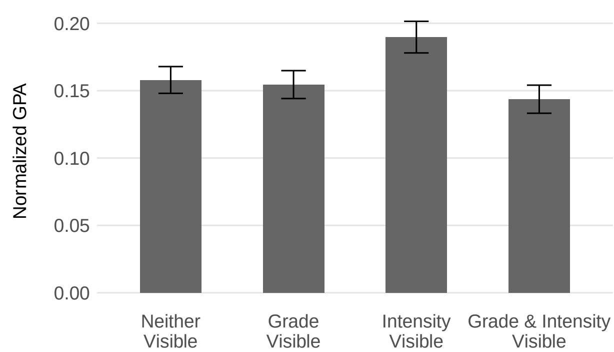

Figure 4. Covariate-adjusted normalized GPA estimates from model (3) in Table 4 by experimental condition. Error bars are ±1 robust standard error. Visible course information was expanded in the interface by default.

Taken together, the effects paint a subtle picture. First, the results suggest that much of the explanation for the overall effect of grade visibility comes from the effect of grade visibility among those students who also experienced intensity visibility by default. Second, the results suggest that intensity

[10] Similarly, using logistic regression controlling for covariates, the grade visibility treatment led to a significant decrease in GPA ($z = -2.09, p = 0.037$), but not the intensity visibility treatment ($z = -0.077, p = 0.938$).

visibility has a positive effect on subsequent GPA, but that this positive effect is countervailed by the additional inclusion of grade visibility. Overall, the effect of the full treatment (both grade and intensity visibility) relative to the full control (neither grade nor intensity visibility) is negative albeit not statistically significant.

DISCUSSION

Students actively seek information about courses and course outcomes as they consider and choose among their academic options. Institutional provision of systematic information describing how prior students fared in similar offerings is a potentially powerful tool for influencing academic choice sets, work expectations, study activity, and grade outcomes. This study was designed as an initial inquiry into this potential.

Our results consistently suggest that exposure to prior course outcome information influences student behavior and subsequent course performance (cf. RQ1). Further, there is no evidence such exposure materially alters the portfolio of courses chosen by students (cf. RQ2), while we do find evidence that the behavior of students within courses is altered (cf. RQ3). Our finding that course outcome information influences course grades without substantially changing course selection adds further support to the conclusion reached by Main and Ost [13], who also did not observe portfolio effects, and stands in contrast to the observed impact on course choice at Cornell [3, 4].

One lens through which we may interpret this result is Pintrich and Zusho's [15] conceptual model of motivation and self-regulated learning in college classrooms. The provision of access to systematic course outcome data can be conceptualized as expanding students' grasp of the classroom context by inducing information about historical grade distributions, time investment, and other factors. This may affect motivational processes such as students' efficacy beliefs about their performance in a course [2]. For example, knowing that a course is leniently graded could raise students' confidence in their ability to earn a high grade. Official course information may also affect self-regulation via meta-cognitive processes involved in goal setting and strategic planning [21, 19]. For example, knowing how much time a course demands can help students allocate their time to achieve personal goals. Together, motivational and self-regulatory processes affect choice, effort, and achievement outcomes, such as course selection, spending more or less time on homework, and ultimately GPA.

While expanding students' grasp of the classroom context by inducing historical course information may have triggered both motivational and self-regulatory processes, it evidently only changed effort regulation and achievement within courses. Results from Study 2 indicate that grade information has a stronger influence on student achievement than time intensity information. This suggests that motivational processes shaping students' sense of self-efficacy [2] with respect to particular courses were dominant in the observed GPA effect; by contrast, time intensity information would be expected to directly support self-regulatory planning processes [21, 19]. To formally disentangle the role of these processes in shaping student behavior within courses, future research might directly measure student perceptions using surveys and assess self-regulatory behavior from behavioral trace data collected by learning management systems.

While our findings are robust and consistent, data limitations moderate the reach of our claims in several ways. First, our study design prevents us from identifying the psychological mechanisms through which information about official course outcomes is translated into changed academic behavior and performance. Course outcome information could influence behavior through a variety of pathways, such as motivational or self-regulatory processes. Our instruments in the present research are not granular enough to distinguish these pathways.

Second, our study observes the effects of access to outcome information in the aggregate. Prior research suggests that there may be heterogeneity in responses to information by gender, race, and social class, because these social identities implicate how students make sense of cues in the academic environment [12, 17]. An essential next step for this research will be to observe whether the magnitude, and perhaps even the direction, of effects of prior information on academic outcomes varies with student demographic characteristics.

Third, we note the peculiarity of our case university, in which most students arrive with very strong academic preparation, enjoy a wealth of academic support services, and almost uniformly graduate in a timely fashion. The effects of tools such as Carta on student behavior may be quite different at schools serving students from a wider range of academic backgrounds or campuses with very different intramural information ecosystems.

Finally, in its current form our inquiry is unique to the U.S. academic world, in which students are expected to explore a wide variety of courses and decide on concentrations of study midway through their undergraduate careers. The interplay between received information about prior and subsequent academic behavior may be different in systems that oblige students to commit to set programs of study upon matriculation.

CONCLUSION

Our two studies offer consistent evidence that knowledge of the distribution of prior students' academic experiences causes changes in behavior that, in the aggregate, produce lower earned grades. This is a powerful insight for educators, because it suggests that the presentation of currently available institutional information can influence students' academic behaviors. Just how educators might use this insight to inform institutional policy is ultimately a normative question. Educators might decide that equity goals compel provision of information about course outcomes to all students because, without it, information access will vary substantially by students' position in peer networks [6] and likely exacerbate information disparities by race and class [1]. At the same time, recognition that access to the grade attainments of prior students can depress subsequent grade attainments might encourage educators to be very thoughtful about how information is displayed on digital platforms and officially defined by campus leaders. Candid discussion of the presentation and interpretation of academic information might itself become a worthy pedagogical project:

an object lesson in the promise and risk inherent in any kind of systematic academic measurement [8, 9].

ACKNOWLEDGMENTS

We thank Michael S. Bernstein, Thomas Black, Sara Cina, Persis S. Drell, Harry J. Elam, Jr., Patricia J. Gumport, Monique H. Harrison, Arik Lifschitz, John C. Mitchell, Andreas Paepcke, Marissa E. Thompson, and Jennifer Widom for helpful feedback and support, and student members of the Carta team for their substantial assistance.

REFERENCES

1. Elizabeth A Armstrong and Laura T Hamilton. 2013. *Paying for the Party*. Harvard University Press.
2. Albert Bandura. 1977. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review* 84, 2 (1977), 191.
3. Talia Bar, Vrinda Kadiyali, and Asaf Zussman. 2008. Quest for knowledge and pursuit of grades: information, course selection, and grade inflation. *Available at SSRN: https://ssrn.com/abstract=1019580 or http://dx.doi.org/10.2139/ssrn.1019580* (2008).
4. Talia Bar, Vrinda Kadiyali, and Asaf Zussman. 2009. Grade information and grade inflation: The Cornell experiment. *The Journal of Economic Perspectives* 23, 3 (2009), 93–108.
5. Elizabeth Bruch and Fred Feinberg. 2017. Decision-making processes in social contexts. *Annual review of sociology* 43 (2017), 207–227.
6. Daniel F Chambliss. 2014. *How college works*. Harvard University Press.
7. Michael D Cohen, James G March, and Johan P Olsen. 1972. A garbage can model of organizational choice. *Administrative science quarterly* (1972), 1–25.
8. Wendy Nelson Espeland and Michael Sauder. 2007. Rankings and reactivity: How public measures recreate social worlds. *American journal of sociology* 113, 1 (2007), 1–40.
9. Wendy Nelson Espeland and Mitchell L Stevens. 2008. A sociology of quantification. *European Journal of Sociology/Archives Européennes de Sociologie* 49, 3 (2008), 401–436.
10. Martin Fishbein. 1979. A theory of reasoned action: some applications and implications. (1979).
11. Valen E Johnson. 2006. *Grade inflation: A crisis in college education*. Springer Science & Business Media.
12. René F Kizilcec, Glenn M Davis, and Geoffrey L Cohen. 2017. Towards Equal Opportunities in MOOCs: Affirmation Reduces Gender & Social-Class Achievement Gaps in China. In *Proceedings of the ACM Conference on Learning at Scale*. ACM, 121–130.
13. Joyce B Main and Ben Ost. 2014. The impact of letter grades on student effort, course selection, and major choice: A regression-discontinuity analysis. *The Journal of Economic Education* 45, 1 (2014), 1–10.
14. Ivana Ognjanovic, Dragan Gasevic, and Shane Dawson. 2016. Using institutional data to predict student course selections in higher education. *The Internet and Higher Education* 29 (2016), 49–62.
15. Paul R Pintrich and Akane Zusho. 2007. Student motivation and self-regulated learning in the college classroom. In *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Springer, 731–810.
16. Peter Smith. 1995. On the unintended consequences of publishing performance data in the public sector. *International Journal of Public Administration* 18, 2-3 (1995), 277–310.
17. Claude M Steele. 1997. A threat in the air: How stereotypes shape intellectual identity and performance. *American psychologist* 52, 6 (1997), 613.
18. Karl E Weick. 1976. Educational organizations as loosely coupled systems. *Administrative science quarterly* (1976), 1–19.
19. Philip H Winne. 1995. Inherent details in self-regulated learning. *Educational Psychologist* 30, 4 (1995), 173–187.
20. Wendy Wood. 2000. Attitude change: Persuasion and social influence. *Annual Review of Psychology* 51, 1 (2000), 539–570.
21. Barry J Zimmerman and Dale H Schunk. 2001. *Self-regulated learning and academic achievement: Theoretical perspectives*. Routledge.