I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13732

Nudging Demand for Academic Support Services: Experimental and Structural Evidence from Higher Education

Todd Pugatch
Nicholas Wilson

SEPTEMBER 2020

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13732

Nudging Demand for Academic Support Services: Experimental and Structural Evidence from Higher Education

Todd Pugatch
Oregon State University and IZA

Nicholas Wilson
Reed College

SEPTEMBER 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Nudging Demand for Academic Support Services: Experimental and Structural Evidence from Higher Education*

More than two of every five students who enroll in college fail to graduate within six years. Prior research has identified ineffective study habits as a major barrier to success. We conducted a randomized controlled advertising experiment designed to increase demand for academic support services among more than 2,100 students at a large U.S. public university. Our results reveal several striking findings. First, the intervention shifted proxies of student attention, such as opening emails and self-reported awareness of service availability. However, the experimental variation indicates that approximately one-third of students are never attentive to student services. Second, advertising increased the use of extra practice problems, but did not affect take-up of tutoring and coaching, the other two services. Structural estimates suggest that transaction costs well in excess of plausible opportunity costs explain the differences in service use. Third, the characteristics of advertising messages matter. Several common nudging techniques—such as text messages, lottery-based economic incentives, and repeated messages—either had no effect or in some cases reduced the effectiveness of messaging.

JEL Classification: A22, D91, I23, M31

Keywords: email, higher education, incentives, nudges, text, randomized control trial

Corresponding author:
Todd Pugatch
School of Public Policy
Oregon State University
303 Ballard Extension Hall
Corvallis, OR 97331
USA
E-mail: todd.pugatch@oregonstate.edu

* This study was registered in the AEA Trial Registry as AEARCTR-0002744 (Pugatch and Wilson 2018a). We thank Jon Chesbro, Maureen Cochran, Marjorie Coffey, Clare Creighton, Allyson Dean, Lindy Foster, Chrysanthemum Hayes, Laura Kawano, Camille Nelson, Mike Nelson, Daniel Newhart, Laura Relyea, Liz Schroeder, Paul Thompson, and Gail Udell for their assistance and comments. Research with human subjects approved by Oregon State University Institutional Review Board, Study #8402. Financial support provided by Oregon State University via an FY19 Individual Learning Innovation Grant. The findings, interpretations, and conclusions expressed in this paper are those of the authors and do not necessarily represent the views of the aforementioned individuals or agencies. All errors are our own.

1 Introduction

More than two of every five students who enroll in college fail to graduate within six years (National Center for Education Statistics 2016). Insufficient study effort can be central to this failure. Students often study far less than recommended not only by faculty, but also their own initial intentions (Oreopoulos and Petronijevic 2019). Distractions such as video games brought by randomly assigned roommates (Stinebrickner and Stinebrickner 2008) or the success of the university's football team (Lindo, Swensen, and Waddell 2012) can reduce study effort and lower grades. Colleges have attempted to close the gap between student intentions and behavior: public 4-year institutions spent 22% of their budgets in 2016-2017 on "academic support, student services, and institutional support" (U.S. Department of Education 2019). Yet many students do not use these services.

Although several behavioral biases may explain insufficient student effort, one unifying explanation is inattention to the future consequences of current study habits (Gabaix 2019). How can universities better focus student attention on changing study habits? Which academic support services do students value most? What messaging methods best capture student attention? When, and how often, should students be reminded to seek help? Do incentives increase take-up of academic services?

To answer these questions, we conducted a field experiment using insights from behavioral economics to test which advertising interventions increase demand for academic support services and improve learning outcomes. The experiment included more than 2,100 students enrolled in introductory economics courses at Oregon State University, the largest public university in the state, during the 2018-2019 academic year. We tested the effects of advertising on use of three different academic support services: tutoring, coaching, and extra practice problems. We varied the medium, frequency, and timing of our advertising, as well as whether it included a lottery-based cash prize offer, allowing for rich variation among treatments.

We present a model of student inattention to the net benefits of academic service use to frame the experiment. We evaluate the interventions using a pre-registered analysis plan (Pugatch and Wilson 2018a). In exploratory analysis not included in our analysis plan, we structurally estimate the parameters of the behavioral model, allowing us to sort students into attention "types" and quantify the perceived costs and benefits of academic service use.

Our results reveal several striking findings. First, the interventions shifted proxies of student attention, such as opening emails and self-reported awareness of service availability. We use experimental variation to characterize attention types, or the proportion of students who are always attentive, never attentive, and compliers. We find that approximately two-thirds of students are always attentive, yielding an estimate of the standard attention parameter *m* (Gabaix 2019) of 0.65. Approximately one-third of students are never attentive to student services. Around 5% or less are compliers, i.e., shift their attention in response to the intervention.

Second, we find increases in service use in response to advertising, driven by extra practice problems, which carried the lowest transaction costs. Although the absolute magnitude of our effect sizes on service use are modest – almost always less than 10 percentage points – the effects often are economically quite meaningful. For example, sending an email advertisement for extra practice problems reduced the proportion of students not using the problems by approximately 50%. By contrast, structural estimates indicate enormous transaction costs for coaching and tutoring, well in excess of plausible opportunity costs. Estimates of the shrouded benefits of academic support services to attentive students are large relative to the common (or base) benefit, yet small relative to the estimated cost, consistent with the modest effect of our messages on service use.

Third, the characteristics of messages matter. Emails were relatively effective, but other common nudging techniques were not. For instance, text advertising was ineffective at increasing take-up for any of the services. Adding a lottery-based economic incentive reduced the effectiveness of messaging. Messaging was particularly effective toward the end of the academic term, but message fatigue occurred for those receiving more than two messages.

These findings contribute to a nascent literature examining nudges to use academic support services.[1] For instance, other studies have analyzed tutoring (Butcher and Visher 2013; ideas42 2015; Paloyo, Rogan, and Siminski 2016; Pugatch and Wilson 2018b; Gordanier, Hauk, and Sankaran 2019), coaching (Bettinger and Baker 2013; Oreopoulos and Petronijevic 2018), goal-setting (Clark et al. 2017; Dobronyi, Oreopoulos, and Petronijevic 2019; Bettinger, Castleman, and Mabel 2019), feedback from faculty (Carrell and Kurlaender 2020), and course-specific practice (Auferoth 2020; Rury and Carrell 2020).

The rich variation in our experiment allows us to build upon this literature in important ways. First, while most studies focus on a single academic support service in isolation,[2] our experiment draws attention to several different services. We can therefore assess the relative demand and perceived transaction costs for each.

Second, we build on the relatively thin literature testing how the characteristics of messages influence human capital investment. Designers of nudges to promote education face important choices about message medium (e.g., email versus text), timing, frequency, and inclusion of incentives, but have little evidence to guide these choices. Only a few previous studies have tested effects on the timing and frequency of messaging (e.g, Pop-Eleches et al. 2011; Cortes et al. 2019a; 2019b). To our knowledge, the only other study testing message medium in this context is Oreopoulos and Petronijevic (2018), which analyzes different methods of coaching students. Studies on financial incentives for academic effort in higher education are also relatively few (Angrist, Lang, and Oreopoulos 2009; Paloyo, Rogan, and Siminski 2016; Pugatch and Wilson 2018b; Barrow, Rouse, and McFarland 2020). More broadly, there are few studies of conditional economic compensation for public service take-up in the United States (e.g., Riccio 2010; Miller et al. 2016; Courtin et al. 2018). We contribute to each of these areas.

Third, we expand the existing experimental evidence on attentiveness (e.g., Hossain and Morgan 2006; Chetty, Looney, and Kroft 2009; DellaVigna and Pollet 2009; Lacetera, Pope, and Sydnor 2012;

[1] See Damgaard and Nielsen (2018) for a review of nudges in education.
[2] There are some notable exceptions (e.g., Oreopoulos and Petronijevic 2019; Balaban and Conway 2020; Page, Lee, and Gehlbach 2020).

Allcott and Wozny 2014; Taubinsky and Rees-Jones 2018; see Gabaix 2019 for a review), by presenting what may be the first explicit evidence from higher education on the attention parameter *m* from this class of behavioral models.

Our experiment also benefits from being easily replicable and immediately relevant for many higher education institutions. The experiment took place among several thousand students at a large public university using interventions (email and text messages) with nearly zero marginal cost. In addition, we focus on existing academic support services. This complements studies of support services designed by researchers which establish proof of concept before scaling-up more broadly.

The remainder of the paper is organized as follows. Section 2 provides a theoretical framework for understanding the decision to use academic support services. Section 3 presents the experimental design. Second 4 describes the statistical methods and data. Section 5 presents our regression results. Section 6 presents our structural estimation. Section 7 provides a discussion. Section 8 concludes.

2 Theoretical Framework

2.1 Student attention and use of academic support services

Consider a student's choice to use an academic support service. When using the service, the student bears an immediate cost but enjoys a future payoff. Let D be an indicator for service use. The student's value function V, expressed in course grade points (e.g., the 0-4 scale common to U.S. universities), is:

$$\begin{aligned} V(D) &= u_1(D) + \delta u_2(D) \\ &= -Dc + \delta(1 + \theta D)\underline{y} \qquad (1) \end{aligned}$$

where u_t is the period-t utility function, with the choice of D occurring in period 1 and associated payoff in period 2; c is the cost of using the service, in terms of the next-best alternative use of time; $\delta < 1$ is the discount factor; θ is the return to the service; and $\underline{y}$ is the baseline course grade in the absence of the service.

The student uses the service if:

$$V(1) \geq V(0)$$

$$-c + \delta(1+\theta)\underline{y} \geq \delta\underline{y}$$

$$\delta\theta\underline{y} \geq c \qquad (2)$$

The student therefore uses the service if the discounted return exceeds the immediate cost. We may define $c^* \equiv \delta\theta\underline{y}$ as the threshold cost below which the student uses the support service.

Now consider a student who may be inattentive to the net benefits of their choice (as in Gabaix 2019). Let *m* be a binary attention parameter, such that *m*=1 represents full attention and *m*=0 represents inattention. An attentive student may perceive the benefits of using a support service differently from an inattentive student, such that:

$$\theta = b + m\tau_b \qquad (3)$$

where b is the benefit of service use commonly known to all students, and τ_b is the shrouded value of service use visible only to attentive students. If $\tau_b > 0$, attentive students perceive academic support services as generating higher returns than inattentive students. If $\tau_b < 0$, the opposite is true: attentive students perceive support services as less beneficial than inattentive students. If $\tau_b = 0$, the model reverts to the standard formulation in which student attention plays no role.[3]

One form of inattention is lack of knowledge about the existence of the academic support service. In our model, a complete lack of knowledge would be equivalent to *b*=0 for a given student, i.e., no commonly known benefit to service use. This case seems implausible in our setting, as all students are informed about the existence of the student services which we examine. Alternatively, a student may forget their initial plan to study (Ericson 2017) or become distracted (Stinebrickner and Stinebrickner 2008; Lindo,

[3] For simplicity, we treat attention *m* as exogenous. We could endogenize attention by letting students choose *m*, with *m*=1 carrying an additional cost. Students would then tradeoff the benefits of attention with its cost.

Swensen, and Waddell 2012). All these cases are consistent with $\tau_b > 0$. All else equal, inattentive students would be less likely to use the service than attentive students.[4]

2.2 Relationship to experimental interventions

We test experimental interventions designed to draw attention to academic support service use. To see how such interventions fit within the framework, let *m* be a function of a binary variable *Z*. Adopting potential outcomes notation:

$$m(Z) = \begin{cases} m_0, & Z = 0 \\ m_1, & Z = 1 \end{cases}$$

$$= (1 - Z)m_0 + Zm_1 \quad (5)$$

In our study, *Z* are randomly assigned advertising messages to use different academic support services. We assume monotonicity, so that $m_1 \geq m_0$. In other words, the experimental interventions cannot make students less attentive. Under this assumption we may define three groups of students:

1. Always attentive: $m_1 = m_0 = 1$
2. Never attentive: $m_1 = m_0 = 0$
3. Compliers: $m_0 = 0, m_1 = 1$

These categories adapt the familiar classifications of Imbens and Angrist (1994) to the model of behavioral inattention described above. The always- and never-attentive students will not change their behavior in response to the experimental interventions. Compliers who receive an intervention will change their objective function from that of an inattentive student ($\theta = b$) to an attentive student ($\theta = b + \tau_b$). In other words, the intervention reveals the shrouded benefit of service use to compliers. Although we cannot observe attention directly, we can proxy for attention (using administrative data on email read receipts and

[4] There are other ways to model the form of inattention, or to model behavioral biases more generally. For instance, we could set $V(D) = -Dc + (m + [1 - m]\beta)\delta(1 + \theta D)\underline{y}$ for β<1, so that the discount factor for inattentive students is $\beta\delta$, as in models of hyperbolic discounting and procrastination (Laibson 1997; O'Donoghue and Rabin 1999). However, the version in equation (3) relates more directly to the experimental treatments in this study. Our encouragement messages focus on the benefits of academic support services, rather than attempting to make students less present-biased.

survey measures of awareness about academic support services, for example) to approximate the function $m(Z)$. Shifts in these outcomes in response to experimental interventions Z represent the share of compliers for these measures of attention.

Our framework considered only a binary Z for simplicity in exposition. Our main empirical results take a reduced-form approach, using the model as motivation but distinguishing among many different characteristics of the intervention. A subsequent section presents results from structurally estimating the model.

3 Research Design

3.1 Context and sampling frame

The study occurred at Oregon State University (OSU). OSU is the largest university in the state, enrolling more than 25,000 students at the flagship Corvallis campus. Success in large introductory courses is highly predictive of degree completion at the university. Economics Principles courses (Introduction to Microeconomics and Introduction to Macroeconomics), which enroll more than 2,500 students per year in sections of 200-250 students each, are of particular concern. The university administration has placed Economics Principles courses on a shortlist of problematic undergraduate courses for two reasons. First, student DFW (grade of D, F, or Withdraw) rates in both courses are among the highest among large-enrollment courses at the university. Second, receiving a DFW in these courses is highly correlated with failure to graduate because at least one of these courses fulfill requirements in 40 undergraduate majors.

Students in Economic Principles may access three types of free academic support services through the university.[5] First, the Academic Success Center (ASC) offers academic coaching to improve student study skills. Coaches are advanced undergraduates trained by professional ASC staff. Students can drop in for coaching or make appointments online. Second, the Economics Tutoring Lab offers drop-in peer

[5] Other than contacting the instructor directly, to our knowledge these are the only free academic support services available to all students in these courses. Other services are either limited to subsets of students (e.g., student-athletes), non-academic (e.g., psychological counseling), or not free (e.g., private tutoring).

tutoring by advanced undergraduate Economics majors. Tutors are paid for their services. Third, Economics Principles instructors provide extra practice problems on their course websites. Academic coaching and tutoring therefore require students to go to specific locations during weekday business hours, while practice problems are available online at any time. Tutoring and practice problems offer course-specific support, while coaching is general.

All three services are free for all students to use. All instructors discuss these services in class and on course websites, particularly at the beginning of the term and around exams. Flyers appear for the Tutoring Lab and Academic Success Center in classrooms and around campus. We lack data on student contact with these messages, but anecdotally we expect their exposure to be widespread.

We conducted the study in 12 (of 13) on-campus sections of Economics Principles in academic year 2018-2019. These 12 sections were taught by three different instructors.[6] Eight sections were introductory microeconomics and four sections were introductory macroeconomics. Each course section met for an academic quarter, 10 weeks.

3.2 Experimental design

Our research design is the randomized controlled trial. We used automated messaging to draw the attention of randomly assigned students to specific actions they could take to improve their performance in Economics Principles courses. All students were eligible for the full range of academic support services and received the current standard for notifications, but subsets of students were encouraged to access particular services. We measured how this advertising influenced service take-up and learning outcomes.

To assign students to groups, we first randomly assigned all study participants to control (25%) or treatment (75%). Within the treatment group, we randomly assigned the characteristics of each treatment: medium, message, incentive, frequency, and timing.

[6] We piloted a separate study in the 13th section. This section was taught by a different instructor, allowing us to simplify data reporting across the separate studies.

• *Medium*: Within the subsample of treated students who provided a mobile phone number, we randomly assigned the message medium as:

1. Email (50%)
2. Text message (50%)

Within the subsample of treated students who did not provide a phone number, we assigned students to the email group.

• *Message*: we assigned an advertising message to each treated student. The messages advertised different academic support services:

1. Academic coaching (33.3%)
2. Economics peer tutoring (33.3%)
3. Extra practice problems (33.3%)

• *Incentive*: we assigned an incentive or no incentive to each treated student:

1. Incentive (50%): The advertising message indicated that the student will be entered in a lottery to receive $250 credit at the campus dining halls and bookstore if they access the support service before a specified date.
2. No incentive (50%)

• *Frequency and timing*: we randomly assigned the frequency and timing of advertising messages to each treated student within the 10 weeks of the term:

1. Week 3 (14.3%)
2. Week 6 (14.3%)
3. Week 9 (14.3%)
4. Weeks 3/6 (14.3%)
5. Weeks 3/9 (14.3%)
6. Weeks 6/9 (14.3%)
7. Weeks 3/6/9 (14.3%)

Figures 1-2 show the emails and texts, respectively. The versions shown include incentives. Versions without incentives simply remove the incentive language, but are otherwise identical. Emails were sent from the accounts of the student's instructor. Texts arrived from a phone number with the same area code as the university. We (the researchers) sent all messages, allowing us to standardize message timing across course sections and conceal student treatment status from instructors.

Note that each treatment characteristic remained constant for each treated student during the term, regardless of message frequency. For instance, if a student was assigned to receive three messages, they were all for the same support service, same incentive treatment, and via the same medium.

Within the subsample of treated students who did not provide a phone number, there were 42 possible combinations of treatment characteristics (3 messages x 2 incentives x 7 frequency/timing combinations). Within the subsample of treated students who provided a phone number, these 42 combinations exist for both media, leaving 84 possible combinations.

The probability that a randomly selected member of each subsample was assigned a particular combination of treatment characteristics is the probability of treatment assignment times the corresponding probabilities of each characteristic (conditional on treatment). For instance, the probability that a student who provided a phone number was assigned to the email/coaching/incentive/Week 9 cell is:

$$Pr(treated)*Pr(email)*Pr(coaching)*Pr(incentive)*Pr(Week\ 9) = .75*.5*.333*.5*.143 = .009$$

The unit of treatment assignment is the individual student.[7] We stratified treatment by course section and class year (freshman/sophomore/other) to ensure balance on these characteristics. The number of treatment combinations did not divide evenly among study participants. The actual sample sizes assigned to each group therefore does not match the *a priori* treatment probabilities exactly. We conducted random

[7] More precisely, the unit is the student-section. Students enrolled in multiple sections within the study can appear in the data multiple times. For simplicity, we refer to students as the unit of observation, unless stated otherwise. We later check sensitivity of results to repeat appearances of the same student.

assignment privately using Stata. Instructors were blinded to student treatment assignment. We report realized treatment assignments in Section 4.7.

4 Statistical Methods and Data

4.1 Effects of Advertising on Academic Support Service Take-Up

We estimate the causal effect of the intent to treat with an advertising message on service take-up using linear OLS regression. We follow an analysis plan submitted prior to analyzing the data (Pugatch and Wilson 2018a).

One specification is:

$$takeup_{ijk} = \alpha + \beta_1 message_{ijk} + \beta_2 incentive_{ijk} + X'_{ij(-k)}\,\Omega + \gamma_j + \varepsilon_{ijk} \qquad (6)$$

where $takeup_{ijk}$ is an indicator variable equal to 1 if student *i* in course section *j* accessed service *k* (e.g., the Economics Tutoring Lab) during the term and 0 otherwise. $message_{ijk}$ is an indicator variable if the student was assigned to receive an advertising message (of any form at any time) for the given service and $incentive_{ijk}$ is an indicator variable if the student was assigned to receive an incentive for the given service in addition to the message. $X'_{ij(-k)}$ is a vector of analogous message and incentive measures for the services other than service k.[8] γ_j are strata fixed effects and ε_{ijk} is an idiosyncratic error term. Because treatment is stratified by course section and class year, the strata fixed effects ensure that we isolate the random variation in treatment assignment across students. These fixed effects also control for any characteristics common within a course section, such as instructor, course subject (microeconomics or macroeconomics), textbook, academic term, etc. They also adjust for characteristics common to class years, for instance if freshmen have different propensities to access academic support services than others. We estimate heteroscedasticity-robust standard errors for all analysis throughout this study.

[8] We include this because the correlation between *X* and *message* is not zero.

Our coefficients of interest are β_1 and β_2. β_1 measures the effect of receiving any message (in the absence of an incentive) on take-up. β_2 captures whether an incentive offer increases take-up above and beyond any effect of the non-incentive component of the message. We also test whether the sum of these coefficients differs from zero, which assesses the combined effect of the message and the incentive on service take-up. We estimate Equation (6) separately for each of the three services and test each of these hypotheses for each of the three services.

A second specification is:

$$takeup_{ijk} = \alpha + \beta_1 email_{ijk} + \beta_2 text_{ijk} + \beta_3 emailincentive_{ijk} + \beta_4 textincentive_{ijk} + X'_{ij(-k)}\Omega + \gamma_j + \varepsilon_{ijk} \qquad (7)$$

where $takeup_{ijk}$ is an indicator variable equal to 1 if student *i* accessed service *k* (e.g., the Economics Tutoring Lab) during the term and 0 otherwise. $email_{ijk}$ is an indicator variable if the student was assigned to receive an email message (at any time) for the given service and $emailincentive_{ijk}$ is an indicator variable if the student was assigned to receive an incentive for the given service in addition to the message. $text_{ijk}$ is an indicator variable if the student was assigned to receive a text message (at any time) for the given service and $textincentive_{ijk}$ is an indicator variable if the student was assigned to receive an incentive for the given service in addition to the message. $X'_{ij(-k)}$ is a vector of analogous message and incentive measures for the services other than service *k*. All other notation follows equation (1).

For Equation (7), we use the sample of individuals who provided phone numbers as our regression sample. We test each treatment coefficient (β_1 through β_4) separately against a null hypothesis of zero. We also test several hypotheses on combinations of these coefficients, as specified in the analysis plan. For brevity, we defer a complete description of these hypotheses to Appendix A.

We repeat all of the analysis thus far in this subsection with a few key modifications. First, we stack all three sets of service outcomes across the same student, resulting in a dataset where each student in our study appears three times, once for each service. We call this dataset the student-service panel. Then we estimate versions of Equations (6) and (7) using student-by-section random effects.[9] We choose random effects for efficiency gains. Treatment assignment should be orthogonal to student characteristics by randomized assignment to treatment, meaning student fixed effects are unnecessary and come at the expense of increased efficiency relative to random effects. These specifications also include service fixed effects, to account for any characteristic of an academic support service common to all students.

By stacking the data and estimating Equations (6) and (7) using random effects OLS, we are able to test whether advertisement increases service use in general, not just for a specific service (e.g., the Academic Coaching Center). We repeat all of the analysis thus far in the sub-section using indicator variables for multiple use of a given service as the outcome instead of indicator variables for using a given service at least once.

4.2 Effects of Message Timing and Frequency on Academic Support Service Take-Up

We estimate the causal effect of the *timing* of the intent to treat with an advertising message on service take-up using the following linear OLS regression:

$$takeup_{ijk} = \alpha + \beta_1 week3_{ijk} + \beta_2 week6_{ijk} + \beta_3 week9_{ijk} + X'_{ij(-k)}\,\Omega + \gamma_j + \varepsilon_{ijk} \qquad (8)$$

where $takeup_{ijk}$ is an indicator variable equal to 1 if student i in section j accessed service k (e.g., the Economics Tutoring Lab) during the term and 0 otherwise. $week3_{ijk}$ is an indicator variable equal to 1 if the student was assigned to receive an advertising message (of any form) in Week 3 for the given service. $week6_{ijk}$ and $week9_{ijk}$ are defined analogously. Note that some students received messages at more than one time. Thus, a student assigned to receive a message in Weeks 3, 6, and 9 would have all of the week

[9] We use student-by-section random effects, rather than student random effects, because the same student can appear multiple times in the dataset if they enroll in more than one course in the study.

indicator variables equal to one. $X'_{ij(-k)}$ is a vector of analogous message timing indicators for the services other than service *k*.

We also estimate the causal effect of the *frequency* of the intent to treat with an advertising message on service take-up. We use the specification:

$$takeup_{ijk} = \alpha + \beta_1 onemessage_{ijk} + \beta_2 twomessages_{ijk} + \beta_3 threeemessages_{ijk} + X'_{ij(-k)}\,\Omega + \gamma_j + \varepsilon_{ijk} \qquad (9)$$

where $takeup_{ijk}$ is as in Equation (8). $onemessage_{ijk}$ is an indicator variable equal to 1 if the student was assigned to receive at least one message (of any form). $twomessages_{ijk}$ and $threeemessages_{ijk}$ are defined analogously. Note that we choose to define these variables additively. Thus, a student assigned to receive a message in Weeks 3, 6, and 9 would have all the frequency indicators set to one. $X'_{ij(-k)}$ is a vector of analogous message and incentive measures for the services other than service *k*.

Appendix A presents a complete description of all hypothesis tests we conduct using Equations (8) and (9). We repeat estimation of Equations (8) and (9) using analogous modifications as for Equations (6) and (7).[10]

4.3 Heterogeneous Effects

10 The methods described in this subsection to estimate the ITT of frequency and timing depart slightly from the analysis plan. The analysis plan collapsed Equations (8) and (9) into a single equation with all timing and frequency indicators. However, in this equation the coefficient on *threemessages* is not identified because it is perfectly collinear with the other timing and frequency indicators. We therefore split the equation into the two specifications described here. For completeness, we present results following the analysis plan in Supplemental Appendix Table SA1. We also present results from a saturated model with all frequency and timing combinations (i.e., separate indicators, for Week 3, Weeks 3 and 6, Weeks 3 and 6, and so on) in Table SA2.

We test for heterogeneous effects by class year, sex, baseline GPA, first generation college student, and self-reported likelihood (at baseline) of using the corresponding academic service. To do so, we estimate a modified version of Equation (6), in which *incentive* is replaced by an interaction between *message* and the aforementioned characteristics. We use the student-service panel version of this regression, with student-by-section fixed effects instead of random effects.[11] We interpret the coefficient estimate on the interaction term as the differential effect for that sub-group.

4.4 Mechanisms

We test for mechanisms in two ways. First, we test whether treatments increased proxies for attention to academic service availability. To do so, we repeat the regressions of Equations (6)-(9), replacing the outcome *takeup* with indicators for "opened email" and for "aware of the corresponding academic service" based on responses to the endline survey.[12]

Second, we test whether the treatments influenced (self-reported) student time use. To do so, we repeat the regressions of Equations (6)-(9), replacing the outcome *takeup* with indicators for how students spent their time during periods of extra effort in Economics Principles. The indicators are for studying less in other courses, sleeping less, reducing time in activities other than sleep, or not spending extra time in Economics Principles. These indicators were constructed from responses to the endline survey.

4.5 Statistical Power

Power calculations included in the analysis plan suggested sufficient statistical power to detect effects of 3.1 percentage points or larger in academic service takeup. We consider these effect sizes both plausible

[11] The student-section indicator is correlated with the interaction between the treatment dummy and the student characteristic, requiring us to use fixed effects instead of random effects.

[12] The endline survey asked students whether they were "aware" of a particular student service (Q11, Appendix B.2). We chose the phrasing "are you aware of this service?" because it is closer to common language than "are you attentive to this service?" Aside from our study communications, all students enrolled in these courses receive information about these three services through instructor announcements in class and flyers for tutoring and academic coaching around campus. Thus, we interpret self-reported awareness as attentiveness.

and policy relevant. The power calculations assumed a comparison between a control group of 750 students against a treatment arm of equal size.[13] We expected to be underpowered to detect effects on academic performance. Nonetheless, we report intent-to-treat effects on student grades later in the paper.[14]

4.6 Data

We collected data on study participants from administrative data, and from baseline and endline surveys. Administrative data on experimental take-up include whether students opened treatment emails, clicked on links within those emails, and the number of visits and dates on which they used academic support services (academic coaching, Economics tutoring, or practice problems). For practice problems, we only observe page views on the course website, not problem completion or performance. The university registrar provided data on student demographics and grades. The baseline survey was conducted in the first two weeks of each term. The endline survey was conducted in Weeks 9-10. To incentivize responses, students earned course credit for completing the surveys. Surveys included questions about likelihood of using academic support services (baseline), awareness of those services (endline), and time use (endline).[15]

To participate in the study, students had to be at least 18 years old, complete the baseline survey, and provide consent. Table 1 presents sample sizes. Panel A shows participation rates. Overall, 2,119 students participated in the study, or 85% of those enrolled. Among participants, 858 (40%) provided a

[13] In other words, three treatment arms corresponding to each academic support service would translate to an overall sample size of 750*4 = 3,000 students. We assumed a significance level (i.e., alpha) of 10%, power (i.e., beta) of 80%. We assumed control group service use of 10%, based on approximations from past sections of Economics Principles.

[14] Our analysis plan also proposed to measure local average treatment effects (LATE) of each support service on course grades by estimating two-stage least squares (2SLS) regressions, instrumenting for service use (e.g., visiting an academic coach) using randomized assignment to the corresponding encouragement. We report this analysis in Supplemental Appendix Table SA4, but caution that the instruments are weak.

[15] The complete surveys appear in Appendix B. Many questions pertain to a companion paper about preferences for majoring in Economics (Pugatch and Schroeder 2020).

mobile phone number, making them eligible for text messages as the treatment medium.[16] Approximately two of three participants were male. Three of five were enrolled in microeconomics.

Panel B shows the distribution of treatments. The first cell reports the size of the control group. The remaining rows report the message type (i.e. which academic support service was advertised), medium, and presence of incentive. The columns report the timing and frequency of messages. The control group included 556 students, or 26% of the total. Coaching, tutoring, and practice problems comprised 26%, 25%, and 23% of the sample, respectively. These proportions are close to our targets of 25% for each group. Nearly 60% of the sample received service advertisement messages by email, while 14% received these messages by text. These uneven proportions arise because students had to provide a phone number to be eligible for the text treatment. When analyzing the effect of message medium on outcomes, we omit students who did not provide a phone number. This restriction mitigates bias from selection into eligibility for the text treatment. Treated students were roughly evenly split between those who received an incentive with their messages and those who did not. Looking across columns in Panel B, treatments were divided nearly evenly among the unique combinations of timing and frequency, with 10-11% in each group.[17]

5 Results

5.1 Randomization balance checks

Table 2 displays the means of baseline characteristics across treatments. We split the table into two parts due to the number of treatment characteristics considered. To check for balance, we regress an indicator for the treatment characteristic in each column on the listed baseline characteristics, plus dummies for randomization strata. The final row of the table reports the p-value from a joint F-test that all of the

[16] Supplemental Appendix Table SA6 reports tests for differences in observable characteristics between students who provided a mobile phone number and those that did not, using the same baseline covariates as the randomization balance checks. We find several significant differences in these characteristics, highlighting the importance of focusing on only those students who provided a mobile phone number when analyzing the effect of message medium.

[17] In Fall quarter 2018, students assigned to emails for Week 6 did not receive these messages. Instead, the Week 6 messages were erroneously sent to the distribution list for Week 3. Text messages went to the proper recipients, however. We report sensitivity of results to this error later in the paper.

regression coefficients on the observable characteristics equal zero. This procedure, and the baseline characteristics chosen, were specified in our analysis plan. None of the p-values fall below conventional significance thresholds (minimum=0.12). This finding is consistent with randomization ensuring that treatment assignment is unrelated to student characteristics.

5.2 Effects of advertising on service attention

For the intervention to influence demand for academic support services, it must first draw attention to those services. Accordingly, we show results for measures of student attention in Table 3. In columns (1)-(2), we test whether students opened treatment emails or clicked links within those emails. These binary outcomes are set to one if a student opened any email or clicked any link during the term, regardless of the number of messages assigned. For these outcomes, we pool messages about different academic support services into a single treatment indicator, rather than include separate variables for each support service. For instance, a student assigned to the coaching treatment and a student assigned to the tutoring treatment would both have the message indicator set to one in Panel A.[18] Columns (1)-(2) also omit students treated with text messages from the estimation sample, as we lack data on these outcomes for texts. For this reason, results for message medium in Panel B are not applicable.

In Panel A, column (1), the message coefficient of 0.662 means that two of every three students assigned to the email treatment opened at least one email message sent as part of the experiment. The coefficient on whether the email included an incentive is positive but not statistically significant. A null effect on the incentive here is unsurprising, as the presence of the incentive was not announced in the message subject. Treated students clicked a link in 0.5% of cases. Though significant at 10%, this effect is small, representing one of every 200 treated students.[19] Incentives did not increase the rate of links clicked.

[18] Results for specific academic services are available upon request.
[19] A caveat to these results is that only the coaching messages included a link allowing students to access the support service (schedule a coaching appointment online, in that case). Redefining the treatment as the coaching message increases the coefficient to .013, meaning 1.3% of students in this treatment arm clicked a link (significant at 10%).

Panel C shows the effect of message timing during the 10-week term. Messages increased rates of opened emails at all times, but had the greatest marginal effect when sent near the beginning and end of the term (Weeks 3 and 9).[20] The pattern for clicking links is the opposite, with the middle of the term (Week 6) being the only significant coefficient. These results may reflect selection into reading emails, with the smaller proportion influenced to read emails in Week 6 being those with higher demand for academic support.

Panel D shows the effect of message frequency. Messages exhibit sharply diminishing marginal returns to opening emails, with the second email increasing opening rates by less than a third of the first email (59 versus 18 percentage points). The third email has no marginal effect on email opening. Message frequency has no effect on clicking links.[21]

The remaining columns of Table 3 show the effect of the intervention on a second proxy for attentiveness: self-reported awareness of academic support services. Columns (3)-(5) show self-reported awareness of specific services, with treatment variables defined specific to each service (controls for assignment to messages for other services are also included in these regressions). Column (6) stacks all services in a student-by-service panel, with service fixed effects and student-section random effects included in the regression. Awareness is a binary variable equal to one if the student indicated on the endline survey that they were aware of the service.[22] In Panel A, column (6) we find that messages increased service awareness by 5.1 percentage points (significant at 1%), relative to a control mean of 65%. Columns (3)-(5) show that this effect was driven by practice problems. Adding an incentive to messages did not increase service awareness in any specification. In Panel B, we find that emails increased service awareness by 11.6 percentage points (column 6; significant at 1%), with roughly equal point estimates for each separate service

[20] Recall that students could receive messages at multiple times during the term. These coefficients therefore represent marginal effects of message assignment in a particular week, conditional on assignment to other weeks. This explains why the coefficients are not as large as the *message* coefficient in Panel A.
[21] When distinguishing between message type, the marginal effect of two coaching messages on clicking links is 2.9 percentage points, significant at 5%. This result is consistent with the pattern of marginal effects for message frequency on email opening.
[22] Endline response rates were 86.5% in the control group and 86.8% in the treatment group. The difference is not statistically significant.

(columns 3-5). The effects of text messages and incentives on service awareness were indistinguishable from zero.

In Panel C, column (6) we find that messages in Weeks 6 and 9 increased service awareness by 3.3 and 4.8 percentage points, respectively. Columns (3)-(5) show that these increases were driven by practice problems, for which the marginal effect of receiving messages in those weeks was 9 percentage points. Weeks 6 and 9 occur after midterm exams in all sections, when students may be more receptive to opportunities for additional practice. In Panel D, we find no statistically significant effects of message frequency on service awareness, with one exception. Sending three messages about practice problems increases awareness of this service by 13.7 percentage points (significant at 1%).

Overall, results in Table 3 demonstrate shifts in student attention to academic support services, consistent with our theoretical framework. We find relatively high rates of email opening among students assigned to advertising messages. Messages also increased service awareness, albeit modestly, with effects driven by email messages and advertising messages to solve practice problems. Text messages and incentives did not increase student awareness of academic support services. Increases in awareness specific to Weeks 6 and 9, when the endline survey and final exams are approaching, also suggest attention plays a role. In Section 6.1, we build on these reduced-form results to characterize student attention in the context of our theoretical model.

5.3 Effects of advertising on service take-up

Did encouragement to use academic support services increase their use? Table 4 shows intent-to-treat effects of message assignment and incentives. In columns (1)-(3), the outcomes are binary indicators for use of the listed service at any time during the term. The reported treatment coefficients are specific to that service, with controls for messages about other services included in the regression. Column (4) "stacks" all services in a student-by-service panel, with service fixed effects and student-section random effects included in the regression. In other words, the specifications mirror those of Table 3, Panel A, but with service use as the outcome. We find no significant effects of advertising messages on use of any service in these restricted specifications. Nor do incentives attached to these messages increase service use.

Table 4, columns (5)-(8) repeat the regressions of the first four columns, but here the outcome is a binary variable set to one if the student used the service more than once. In the panel regression of column (8), we find that advertising messages increased multiple service use by 1.2 percentage points (significant at the 10% level). A significant increase in multiple service use, but not any service use, is consistent with increased attention, as some students already using the service are induced to use it more intensively. This effect is driven by practice problems, which increased 3.4 percentage points in response to messages (column 7). This effect is small relative to the 90% of the control group that viewed practice problems multiple times. It is nonetheless notable, as this effect closes more than one-third of the gap between universal multiple use by all students and actual use. Adding incentives to messages had no differential effect on multiple service use.[23]

Table 5 tests whether service use responds differently based on message medium. Sample sizes drop substantially in these regressions because estimation is limited to those who provided a mobile phone number. In the panel specification of column (4), we find an increase of 2.9 percentage points in response to email messages (significant at 5%). We find no significant change in response to texts, or for incentives attached to either message medium. Although the email and text coefficients do not differ significantly from each other (H7 at bottom of table), the effect of emails with incentive does differ from texts with incentive (p=.04). Increased service use from emails is driven by a 6.7 percentage point increase in use of practice problems, significant at 1% (column 3). Another notable result is a negative and significant effect of texts with incentive on tutoring (column 2, H6 at bottom of table).

Table 5, columns (5)-(8) show a similar pattern for multiple service use, but with larger magnitudes. Emails increased multiple service use by 3.8 percentage points overall, driven by an 11 percentage point increase for practice problems. This increase implies universal multiple use of practice problems for

[23] Encouragement messages might also induce students to substitute one type of service for another. We suppress results for cross-service effects for brevity in Table 4, but find no evidence of substitution across services.

students sent advertisements via email. However, we find a negative and significant effect of similar magnitude (9 percentage points) when encouragement emails for practice problems include an incentive. Students may have perceived the incentive as a gimmick, steering them away from the practice problems. We find no effect of text messages on multiple service use.

5.4 Effects of message timing and frequency on service take-up

Does the timing or frequency of messages affect service use? Table 6, Panel A presents results for timing. All significant coefficients for timing are for Week 9, the final period of messages and the closest to the final exams in Week 11. Receiving a message in Week 9 increases multiple service use by 1.6 percentage points (significant at 5%; column 8). As before, the effect is driven by practice problems, for which a message in Week 9 generates a 2.9 percentage point increase in use and 4.1 percentage point increase in multiple uses. We also find several cases in which a message in Week 9 has a significantly greater effect than in Week 3 or Week 6, further strengthening the case for messages late in the term.[24]

In Panel B, we find the marginal effect of a third message is to *reduce* service use by 2.1 percentage points (significant at 10%; column 4). This effect is driven by a 1.7 percentage point decrease in coaching, a magnitude that exceeds the control group mean of 1.3% (column 1). We also find that the overall effect of receiving three messages is negative and significant for tutoring (column 2, H5 at bottom of table). For multiple service use, we again find a negative marginal effect for the third message, this time a 2.9 percentage point decline in multiple tutoring visits. However, we find a 1.8 percentage point increase in overall multiple service use in response to a second message (column 8). The results suggest that while two messages may be effective to promote academic service use, the third message can turn students away.[25]

---

[24] We also explore whether students use academic services within two weeks of each message, as incentivized by the lottery version of messages. We fail to find robust evidence that the lottery message affected this outcome (Supplemental Appendix Table SA7).

[25] We check robustness of our results to multiple appearances in the data by the same student. During the three terms of the study, 342 students (of 1,777, or 19%) take more than one Economics Principles course. An additional 178 students participated in a pilot of this study in Spring 2018. For the robustness check, we keep only those students whose first exposure to the project was during the analysis period of academic year 2018-2019. If a student enrolled

5.5 Grade outcomes

We report treatment effects on academic performance in Supplemental Appendix Table SA5. The outcomes are student course grade (0-4 scale) and "DFW," which is an indicator for a grade of D, F, or withdrawal from the course. Columns (1)-(2) report intent to treat effects using OLS regression, as in other analyses. Columns (3)-(4) define treatment as opening an email, using treatment assignment as instrumental variables in two-stage least squares (2SLS) regressions. Across most specifications, treatment effects have positive sign for grades and negative for DFW, as expected. However, few effects are statistically significant.[26] Lack of significant results is unsurprising, as the modest magnitudes of effects on academic service use leave us underpowered to detect changes in academic performance. The null effects are also consistent with Oreopoulos and Petronijevic (2019) and Page, Lee, and Gehlbach (2020), which find little effect of a range of interventions on academic performance.

5.6 Heterogeneity analyses

Did the intervention differentially affect particular subgroups of students? Table 7 presents tests for heterogeneous treatment effects by observable student characteristics. We limit the analysis by collapsing all treatments into a single indicator within the student-service panel specification, reducing the number of interactions to test. We interact treatment with indicators for freshman; sophomore; male; GPA below sample median; first generation student; and low service demand, defined as being at or below median demand for coaching and tutoring in the baseline survey.[27] Freshman and sophomore interactions appear

---

in both microeconomics and macroeconomics in their first term in the data, we keep data from the microeconomics course. Our results remain largely unchanged using this restricted sample (n=1,599 students). One notable difference is that a third message reduces multiple service use by 2.4 percentage points (significant at 10%), whereas this coefficient was not statistically significant in the full sample (Supplemental Appendix Tables SA8-SA10).

[26] The few exceptions are difficult to interpret and may therefore be spurious. For instance, the marginal effect of a second message on DFW rates is positive.

[27] The survey asked, "Suppose you encounter difficulties in a course. If free [peer tutoring/academic coaching] were available, how likely would you be to use it?" Responses were "definitely wouldn't use," "unlikely to use," "not sure if I would use or not," "might use," and "definitely would use" (Q15-16, Appendix B.1). We coded responses

in the same regression, while all other characteristics appear in separate regressions. All subgroups considered were listed in the analysis plan. Regressions include student and service fixed effects.

For any service use as the outcome, we find no statistically significant interaction terms (columns 1-6). This means that no subgroup had a distinct treatment effect from its respective omitted category, e.g, males compared to females, first generation to non-first generation, etc. However, for several subgroups we find a significant effect compared to the same subgroup within the control group, as reported by the p-value of the treatment effect plus interaction at the bottom of the table. We find positive, statistically significant effects for freshmen, below median GPA, and low service demand, in comparison to students from these subgroups in the control group.

For multiple service use as the outcome, we find an increase of 1.1 percentage points among all students (significant at 5%; column 7). We also find positive and significant effects for freshmen, males, GPA below median, not first generation, and low service demand (columns 1-6).[28] The pattern of positive effects among subgroups is largely internally consistent. Most subgroups with positive and significant responses to treatment are likely to have less knowledge of (e.g., freshmen) and greater need for academic support services (e.g., below median GPA). It is therefore sensible that these groups would be most responsive to advertising about these services. The lone exception is first generation students, for whom we might expect greater response via the same logic. In this case, first generation students may already seek out these services regardless of messaging, or otherwise face greater barriers to their use than such messages can mitigate.[29] The pattern of results also suggest that both the information and (pure) attention channels matter, as groups likely to have less information (e.g., freshman) and more (e.g., not first generation) information than their counterparts both respond to the intervention.

1-5 in ascending order of demand. We calculate the mean response to the coaching and tutoring versions of this question. “Low service demand” is a dummy for falling at or below the sample median of this variable.

[28] All inferences based on p-values at bottom of table, except not first generation, which is from the coefficient for the main effect of treatment.

[29] First generation students are eligible for additional academic support. For instance, they may use TRIO Student Support Services, a federal program offering academic guidance, professional development, peer mentorship, technology access, student success seminars and additional services. We lack data on which students access TRIO or similar services.

5.7 Analysis of mechanisms

Table 3 presented results of the intervention on proxies for service attentiveness, a key potential mechanism underlying changes in service use. The intervention might also influence student time use. Students who use the services advertised by the intervention have less time to spend on other activities. Even those who fail to use these services may nonetheless alter their study habits in other ways in response to treatment. Table 8 presents results for student time use, with outcomes from the endline survey (Q12, Appendix B.2). The format of the table follows Table 3, with results for message type and incentive, medium, timing, and frequency in Panels A-D, respectively.

We find no evidence of altered time use in response to treatment, with a single exception. Students assigned three messages were 3.7 percentage points less likely to report spending no extra time studying economics, significant at 10%. In other words, sending a third message made students more likely to report additional study effort in economics. This result is interesting in light of our findings on service use in response to message frequency (Table 6, Panel B), in which the only significant marginal effects of the third message were negative, i.e., reduced service use. A potential explanation is that the third message led students to use academic services less, but study more on their own. We interpret this result cautiously, however. In addition to relatively low precision, the self-reported nature of the outcome may reflect social desirability bias among students reminded the most frequently to increase their studying.[30]

6 Structural estimation

[30] We also check robustness to misclassification of treatments. In Fall quarter 2018, students assigned to emails for Week 6 did not receive these messages. Instead, the Week 6 messages were erroneously sent to the distribution list for Week 3. Text messages went to the proper recipients, however. Due to the error, 13% of observations fall into a group different from their randomly assignment along at least one dimension of treatment (though only 3% switch between any treatment and control). We reanalyze the data by regressing outcomes on treatments received, using the random assigned treatments as instruments in two-stage least squares estimation. Results remain largely unchanged. However, the magnitude and precision of most significant coefficients increase, which is unsurprising as the procedure corrects for measurement error. Differences are particularly pronounced for the Week 6 coefficients on message takeup and awareness in Table 3, Panel C. This result is also sensible, because the distribution error reclassified Week 6 email recipients the most. See Supplemental Appendix Tables SA11-SA17.

6.1 Attention types

In this section, we apply the model of academic service use from Section 2 to the data. As a first step, we characterize the function *m(Z)*, where *m* is the binary attention parameter and *Z* indicates assignment to our experimental intervention. As described in Section 2.2, we can classify students as always attentive, never attentive, or compliers based on their (potential) response to the experiment. Student attention *m* is not observed. However, we can proxy for attention using survey data on service awareness.

Following Kowalski (2016), we use the experimental variation to calculate the proportion of the sample within each group. We classify control group students who report awareness of an academic service at endline as always attentive; their proportion can be written as $\Pr(m=1|Z=0)$. Students assigned an advertising message for a service but who are unaware of that service at endline are never attentive; their proportion is $\Pr(m=0|Z=1)$. The proportion of compliers is the difference in service awareness between treatment and control students, $\Pr(m=1|Z=1) - \Pr(m=1|Z=0) = \Delta m(Z)$, i.e., the first stage. Assuming no defiers (students who would be unaware if treated but aware if untreated), the proportions of always attentive, never attentive, and compliers sums to one.

Table 9 shows the proportions of the sample within each group, separately for each academic service.[31] In column (1), we find that 55% of students are always attentive to coaching services, 45% are never attentive, and 1% are compliers (the sum exceeds one due to rounding error). In other words, the experiment barely changed student attention to coaching services. We find greater student attention, and a higher response to the experiment, for tutoring and practice problems in columns (2)-(3). For these services, always attentive students represent around 70% of the sample. Compliers are 7% for each service. Finally, in column (4), we stack all services in a student-service panel, so that estimates represent a weighted average of the three services. In this case, always attentive students are 65% of the sample and compliers 5%.

[31] To ensure we rely on the random variation from the experiment, we adjust for randomization strata when calculating all proportions. Specifically, to find the proportion of always attentive students, we restrict the sample to the control group, then regress the awareness dummy on a constant, after demeaning by randomization strata. The proportion of never attentive follows the same procedure, altering the outcome to (1-awareness) and limiting the sample to the treatment group. The proportion of compliers is the coefficient on an indicator for treatment in the full sample.

Another way of viewing results in Table 9 is to focus on the share of never attentive students. Across services, never attentive students range from 22%-45%, with a weighted average of 30%. In other words, nearly one third of students are inattentive to the availability of free academic support services, even if encouraged to use them. The encouragement shifts attention for less than 10% of students across all services.[32]

Do our results merely reflect an information channel, rather than attention? To check robustness, we limit the sample to students attentive at baseline and repeat the exercise of Table 9. This sample restriction should narrow results to the pure attention, rather than information, channel.[33] We find similar results as for the full sample, with the share of compliers for extra practice problems increasing to 15%, suggesting even greater effects on attention than for the full sample.[34]

6.2 Structural parameter estimates

Having characterized the attention function *m(Z)*, we now turn to estimating the parameters of the service choice model. To make estimation feasible, we add a stochastic component to the value function of equation (1):

$$V(D) = -Dc + \delta(1 + \theta D)\underline{y} + \varepsilon_D \qquad (10)$$

In other words, we modify the model to include a choice-specific shock ε, so that service use is no longer a deterministic function of the model parameters. Recall from equation (3) that $\theta = b + m\tau_b$, where

[32] Results for the share of never attentive students are similar when using data on opening encouragement emails as the attention proxy. In this case, compliers represent 67-74% of the sample, depending on the service, a much larger share than the awareness proxy. However, in this case we cannot define always attentive students, because by definition no control group students receive an email to which they can pay attention. The treated students who do not open encouragement emails are therefore never attentive, with sample shares between 26-33%. Full results available upon request.

[33] For coaching and tutoring, we define attentive at baseline defined as indicating plans to use the service in the baseline survey (Q15-16, Appendix B.1). For extra practice problems, we define attentive at baseline defined as viewing extra practice problems during Weeks 3-5 even though the student was not assigned an encouragement message in Week 3.

[34] We also find similar results for this subsample when repeating the exercise of Table 10 in the next subsection. See supplemental appendix Tables SA18-SA19.

b is the benefit of service use commonly known to all students, and τ_b is the shrouded value of service use visible only to attentive students.

Within the model, students choose whether to use an academic support service by comparing their expected grades with and without service use, net of costs. Because grades are observed for only one choice for each student, we must calculate a counterfactual grade for each student to generate the relative values of $V(1)$ and $V(0)$. By definition, service use is endogenous, making comparisons of grades between users and non-users of services problematic. In principle, we could use the experimentally-induced variation in service use to generate an unbiased estimate of b, the common return to service use, using two-stage least squares (2SLS). Unfortunately, the instruments are weak.[35] We therefore estimate the model under a range of calibrated values for b. We choose b=0.001, b=0.01, and b=0.1, i.e., we assume service use generates negligible returns or returns as large as 10%, equivalent to the difference between grades of B (3.0) and B+ (3.3). We calibrate the discount factor δ=0.95.[36] Finally, we assume the shock ε is distributed Type I Extreme Value, allowing us to estimate the model using logistic regression.

Our goal is to estimate c, the cost to service use (in GPA units), and τ_b, the shrouded benefit of service use to attentive students. Within this structure, the cost parameter c is simply the constant term in the logistic regression for service use D. The shrouded benefit τ_b is the coefficient on the interaction between (discounted) grades and attention, $\delta m\underline{y}$. As with grades, attention m (and its proxies) are likely correlated with unobserved student characteristics, generating bias. We therefore replace m with Z, an indicator for random assignment to an advertising message for the service, allowing us to rely on experimental variation when estimating τ_b. Because $m=m(Z)$, this approach is analogous to estimation of the intent to treat (ITT), using the attention shifter Z directly in estimation. We convert this estimate to a

[35] First-stage F-statistics achieve a maximum value of 5.7 under various specifications. 2SLS estimates of returns are implausibly large in magnitude, often exceeding 4 (on a 4-point scale) or 100%, and changing sign across specifications (Supplemental Appendix Table SA4). We find these results implausible.
[36] Results (available upon request) change little when choosing δ=0.99.

local average treatment effect (LATE) by scaling the ITT by the share of compliers, i.e., by dividing by $\Delta m(Z)$ from Table 9. Appendix C provides further details on parameter estimation.

Structural estimates of the cost c and shrouded benefit τ_b of service use appear in Table 10. Each column represents a different academic support service, with column (4) reporting results from stacking all services into a student-service panel. Each panel of the table uses a different calibrated value of b, the common benefit of service use. In Panel A, we assume b=0.001, a negligible rate of return. Panel A, column (1), reports a cost of using academic coaching of 4.3 grade points, on a 0-4 scale. In other words, even if coaching moved the course grade from F (0.0) to A (4.0), students behave as though visiting an academic coach is more costly. Though extremely large, this parameter value is perhaps unsurprising given the near-zero use of coaching in the sample. The ITT for the shrouded benefit of coaching is -0.13, i.e., the advertising message reduces the perceived benefits of coaching by 13%. However, this point estimate is not statistically distinguishable from zero, nor is the corresponding LATE.

For tutoring, the estimated cost is -3.0 (e.g., the difference between course grades of A and D), also reflecting the relatively low take-up of this service (Panel A, column 2). The shrouded benefit is positive (ITT=.05, or 5%), but not statistically significant. For practice problems, the cost reverses sign (Panel A, column 3). The point estimate of -2.78 means students derive a consumption value of 2.78 GPA points from using practice problems. This parameter justifies the near-universal use of practice problems in the sample under the assumption of negligible common benefits. The ITT of the shrouded benefit for practice problems is 0.23 (significant at 5%). In other words, the experiment revealed additional value of practice problems to students. In column (4), estimates from the student service panel largely align with those of coaching and tutoring: high costs and shrouded benefits too noisy to distinguish from zero.

In Panels B-C, estimated costs grow relative to Panel A, because higher costs are necessary to reconcile low take-up under the assumed increase in benefits. Shrouded benefits remain similarly noisy.

Across all scenarios, LATE estimates tend to be implausibly large in magnitude, often exceeding one (100%) in absolute value.[37]

7 Discussion

We find shifts in student attention to academic support services in response to advertising messages, consistent with our theoretical framework. Most students sent advertisements by email read those messages. Messages also increased self-reported awareness of student services. Although magnitudes were modest, the results suggest some shifts in student attention in response to the intervention.

Experimental variation allows us to characterize the proportion of always attentive (approximately two-thirds), never attentive (one-third), and compliers (one-twentieth) in the sample. The proportion always attentive may be compared to the mean of the attention parameter m in Equation (3). Previous studies have found estimates of the attention parameter m from other settings ranging from 0 to 0.82, with an average of 0.44 (Gabaix 2019). Our results suggest that university students may be more attentive than some groups (e.g., eBay buyers purchasing iPods; Hossain and Morgan 2006), although not as attentive as other groups (e.g., automobile buyers; Lacetera, Pope, and Sydnor 2012).

Within the theoretical framework, shifts in student attention are insufficient to change the decision about whether to use academic support services. Newly attentive students must also value the expected benefits of service use as greater than the costs. In the data, we do find increases in academic support service use in response to advertising messages. These increases were not universal, however, but instead driven by practice problems. Unlike coaching and tutoring, practice problems could be accessed immediately online following receipt of an advertising message, without requiring a trip to a separate location. Nor do practice problems require interaction and the implied revelation of academic difficulties to peers. This difference in transaction costs among services could explain the stronger effects for practice problems.

[37] To check if we overestimated the shrouded benefit by setting the common benefit too low, we set b=0.5, or the difference between a grade of C (2.0) and B (3.0). In this scenario, costs increase still further, while the shrouded benefit to attentive students turns negative for all services. We find this result implausible, consistent with the implausibly large common benefit assumed.

Email was an effective advertising medium, whereas text messages were not. One explanation for this difference is that the email messaging appeared to come from the instructor, whereas the text messaging appeared to come from an anonymous source. Another explanation is that email access is more closely linked to practice problem access. That is, the platform for accessing the practice problems is readily accessed via laptop/desktop computer, where access via smartphone is more limited.

Perhaps surprisingly, the offer of a financial incentive did not additionally increase take-up, and in many specifications, actually reduced take-up relative to messaging without an economic incentive. It is possible that the incentive offer crowded out intrinsic motivation (e.g., List, Livingston, and Neckermann 2018). Alternatively, it may have sounded like a scam, as financial rewards for studying are not offered elsewhere by the university.

Study participants exhibited some evidence of message fatigue. In general, two messages were more effective than a single message, yet three messages did not have a statistically significant effect. This pattern suggests judicious use of reminders to increase the desired behavior, consistent with other studies among parents of younger students (Cortes et al. 2019a; 2019b). In light of message fatigue, ascertaining the optimal timing of a limited number of messages becomes more valuable. Our results suggest that messaging about academic support services is most effective toward the end of the academic term. Although the time available for increased learning may not be as great as support service take-up toward the beginning of the term, messaging at the beginning of the term had no effect on support service take-up.

Student subgroups responding most strongly to advertising messages included freshmen, males, GPA below median, not first generation, and low initial service demand. With the exception of not first generation, effects among these subgroups are sensible in the context of our model. Greater differences between attentive and inattentive decisions are likely for freshmen, by virtue of their inexperience, and males, due to their susceptibility to distraction (Stinebrickner and Stinebrickner 2008; Lindo, Swensen, and Waddell 2012). For students with below median GPA or low initial service demand, the intervention should lead to greater revisions in the perceived returns to service use. By comparison, students with higher initial grades or demand would be less likely to update their perceived net benefits in response to messages. Low take-up among first generation students is consistent with our structural estimates of implicitly high costs to service use, as the obstacles these students face may exceed the light-touch encouragement provided by the intervention.

Our effect sizes on service use are fairly modest in absolute terms, yet often are economically meaningful. For example, the point estimates for the email indicator variable in Table 5 indicate that the email advertising (without the incentive) virtually eliminated the proportion of students not accessing these problems. Similarly, the point estimates in the student-by-services panel in Table 6, Panel B indicate that sending three emails eliminated the beneficial effect of two emails.[38]

The most effective intervention – two emails (without an incentive) reminding students to access practice problems – was highly cost-effective. Creating and sending these two emails took approximately 30 minutes or less. In a 200-person class, an approximately 5 percentage point increase in the proportion of students accessing practice problems is an increase of 10 students. Suppose instructor time costs $50 per hour. Then our most effective intervention cost approximately $2.50 per student induced to access the support service. This cost compares favorably to other interventions. For instance, Pugatch and Wilson (2018b) report costs of $3.32-$14.58 per additional tutoring visit, while Paloyo, Rogan, and Siminski (2016) report a cost of $30.72.

Our structural estimates reveal three stylized facts. First, shrouded benefits are large relative to the common (or base) benefit (e.g., often exceeding the common benefit), yet are only statistically significant for practice problems. Second, the perceived costs of coaching and tutoring are incredibly high. Measured in terms of GPA, they are equal to the difference between an average-to-above-average letter grade (e.g., A or B) and unsatisfactory performance (e.g., D). These costs exceed plausible opportunity costs such as

[38] A potential objection to our interpretation is that the results are spurious, given the large number of hypotheses we test. When we apply a multiple hypothesis correction to Tables 4-6 to control the false discovery rate (Benjamini, Krieger, and Yekutieli 2006), treatment coefficients significant at 10% or 5% based on conventional p-values lose significance, but those significant at 1% remain so after the correction. This suggests the basic pattern of results holds, albeit with caution when interpreting results. Results available upon request.

one hour of work at a minimum wage job.[39] Third, the perceived cost of practice problems is negative, indicating students enjoy consumption value from this service.

Given the dramatically large and noisy structural parameter estimates, what do we learn from this exercise? One possible conclusion is that the structural model is improperly specified.[40] Certainly this is the case if we wish to interpret the parameter estimates literally. But rather than focus narrowly on parameter magnitudes, we interpret the results as illustrative of behavioral inertia among college students in response to nudges (as in Oreopoulos and Petronijevic 2019). Our experiment "moved the needle" on academic support service use, as we showed in Section 5. But these changes occurred at the margin. The vast majority of students behave as though service use carries enormous transaction costs far beyond the opportunity cost of time and minimal benefits, even when told otherwise by nudges from their instructors.

8 Conclusion

We conducted a large-scale field experiment at a major public university in the United States to test messaging designed to increase demand for academic support services. Several policy recommendations follow from our results. First, automated emails from instructors sent toward the end of the term advertising extra practice problems were effective at promoting student engagement with this service, suggesting universities scale up this messaging approach. Second, two seemingly "cutting edge" solutions – economic incentives and text messaging – do not appear to be effective, at least in our study setting. Third, repeated messages increased multiple service use, yet evidence of message fatigue suggests messaging in moderation.

Future research should examine whether advertising-induced use of academic support services improves learning outcomes. Standard selection models would predict that individuals induced to use these

[39] During the study period, the (non-Portland Metro area) Oregon minimum wage was $10.50. https://hr.oregonstate.edu/employees/administrators-supervisors/classification-compensation/new-oregon-minimum-wage-rate (accessed August 8, 2020).
[40] We also attempted to estimate the model by including a shrouded cost for attentive students, i.e., modeling costs as $-c + m\tau_c$. Parameter estimates for both τ_c and τ_b were implausibly large in magnitude and statistically indistinguishable from zero.

services only through advertising may be those students with smaller expected benefits compared to students who choose to use these services in the absence of additional advertising. However, behavioral biases – such as inattention due to present bias – may mean that advertising-induced tutoring can have a large impact on learning outcomes. A major barrier to this question in a nudge-based experiment – one that we were not able to overcome – is that the nudge must generate sufficiently large increases in take-up to ensure sufficient power to measure an effect on learning.

9 References

Allcott, Hunt, and Nathan Wozny. 2014. "Gasoline Prices, Fuel Economy, and the Energy Paradox." *Review of Economics and Statistics* 96 (5): 779–795.

Angrist, Joshua, Daniel Lang, and Philip Oreopoulos. 2009. "Incentives and Services for College Achievement: Evidence from a Randomized Trial." *American Economic Journal: Applied Economics*, 136–163.

Auferoth, Florian. 2020. "Who Benefits from Nudges for Exam Preparation? An Experiment." SSRN Scholarly Paper ID 3424583. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3424583.

Balaban, Rita, and Patrick Conway. 2020. "A Test of Enhancing Learning in Economics through Nudges." *AEA Papers and Proceedings* 110 (May): 289–93. https://doi.org/10.1257/pandp.20201050.

Barrow, Lisa, Cecilia E. Rouse, and Amanda McFarland. 2020. "Who Has the Time? Community College Students' Time-Use Response to Financial Incentives." SSRN Scholarly Paper ID 3584755. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3584755.

Benjamini, Yoav, Abba M. Krieger, and Daniel Yekutieli. 2006. "Adaptive Linear Step-up Procedures That Control the False Discovery Rate." *Biometrika* 93 (3): 491–507. https://doi.org/10.1093/biomet/93.3.491.

Bettinger, Eric P., and Rachel B. Baker. 2013. "The Effects of Student Coaching An Evaluation of a Randomized Experiment in Student Advising." *Educational Evaluation and Policy Analysis*, 0162373713500523.

Bettinger, Eric P., Benjamin L. Castleman, and Zachary Mabel. 2019. "Finishing the Last Lap: Experimental Evidence on Strategies to Increase College Completion for Students At Risk of Late Departure."

Butcher, Kristin F., and Mary G. Visher. 2013. "The Impact of a Classroom-Based Guidance Program on Student Performance in Community College Math Classes:" *Educational Evaluation and Policy Analysis*, September. https://doi.org/10.3102/0162373713485813.

Carrell, Scott E, and Michal Kurlaender. 2020. "My Professor Cares: Experimental Evidence on the Role of Faculty Engagement." Working Paper 27312. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w27312.

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and Taxation: Theory and Evidence." *The American Economic Review* 99 (4): 1145–77.

Clark, Damon, David Gill, Victoria Prowse, and Mark Rush. 2017. "Using Goals to Motivate College Students: Theory and Evidence from Field Experiments." Working Paper 23638. National Bureau of Economic Research. https://doi.org/10.3386/w23638.

Cortes, Kalena E, Hans D.U. Fricke, Susanna Loeb, David S Song, and Benjamin N York. 2019a. "When Behavioral Barriers Are Too High or Low – How Timing Matters for Parenting Interventions." Working Paper 25964. National Bureau of Economic Research. https://doi.org/10.3386/w25964.

Cortes, Kalena E., Hans Fricke, Susanna Loeb, David S. Song, and Benjamin N. York. 2019b. "Too Little or Too Much? Actionable Advice in an Early-Childhood Text Messaging Experiment." *Education Finance and Policy*, December, 1–44. https://doi.org/10.1162/edfp_a_00304.

Courtin, Emilie, Peter Muennig, Nandita Verma, James A. Riccio, Mylene Lagarde, Paolo Vineis, Ichiro Kawachi, and Mauricio Avendano. 2018. "Conditional Cash Transfers and Health of Low-Income Families in the US: Evaluating the Family Rewards Experiment." *Health Affairs* 37 (3): 438–446.

Damgaard, Mette Trier, and Helena Skyt Nielsen. 2018. "Nudging in Education." *Economics of Education Review*, March. https://doi.org/10.1016/j.econedurev.2018.03.008.

DellaVigna, Stefano, and Joshua M. Pollet. 2009. "Investor Inattention and Friday Earnings Announcements." *The Journal of Finance* 64 (2): 709–749.

Dobronyi, Christopher R., Philip Oreopoulos, and Uros Petronijevic. 2019. “Goal Setting, Academic Reminders, and College Success: A Large-Scale Field Experiment.” *Journal of Research on Educational Effectiveness* 12 (1): 38–66.

Ericson, Keith Marzilli. 2017. “On the Interaction of Memory and Procrastination: Implications for Reminders, Deadlines, and Empirical Estimation.” *Journal of the European Economic Association* 15 (3): 692–719. https://doi.org/10.1111/%28ISSN%291542-4774/issues.

Gabaix, Xavier. 2019. “Behavioral Inattention.” In *Handbook of Behavioral Economics: Applications and Foundations 1*, 2:261–343. Elsevier.

Gordanier, John, William Hauk, and Chandini Sankaran. 2019. “Early Intervention in College Classes and Improved Student Outcomes.” *Economics of Education Review* 72 (C): 23–29.

Hossain, Tanjim, and John Morgan. 2006. ... “... Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on Ebay.” *The BE Journal of Economic Analysis & Policy* 6 (2).

ideas42. 2015. “Increasing Use of On-Campus Tutoring: Helping Students Achieve Academic Success.” http://www.ideas42.org/wp-content/uploads/2015/12/WK-Brief.pdf.

Imbens, Guido W., and Joshua D. Angrist. 1994. “Identification and Estimation of Local Average Treatment Effects.” *Econometrica* 62 (2): 467–475.

Kowalski, Amanda Ellen. 2016. “Doing More When You’re Running Late: Applying Marginal Treatment Effect Methods to Examine Treatment Effect Heterogeneity in Experiments.” SSRN Scholarly Paper ID 2800873. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2800873.

Lacetera, Nicola, Devin G. Pope, and Justin R. Sydnor. 2012. “Heuristic Thinking and Limited Attention in the Car Market.” *American Economic Review* 102 (5): 2206–36.

Laibson, David. 1997. “Golden Eggs and Hyperbolic Discounting.” *The Quarterly Journal of Economics* 112 (2): 443–78. https://doi.org/10.1162/003355397555253.

Lindo, Jason M., Isaac D. Swensen, and Glen R. Waddell. 2012. “Are Big-Time Sports a Threat to Student Achievement?” *American Economic Journal: Applied Economics* 4 (4): 254–274.

List, John A., Jeffrey A. Livingston, and Susanne Neckermann. 2018. “Do Financial Incentives Crowd out Intrinsic Motivation to Perform on Standardized Tests?” *Economics of Education Review* 66 (October): 125–36. https://doi.org/10.1016/j.econedurev.2018.08.002.

Miller, Cynthia, Rhiannon Miller, Nandita Verma, Nadine Dechausay, Edith Yang, Timothy Rudd, Jonathan Rodriguez, and Sylvie Honig. 2016. “Effects of a Modified Conditional Cash Transfer Program in Two American Cities: Findings from Family Rewards 2.0.” *New York: MDRC, September*.

National Center for Education Statistics. 2016. “Digest of Education Statistics.” 2016. https://nces.ed.gov/programs/digest/d16/tables/dt16_326.10.asp.

O’Donoghue, Ted, and Matthew Rabin. 1999. “Doing It Now or Later.” *American Economic Review* 89 (1): 103–24.

Oreopoulos, Philip, and Uros Petronijevic. 2018. “Student Coaching: How Far Can Technology Go?” *Journal of Human Resources* 53 (2): 299–329. https://doi.org/10.3368/jhr.53.2.1216-8439R.

———. 2019. “The Remarkable Unresponsiveness of College Students to Nudging And What We Can Learn from It.” Working Paper 26059. National Bureau of Economic Research. https://doi.org/10.3386/w26059.

Page, Lindsay C., Jeonghyun Lee, and Hunter Gehlbach. 2020. “Conditions under Which College Students Can Be Responsive to Nudging.” *EdWorkingPapers.Com*. Annenberg Institute at Brown University. https://www.edworkingpapers.com/ai20-242.

Paloyo, Alfredo R., Sally Rogan, and Peter Siminski. 2016. “The Effect of Supplemental Instruction on Academic Performance: An Encouragement Design Experiment.” *Economics of Education Review* 55 (Supplement C): 57–69. https://doi.org/10.1016/j.econedurev.2016.08.005.

Pop-Eleches, Cristian, Harsha Thirumurthy, James P. Habyarimana, Joshua G. Zivin, Markus P. Goldstein, Damien De Walque, Leslie Mackeen, Jessica Haberer, Sylvester Kimaiyo, and John

Sidle. 2011. “Mobile Phone Technologies Improve Adherence to Antiretroviral Treatment in a Resource-Limited Setting: A Randomized Controlled Trial of Text Message Reminders.” *AIDS (London, England)* 25 (6): 825.

Pugatch, Todd, and Elizabeth Schroeder. 2020. “Promoting Female Interest in Economics: Limits to Nudges.”

Pugatch, Todd, and Nicholas Wilson. 2018a. “Addressing Barriers to Student Success in Higher Education.” American Economic Association. https://doi.org/10.1257/rct.2744-2.0.

———. 2018b. “Nudging Study Habits: A Field Experiment on Peer Tutoring in Higher Education.” *Economics of Education Review* 62 (Supplement C): 151–61. https://doi.org/10.1016/j.econedurev.2017.11.003.

Riccio, James. 2010. “Sharing Lessons from the First Conditional Cash Transfer Program in the United States.” *Ann Arbor, MI: National Poverty Center*.

Rury, Derek, and Scott Carrell. 2020. “Knowing What It Takes: The Effect of Information About Returns to Studying on Study Effort and Achievement,” July. http://conference.iza.org/conference_files/edu_2020/rury_d29540.pdf.

Stinebrickner, Ralph, and Todd R. Stinebrickner. 2008. “The Causal Effect of Studying on Academic Performance.” *The BE Journal of Economic Analysis & Policy* 8 (1). http://www.degruyter.com/view/j/bejeap.2008.8.1/bejeap.2008.8.1.1868/bejeap.2008.8.1.1868.xml.

Taubinsky, Dmitry, and Alex Rees-Jones. 2018. “Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment.” *The Review of Economic Studies* 85 (4): 2462–2496.

U.S. Department of Education. 2019. “The Condition of Education 2019.” NCES 2019-144. National Center for Education Statistics. https://nces.ed.gov/programs/coe/indicator_cue.asp.

10 Appendix A: hypothesis tests

This appendix describes in greater detail the hypotheses we test for Equations (6)-(9), our main estimating equations. We follow an analysis plan submitted prior to analyzing the data (Pugatch and Wilson 2018a).

We test the following hypotheses from Equation (6):

$$takeup_{ijk} = \alpha + \beta_1 message_{ijk} + \beta_2 incentive_{ijk} + X'_{ij(-k)}\,\Omega + \gamma_j + \varepsilon_{ijk} \qquad (6)$$

1. $\beta_1 = 0,$
2. $\beta_2 = 0,$
3. $\beta_1 + \beta_2 = 0.$

The first hypothesis is that receiving any message (in the absence of an incentive) increases take-up. The second hypothesis is that receiving an incentive offer increases take-up above and beyond any effect of the non-incentive component of the message. The third hypothesis is that the combined effect of the message and the incentive is greater than zero. We estimate Equation (6) separately for each of the three services and test each of these hypotheses for each of the three services.

We test the following hypotheses for Equation (7):

$$takeup_{ijk} = \alpha + \beta_1 email_{ijk} + \beta_2 text_{ijk} + \beta_3 emailincentive_{ijk} + \beta_4 textincentive_{ijk} + X'_{ij(-k)}\Omega + \gamma_j + \varepsilon_{ijk} \qquad (7)$$

1. $\beta_1 = 0,$
2. $\beta_2 = 0,$
3. $\beta_3 = 0,$
4. $\beta_4 = 0,$
5. $\beta_1 + \beta_3 = 0,$
6. $\beta_2 + \beta_4 = 0,$
7. $\beta_1 = \beta_3,$
8. $\beta_2 = \beta_4,$

9. $\beta_1 + \beta_3 = \beta_2 + \beta_4$,

10. $\beta_1 + \beta_3 = 0 \; and \; \beta_2 + \beta_4 = 0 \; (jointly)$.

The first hypothesis is that receiving any email (in the absence of an incentive) increases take-up. The second hypothesis is that receiving any text (in the absence of an incentive) increases take-up. The third hypothesis is that receiving an email incentive offer increases take-up above and beyond any effect of the non-incentive component of the message. The fourth hypothesis is that receiving a text incentive offer increases take-up above and beyond any effect of the non-incentive component of the message. The fifth hypothesis is that the combined effect of the email message and the incentive is greater than zero. The sixth hypothesis is that the combined effect of the text message and the incentive is greater than zero. The (null version of the) seventh hypothesis is that the effect of any email is equal to the effect of any text. The (null version of the) eighth hypothesis is that the effect of any email incentive is equal to the effect of any text incentive. The (null version of the) ninth hypothesis is that the combined effect of any email and the incentive is equal to the combined effect of any text and the incentive. The (null version of the) tenth hypothesis is that the combined effect of any email and the incentive is equal to zero and the combined effect of any text and the incentive is equal to zero. We estimate Equation (7) separately for each of the three services and test each of these hypotheses for each of the three services.

We test the following hypotheses from Equation (8):

$$takeup_{ijk} = \alpha + \beta_1 week3_{ijk} + \beta_2 week6_{ijk} + \beta_3 week9_{ijk} + X'_{ij(-k)}\,\Omega + \gamma_j + \varepsilon_{ijk} \qquad (8)$$

1. $\beta_1 = 0$,
2. $\beta_2 = 0$,
3. $\beta_3 = 0$,
4. $\beta_1 = \beta_2$,
5. $\beta_1 = \beta_3$,
6. $\beta_2 = \beta_3$

The first hypothesis is that receiving a message (of any form) in Week 3 increases take-up. The second hypothesis is that receiving a message (of any form) in Week 6 increases take-up. The third hypothesis is that receiving a message (of any form) in Week 9 increases take-up. The fourth hypothesis is that being assigned to receive any message in Week 3 has the same effect on take-up as being assigned to receive any message in Week 6. The fifth hypothesis is that being assigned to receive any message in Week 3 has the same effect on take-up as being assigned to receive any message in Week 9. The sixth hypothesis is that being assigned to receive any message in Week 6 has the same effect on take-up as being assigned to receive any message in Week 9.

We test the following hypotheses from Equation (9):

$$takeup_{ijk} = \alpha + \beta_1 onemessage_{ijk} + \beta_2 twomessages_{ijk} + \beta_3 threemessages_{ijk} + X'_{ij(-k)}\,\Omega + \gamma_j + \varepsilon_{ijk} \qquad (9)$$

1. $\beta_1 = 0$,
2. $\beta_2 = 0$,
3. $\beta_3 = 0$,
4. $\beta_1 + \beta_2 = 0$,
5. $\beta_1 + \beta_2 + \beta_3 = 0$

The first hypothesis is that receiving one message (of any form) increases take-up. The second hypothesis is that receiving two messages (of any form) increases take-up compared to receiving only one message. The third hypothesis is that receiving three messages (of any form) increases take-up compared to receiving only two messages. The fourth hypothesis is that being assigned to receive two messages has the same effect on take-up as being in the control group (zero messages). The fifth hypothesis is that being assigned to receive three messages has the same effect on take-up as being in the control group (zero messages).

Appendix B: student surveys

B.1: baseline survey

1. Have you declared a major?
    a. Yes
    b. No
    c. Unsure

2. What is your intended major?
    a. Economics
    b. Liberal Arts (other than Economics)
    c. Engineering
    d. Business
    e. Science
    f. Other

3. What is your intended minor?
    a. No intended minor
    b. Economics
    c. Other
    d. Unsure if I will take a minor

4. On a scale of 0-100, how likely are you to major in economics? (0=definitely not, 100=definitely yes, 50=50% chance, etc.)

5. On a scale of 0-100, how likely are you to minor in economics? (0=definitely not, 100=definitely yes, 50=50% chance, etc.)

6. Thinking about the Economics major, what is its single biggest appeal to you?
    a. Leads to a fulfilling career
    b. Leads to more future income
    c. Prestigious
    d. Increases chance of admission to graduate school
    e. Fun to study
    f. Other

7. Thinking about the Economics major, what is its single biggest drawback for you?
    a. Too difficult because of math
    b. Too difficult for reason besides math
    c. Too focused on making money
    d. Lack of diversity in faculty who teach Economics
    e. Lack of diversity in students who study Economics
    f. Too boring or not relevant to my life
    g. Other

8. Prior to this course, how many economics courses have you taken in college? (Include courses from previous terms only. Do not include this course or courses in which you are currently enrolled.)

9. What is your best guess of how much the median economics graduate earns in their first year after graduation?
    a. Less than $30,000
    b. $30,000-$35,000
    c. $35,001-$40,000
    d. $40,001-$45,000
    e. $45,001-$50,000
    f. More than $50,000

10. What is your best guess of how much the median economics graduate earns 15 years after graduation?
    a. Less than $50,000
    b. $50,000-$60,000
    c. $60,001-$70,000
    d. $70,001-$80,000
    e. $80,001-$90,000
    f. $90,001-$100,000
    g. More than $100,000

11. Before taking a midterm exam, do you usually know your current grade in a course?
    a. Yes, always
    b. Yes, but only if the professor reports it
    c. No
    d. Depends on the course

12. Before taking a midterm exam, do you usually know what score you need to achieve a final grade higher than your current grade?
    a. Yes, always
    b. Yes, but only if the professor reports it
    c. No
    d. Depends on the course

13. For a typical course, when do you usually begin studying for an exam?
    a. The night before
    b. The week before
    c. More than a week before

14. Suppose you encounter difficulties in a course. You try studying harder but still do not improve your performance. What is the next strategy you will take to improve your grade?
    a. Visit office hours
    b. Study with friends or classmates
    c. Find a peer tutor (if available for free)
    d. Seek academic coaching/study skills help
    e. None of these; I keep studying on my own

15. Suppose you encounter difficulties in a course. If free peer tutoring were available, how likely would you be to use it?
    a. Definitely wouldn’t use
    b. Unlikely to use
    c. Not sure if I would use or not

d. Might use
e. Definitely would use

16. Suppose you encounter difficulties in a course. If free academic coaching were available, how likely would you be to use it?
    a. Definitely wouldn't use
    b. Unlikely to use
    c. Not sure if I would use or not
    d. Might use
    e. Definitely would use

17. What is your expected grade in this course?
    a. A
    b. B
    c. C
    d. D
    e. F

Appendix B.2: endline survey

1. Have you declared a major?
    a. Yes
    b. No
    c. Unsure

2. What is your intended major?
    a. Economics
    b. Liberal Arts (other than Economics)
    c. Engineering
    d. Business
    e. Science
    f. Other

3. What is your intended minor?
    a. No intended minor
    b. Economics
    c. Other

4. On a scale of 0-100, how likely are you to major in economics? (0=definitely not, 100=definitely yes, 50=50% chance, etc.)

5. On a scale of 0-100, how likely are you to minor in economics? (0=definitely not, 100=definitely yes, 50=50% chance, etc.)

6. Thinking about the Economics major, what is its single biggest appeal to you?
    a. Leads to a fulfilling career
    b. Leads to more future income
    c. Prestigious
    d. Increases chance of admission to graduate school

    e. Fun to study
    f. Other

7. Thinking about the Economics major, what is its single biggest drawback for you?
    a. Too difficult because of math
    b. Too difficult for reason besides math
    c. Too focused on making money
    d. Lack of diversity in faculty who teach Economics
    e. Lack of diversity in students who study Economics
    f. Too boring or not relevant to my life
    g. Other

8. Before taking this survey, had you seen the video "A career in Economics…it's much more than you think"?
    a. Yes
    b. No

9. What is your best guess of how much the median economics graduate earns in their first year after graduation?
    a. Less than $30,000
    b. $30,000-$35,000
    c. $35,001-$40,000
    d. $40,001-$45,000
    e. $45,001-$50,000
    f. More than $50,000

10. What is your best guess of how much the median economics graduate earns 15 years after graduation?
    a. Less than $50,000
    b. $50,000-$60,000
    c. $60,001-$70,000
    d. $70,001-$80,000
    e. $80,001-$90,000
    f. $90,001-$100,000
    g. More than $100,000

11. Prior to taking this survey, were you aware of the following free services? (Check all that apply.)
    a. Extra practice problems on Canvas
    b. Office hours
    c. Economics Tutoring Lab
    d. Academic Coaching at the Academic Success Center

12. Think about times during this term when you have spent more time than usual on this course (for instance, studying for an exam). How did you add this extra time into your schedule? Choose the answer that accounts for the largest amount of your extra time, even if several answers are possible.
    a. Studied less in other classes
    b. Slept less
    c. Reduced time in other activities besides sleep (for instance, leisure time, time at a job, or hygiene/household maintenance)
    d. I didn't spend any extra time on this class

13. What is your expected grade in this course?
    a. A
    b. B
    c. C
    d. D
    e. F

Appendix C: structural estimation details

From Equation (10), the choice-specific value functions are:

$$V(1) = -c + \delta(1+\theta)\underline{y} + \varepsilon_1$$

$$= -c + \delta(1 + [b + m\tau_b])\underline{y} + \varepsilon_1 \quad \text{(A1)}$$

$$V(0) = \delta\underline{y} + \varepsilon_0 \quad \text{(A2)}$$

The net value of using an academic support service is therefore:

$$V(1) - V(0) = -c + \delta(b + m\tau_b)\underline{y} + (\varepsilon_1 - \varepsilon_0) \quad \text{(A3)}$$

The student uses the service if this net value in (A3) weakly exceeds zero, D=$\mathbf{1}[V(1) - V(0)] \geq 0)$. Under the assumption of Type I Extreme Value errors, the corresponding probability takes the logit form:

$$\Pr(D = 1) = \frac{\exp\left[-c+\delta(b+m\tau_b)\underline{y}\right]}{1+\exp\left[-c+\delta(b+m\tau_b)\underline{y}\right]} \quad \text{(A4)}$$

The corresponding log likelihood function is:

$$\ln L = \sum_{i=1}^{n} D_i \ln \Pr(D_i = 1) + (1 - D_i)\ln(1 - \Pr(D_i = 1)) \quad \text{(A5)}$$

To estimate the parameters of the likelihood function using (A5), we calibrate δ=0.95 and *b*. We use *b*=0.001, *b*=0.01, or *b*=0.1, depending on the specification. Grades $\underline{y}$ are equal to observed grades if *D*=0, or observed grades discounted by 1/(1+*b*) if *D*=0.

We replace *m* with *Z* in order to rely on the random variation induced by the experiment. We then choose *c* and τ_b to maximize the likelihood function (A5). The value of τ_b obtained in this procedure is analogous to the intent to treat (ITT), as it relies directly on variation in random assignment to advertising messages. We convert this estimate to a local average treatment effect (LATE) by scaling it by the share of compliers with the intervention, i.e., by dividing by Δ*m(Z).*

11 Tables

Table 1: Sample sizes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Study participation										
Term	Fall 2018	Winter 2019	Spring 2019	Total						
enrolled	879	803	805	2,487						
participated	765	673	681	2,119						
participation rate	0.87	0.84	0.85	0.85						
among participants:					Proportion					
provided mobile number	308	284	266	858	0.40					
male	499	424	449	1,372	0.65					
female	266	249	232	747	0.35					
microeconomics	518	394	350	1,262	0.60					
macroeconomics	247	279	331	857	0.40					
Panel B: Treatment status										
Timing & frequency of messages	control	Week 3	Week 6	Week 9	Weeks 3/6	Weeks 3/9	Weeks 6/9	Weeks 3/6/9	Total	Proportion
control group	556								556	0.26
message										
coaching		78	81	66	83	73	80	80	541	0.26
tutoring		75	85	73	73	81	72	71	530	0.25
practice problems		78	67	76	71	70	76	54	492	0.23
medium										
email		190	186	173	182	182	182	161	1,256	0.59
text		41	47	42	45	42	46	44	307	0.14
incentive										
none		113	126	109	113	109	119	119	808	0.38
lottery for $250		118	107	106	114	115	109	86	755	0.36
Total	556	231	233	215	227	224	228	205	2,119	1.00
Proportion	0.26	0.11	0.11	0.10	0.11	0.11	0.11	0.10	1.00	

Table shows sample sizes by indicated characteristics. Participated means student was at least 18 years old, consented to participate in study, and completed baseline survey.

Table 2(a): Baseline balance

treatment arm			message			medium (if mobile provided)		incentive	
	control	treated	coaching	tutoring	practice	email	text	none	lottery
Baseline variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
freshman	0.25	0.25	0.25	0.27	0.23	0.28	0.31	0.26	0.24
	[0.43]	[0.43]	[0.43]	[0.44]	[0.42]	[0.45]	[0.46]	[0.44]	[0.43]
sophomore	0.45	0.45	0.44	0.44	0.47	0.41	0.43	0.45	0.44
	[0.50]	[0.50]	[0.50]	[0.50]	[0.50]	[0.49]	[0.50]	[0.50]	[0.50]
other class year	0.30	0.30	0.31	0.29	0.30	0.31	0.26	0.28	0.32
	[0.46]	[0.46]	[0.47]	[0.45]	[0.46]	[0.46]	[0.44]	[0.45]	[0.47]
female	0.37	0.35	0.38	0.33	0.32	0.36	0.35	0.35	0.34
	[0.48]	[0.48]	[0.49]	[0.47]	[0.47]	[0.48]	[0.48]	[0.48]	[0.47]
First generation	0.11	0.11	0.10	0.11	0.10	0.08	0.10	0.11	0.10
	[0.31]	[0.31]	[0.31]	[0.32]	[0.30]	[0.26]	[0.30]	[0.31]	[0.30]
white	0.56	0.60	0.59	0.62	0.59	0.57	0.55	0.60	0.60
	[0.50]	[0.49]	[0.49]	[0.49]	[0.49]	[0.50]	[0.50]	[0.49]	[0.49]
Asian/Pacific Islander	0.08	0.06	0.07	0.07	0.06	0.04	0.07	0.06	0.07
	[0.27]	[0.25]	[0.25]	[0.25]	[0.24]	[0.19]	[0.25]	[0.23]	[0.26]
Hispanic	0.08	0.09	0.09	0.08	0.09	0.08	0.08	0.10	0.08
	[0.28]	[0.28]	[0.29]	[0.27]	[0.29]	[0.27]	[0.27]	[0.29]	[0.27]
other race	0.28	0.25	0.26	0.24	0.26	0.31	0.30	0.25	0.25
	[0.45]	[0.44]	[0.44]	[0.43]	[0.44]	[0.46]	[0.46]	[0.44]	[0.44]
high school GPA	3.48	3.47	3.46	3.45	3.49	3.45	3.46	3.45	3.48
	[0.37]	[0.41]	[0.44]	[0.39]	[0.38]	[0.37]	[0.38]	[0.41]	[0.40]
GPA at Oregon State	3.07	3.07	3.07	3.05	3.10	3.05	3.01	3.05	3.10
	[0.57]	[0.55]	[0.53]	[0.57]	[0.55]	[0.56]	[0.56]	[0.58]	[0.52]
N	556	1,563	541	530	492	334	307	808	755
F-test (p-value)		0.50	0.61	0.56	0.39	0.29	0.96	0.33	0.50

Table shows means of baseline characteristic by treatment arm. Standard deviations in brackets. Message medium restricted to sample providing mobile phone number. Final row reports results of regression of treatment dummy from each column on all baseline characteristics (p-value from F-test of all baseline characteristics reported). Regressions include dummies for randomization strata and missing GPA variables (GPA imputed to zero if missing).

Table 2(b): Baseline balance

treatment arm	frequency & timing						
	Week 3	Week 6	Week 9	Weeks 3/6	Weeks 3/9	Weeks 6/9	Weeks 3/6/9
Baseline variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
freshman	0.29	0.24	0.27	0.25	0.23	0.25	0.24
	[0.45]	[0.43]	[0.44]	[0.43]	[0.42]	[0.43]	[0.43]
sophomore	0.44	0.44	0.45	0.45	0.44	0.48	0.44
	[0.50]	[0.50]	[0.50]	[0.50]	[0.50]	[0.50]	[0.50]
other class year	0.28	0.32	0.28	0.30	0.33	0.28	0.32
	[0.45]	[0.47]	[0.45]	[0.46]	[0.47]	[0.45]	[0.47]
female	0.33	0.39	0.37	0.35	0.32	0.36	0.29
	[0.47]	[0.49]	[0.48]	[0.48]	[0.47]	[0.48]	[0.46]
First generation	0.11	0.13	0.08	0.12	0.09	0.11	0.09
	[0.31]	[0.34]	[0.28]	[0.32]	[0.29]	[0.31]	[0.29]
white	0.61	0.65	0.58	0.58	0.56	0.61	0.58
	[0.49]	[0.48]	[0.49]	[0.50]	[0.50]	[0.49]	[0.50]
Asian/Pacific Islander	0.04	0.05	0.08	0.06	0.08	0.08	0.06
	[0.19]	[0.22]	[0.27]	[0.23]	[0.27]	[0.28]	[0.24]
hispanic	0.09	0.08	0.05	0.12	0.09	0.10	0.08
	[0.29]	[0.27]	[0.22]	[0.32]	[0.29]	[0.30]	[0.28]
other race	0.26	0.22	0.29	0.25	0.27	0.21	0.27
	[0.44]	[0.41]	[0.45]	[0.44]	[0.45]	[0.41]	[0.45]
high school GPA	3.52	3.51	3.47	3.41	3.46	3.48	3.43
	[0.36]	[0.39]	[0.38]	[0.49]	[0.44]	[0.38]	[0.38]
GPA at Oregon State	3.10	3.09	3.10	3.05	3.05	3.09	3.03
	[0.55]	[0.58]	[0.57]	[0.57]	[0.57]	[0.52]	[0.50]
N	231	233	215	227	224	228	205
F-test (p-value)	0.22	0.15	0.12	0.32	0.83	0.48	0.46

Table shows means of baseline characteristic by treatment arm. Standard deviations in brackets. Final row reports results of regression of treatment dummy from each column on all baseline characteristics (p-value from F-test of all baseline characteristics reported). Regressions include dummies for randomization strata and missing GPA variables (GPA imputed to zero if missing).

Table 3: Message takeup and service awareness

	opened email	clicked link	awareness: coaching	awareness: tutoring	awareness: practice	awareness: student-service panel
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: message, incentive						
message	0.662	0.005	0.007	0.056	0.065	0.051
	(0.018)***	(0.003)*	(0.040)	(0.034)	(0.034)*	(0.018)***
message*incentive	0.023	0.007	0.020	0.011	0.005	-0.003
	(0.026)	(0.005)	(0.047)	(0.040)	(0.040)	(0.024)
N	1,812	1,812	1,838	1,838	1,838	5,514
control mean	0.00	0.00	0.54	0.70	0.72	0.65
Panel B: message medium						
email			0.119	0.119	0.133	0.116
			(0.086)	(0.071)*	(0.072)*	(0.041)***
text			-0.092	0.080	-0.091	-0.041
			(0.080)	(0.080)	(0.083)	(0.042)
email, with incentive			-0.066	0.045	-0.007	-0.035
			(0.106)	(0.085)	(0.087)	(0.052)
text, with incentive			0.015	0.054	0.200	0.058
			(0.108)	(0.096)	(0.097)**	(0.057)
N			737	737	737	2,211
control mean			0.47	0.61	0.70	0.59
Panel C: message timing						
Week 3	0.406	0.003	0.028	0.067	-0.052	0.013
	(0.022)***	(0.004)	(0.039)	(0.034)**	(0.034)	(0.020)
Week 6	0.183	0.011	0.012	-0.008	0.090	0.033
	(0.022)***	(0.004)***	(0.038)	(0.034)	(0.032)***	(0.019)*
Week 9	0.324	0.003	-0.005	0.047	0.095	0.048
	(0.022)***	(0.004)	(0.040)	(0.033)	(0.032)***	(0.020)**
N	1,812	1,812	1,838	1,838	1,838	5,514
control mean	0.00	0.00	0.54	0.70	0.72	0.65
Panel D: message frequency						
at least one message	0.587	0.004	-0.016	0.041	0.053	0.024
	(0.021)***	(0.003)	(0.043)	(0.036)	(0.037)	(0.019)
at least two messages	0.176	0.006	0.055	0.029	0.001	0.035
	(0.028)***	(0.005)	(0.050)	(0.043)	(0.043)	(0.026)
three messages	-0.001	0.009	-0.001	0.037	0.137	0.044
	(0.038)	(0.011)	(0.070)	(0.060)	(0.053)***	(0.036)
N	1,812	1,812	1,838	1,838	1,838	5,514
control mean	0.00	0.00	0.54	0.70	0.72	0.65

Sample is all participants in study, academic year 2018-2019. Exception is Panel B, which includes only those eligible for text message treatment. Data on email outcomes is administrative. Awarness outcomes are self-reports from endline survey. Regressions for opened email and clicked link (columns 1-2) exclude those randomly assigned to receive text messages. Regressions for service-specific awareness (columns 3-5) report coefficient on message specific to that service. For example, the table reports the coefficient on the tutoring message when tutoring is the outcome. Variables defined additively, so that in Panel C a student assigned to receive a message in Weeks 3, 6, and 9 would have all of the week indicator variables equal to one. In Panel D, a student assigned two messages would have the "one message" and "two message" indicators equal to one, while a student assigned to three messages would have all message frequency indicators set to one. Regressions in columns 3-5 also control for timing or frequency of messages for other services in Panels C-D, respectively. Student x service panel regressions in column 6 stacks all student-by-section-by-service observations to form a panel. Student x service regressions include service fixed effects and student-section random effects. Robust standard errors in parentheses. All regressions include dummies for randomization strata. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 4: Academic service use and advertising messages

	any service use				multiple service use			
	coaching	tutoring	practice	student x service panel	coaching	tutoring	practice	student x service panel
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
message	0.002	-0.015	0.016	0.007	-0.002	-0.006	0.034	0.012
	(0.009)	(0.016)	(0.016)	(0.007)	(0.002)	(0.010)	(0.020)*	(0.006)*
message*incentive	-0.003	0.006	-0.006	-0.003	0.000	0.012	-0.021	-0.004
	(0.010)	(0.019)	(0.019)	(0.009)	(0.000)	(0.013)	(0.024)	(0.009)
N	2,119	2,119	2,119	6,357	2,119	2,119	2,119	6,357
control mean	0.013	0.059	0.939	0.340	0.002	0.022	0.903	0.310
message+incentive	-0.001	-0.009	0.010	0.000	-0.002	0.006	0.013	0.010
se(message+incentive)	(0.008)	(0.017)	(0.018)	(0.010)	(0.002)	(0.012)	(0.023)	(0.010)
p-value	0.90	0.58	0.57	0.57	0.31	0.62	0.58	0.30

Sample is all participants in study, AY2018-2019. Treatment corresponds to service measured as outcome, i.e., message refers to coaching message when coaching is the outcome. Regressions in columns 1-3, 5-7 also control for messages and message*incentive for other services. Student x service panel regressions in columns 4/8 stack all student-by-section-by-service observations to form a panel. Student x service regressions include service fixed effects and student-section random effects. All regressions also include dummies for randomization strata. Robust standard errors in parentheses. Message + incentive reports sum of coefficients, with standard error and p-value below. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 5: Academic service use and message medium

	any service use				multiple service use			
	coaching	tutoring	practice	student x service panel	coaching	tutoring	practice	student x service panel
message type	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
email	0.006	0.009	0.067	0.029	0.001	0.016	0.110	0.038
	(0.021)	(0.038)	(0.018)***	(0.014)**	(0.001)	(0.025)	(0.024)***	(0.010)***
text	-0.009	-0.001	-0.019	-0.001	-0.001	0.003	0.032	0.013
	(0.008)	(0.040)	(0.040)	(0.017)	(0.001)	(0.024)	(0.042)	(0.015)
email, with incentive	-0.012	0.047	-0.053	-0.007	-0.001	0.019	-0.089	-0.022
	(0.019)	(0.055)	(0.034)	(0.022)	(0.002)	(0.038)	(0.044)**	(0.018)
text, with incentive	0.016	-0.044	-0.021	-0.024	0.000	-0.013	-0.053	-0.024
	(0.020)	(0.037)	(0.055)	(0.023)	(0.001)	(0.022)	(0.059)	(0.021)
N	858	858	858	2,574	858	858	858	2,574
control mean	0.009	0.060	0.935	0.330	0.000	0.018	0.894	0.300
H5: email + incentive	0.32	0.22	0.71	0.21	0.61	0.26	0.65	0.32
H6: text + incentive	0.75	0.01	0.40	0.12	0.59	0.23	0.69	0.46
H7: email = text	0.42	0.85	0.02	0.16	0.38	0.97	0.00	0.14
H8: email incentive = text incentive	0.29	0.17	0.63	0.59	0.85	0.73	0.34	0.93
H9: email + incentive = text + incentive	0.51	0.02	0.33	0.04	0.91	0.12	0.50	0.20
H10: email + incentive = 0, text + incentive = 0	0.47	0.00	0.61	0.12	0.82	0.13	0.80	0.44

Sample is all participants who provided mobile phone number, AY2018-2019. Treatment corresponds to service measured as outcome, i.e., email/text refers to coaching email/text when coaching is the outcome. Regressions in columns 1-3, 5-7 also control for email/text and email/text*incentive for other services. Student x service panel regressions in columns 4/8 stack all student-by-section-by-service observations to form a panel. Student x service regressions include service fixed effects and student-section random effects. All regressions also include dummies for randomization strata. Robust standard errors in parentheses. p-values of hypothesis tests reported at bottom of table. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 6: Academic service use, message timing, and message frequency

	any service use				multiple service use			
	coaching	tutoring	practice	student x service panel	coaching	tutoring	practice	student x service panel
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: message timing								
Week 3	0.001	-0.024	-0.016	-0.010	-0.001	-0.001	0.011	0.004
	(0.009)	(0.016)	(0.017)	(0.008)	(0.001)	(0.011)	(0.019)	(0.007)
Week 6	0.000	-0.005	0.006	0.004	-0.001	-0.006	-0.009	-0.003
	(0.008)	(0.016)	(0.016)	(0.008)	(0.001)	(0.010)	(0.021)	(0.007)
Week 9	-0.005	0.001	0.029	0.011	-0.001	0.007	0.041	0.016
	(0.008)	(0.016)	(0.015)*	(0.008)	(0.001)	(0.011)	(0.019)**	(0.007)**
N	2,119	2,119	2,119	6,357	2,119	2,119	2,119	6,357
control mean	0.013	0.059	0.939	0.340	0.002	0.022	0.903	0.310
H1: Week 3 = Week 6	0.96	0.50	0.34	0.27	0.90	0.78	0.50	0.53
H2: Week 3 = Week 9	0.71	0.36	0.08	0.09	0.83	0.65	0.33	0.29
H3: Week 6 = Week 9	0.69	0.78	0.38	0.53	0.91	0.39	0.12	0.09
Panel B: message frequency								
at least one message	0.000	0.001	0.005	0.008	-0.002	-0.005	0.003	0.002
	(0.009)	(0.019)	(0.019)	(0.008)	(0.002)	(0.011)	(0.024)	(0.008)
at least two messages	0.004	-0.016	0.018	0.001	0.000	0.015	0.040	0.018
	(0.012)	(0.021)	(0.020)	(0.010)	(0.000)	(0.015)	(0.025)	(0.009)*
three messages	-0.017	-0.026	-0.015	-0.021	0.000	-0.029	-0.015	-0.017
	(0.009)*	(0.020)	(0.035)	(0.012)*	(0.000)	(0.012)**	(0.039)	(0.011)
N	2,119	2,119	2,119	6,357	2,119	2,119	2,119	6,357
control mean	0.013	0.059	0.939	0.340	0.002	0.022	0.903	0.310
H4: at least two messages (total effect)	0.64	0.36	0.16	0.24	0.31	0.43	0.03	0.00
H5: three messages (total effect)	0.01	0.02	0.81	0.23	0.35	0.00	0.45	0.80

Sample is all participants, AY2018-2019. Treatment corresponds to service measured as outcome, i.e., in Panel A, Week 3 refers to coaching message in Week 3 when coaching is the outcome. In Panel B, at least one message refers to at least one coaching message when coaching is the outcome. Variables defined additively, so that in Panel A a student assigned to receive a message in Weeks 3, 6, and 9 would have all of the week indicator variables equal to one. In Panel B, a student assigned two messages would have the "one message" and "two message" indicators equal to one, while a student assigned to three messages would have all message frequency indicators set to one. Regressions in columns 1-3, 5-7 also control for timing (Panel A) or frequency (Panel B) of messages for other services. Student x service panel regressions in columns 4/8 stack all student-by-section-by-service observations to form a panel. Student x service regressions include service fixed effects and student-section random effects. All regressions also include dummies for randomization strata. Robust standard errors in parentheses. p-values of hypothesis tests reported at bottom of table. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 7: Heterogeneous treatment effects

	any service use						multiple service use					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
message	0.008	-0.001	0.004	-0.001	0.008	0.000	0.011	0.002	0.010	0.004	0.013	0.005
	(0.006)	(0.011)	(0.011)	(0.008)	(0.006)	(0.012)	(0.005)**	(0.010)	(0.009)	(0.007)	(0.006)**	(0.011)
message*freshman		0.022						0.019				
		(0.016)						(0.015)				
message*sophomore		0.009						0.010				
		(0.014)						(0.013)				
message*male			0.006						0.002			
			(0.013)						(0.011)			
message*GPA below median				0.018						0.016		
				(0.012)						(0.011)		
message*first generation					-0.004						-0.022	
					(0.020)						(0.020)	
message*low service demand						0.011						0.009
						(0.014)						(0.013)
N	6,357	6,357	6,357	6,357	6,357	6,357	6,357	6,357	6,357	6,357	6,357	6,357
control mean	0.34	0.34	0.34	0.34	0.34	0.34	0.31	0.31	0.31	0.31	0.31	0.31
p-value(message+interaction 1)		0.07	0.14	0.04	0.82	0.08		0.06	0.08	0.02	0.67	0.03
p-value(message+interaction 2)		0.42						0.14				

Sample is all participants in study, AY2018-2019. Unit of analysis is student-section-service panel. All regressions include student-section, randomization strata, and service fixed effects. All interactions measured at baseline. Low service demand is at/below median demand for tutoring and academic coaching, from baseline survey. Robust standard errors in parentheses. p-values of test of linear below. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 8: Time use and advertising messages

	reduced time spent:			no extra time in Economics
	studying other subjects	sleeping	other activity	
	(1)	(2)	(3)	(4)
Panel A: incentive				
message	-0.006	0.006	-0.002	0.001
	(0.022)	(0.026)	(0.030)	(0.016)
message*incentive	-0.008	-0.023	0.041	-0.009
	(0.020)	(0.024)	(0.027)	(0.015)
N	1,838	1,838	1,838	1,838
control mean	0.17	0.26	0.50	0.08
Panel B: message medium				
email	-0.033	-0.051	0.045	0.038
	(0.040)	(0.049)	(0.057)	(0.031)
text	0.021	-0.053	0.048	-0.023
	(0.044)	(0.050)	(0.058)	(0.026)
email, with incentive	-0.019	0.048	0.019	-0.046
	(0.040)	(0.053)	(0.061)	(0.031)
text, with incentive	-0.006	-0.018	-0.016	0.048
	(0.048)	(0.052)	(0.064)	(0.031)
N	737	737	737	737
control mean	0.16	0.27	0.51	0.06
Panel C: message timing				
Week 3	-0.011	0.030	-0.028	0.010
	(0.017)	(0.021)	(0.024)	(0.013)
Week 6	-0.004	0.006	0.011	-0.012
	(0.017)	(0.021)	(0.024)	(0.013)
Week 9	0.008	-0.018	0.020	-0.011
	(0.018)	(0.021)	(0.024)	(0.013)
N	1,838	1,838	1,838	1,838
control mean	0.17	0.26	0.50	0.08
Panel D: message frequency				
at least one message	-0.006	-0.025	0.033	-0.004
	(0.023)	(0.026)	(0.031)	(0.017)
at least two messages	-0.016	0.037	-0.029	0.009
	(0.021)	(0.025)	(0.029)	(0.016)
three messages	0.036	-0.010	0.011	-0.037
	(0.034)	(0.037)	(0.044)	(0.021)*
N	1,838	1,838	1,838	1,838
control mean	0.17	0.26	0.50	0.08

Sample is all participants in study, academic year 2018-2019. Exception is Panel B, which includes only those eligible for text message treatment. Outcomes are self-reports from endline survey. Panel A reports treatment effects for any message, regardless of academic support service encouraged in message. Variables defined additively, so that in Panel C a student assigned to receive a message in Weeks 3, 6, and 9 would have all of the week indicator variables equal to one. In Panel D, a student assigned two messages would have the "one message" and "two message" indicators equal to one, while a student assigned to three messages would have all message frequency indicators set to one. Robust standard errors in parentheses. All regressions include dummies for randomization strata. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 9: Attention

service category		coaching (1)	tutoring (2)	extra practice (3)	any service (4)
always attentive	*Pr(M=1|Z=0)*	0.55	0.69	0.71	0.65
		(0.01)	(0.01)	(0.01)	(0.01)
never attentive	*Pr(M=0|Z=1)*	0.45	0.25	0.22	0.30
		(0.02)	(0.02)	(0.02)	(0.01)
compliers	*Pr(M=1|Z=1) -*	0.01	0.07	0.07	0.05
	Pr(M=1|Z=0)	(0.03)	(0.02)	(0.02)	(0.02)

Table shows proportion of each type, by academic service. Sample is all students who completed endline survey. Always attentive is defined as a student indicating awareness of the service in endline survey, conditional on being in the control group for receiving an advertising message for that service. Never attentive is defined as a student not indicating awareness of service in endline survey, conditional on being in the treatment group for advertising messages for that service. Always/never attentive proportions and standard error obtained from regressing always/never attentive dummy, demeaned by randomization strata, on a constant. Proportion compliers and standard error from regression of awareness of service on assignment to treatment for that service, controlling for randomization strata. Regressions for "any service" (column 4) stack observations into a student-service panel.

Table 10: Structural parameter estimates

	coaching	tutoring	practice problems	all
	(1)	(2)	(3)	(4)
Panel A: *b*=0.001				
-*c*	-4.31	-3.00	2.78	-4.43
	(0.22)***	(0.12)***	(0.11)***	(0.21)***
τ_b				
ITT	-0.13	0.05	0.23	0.09
	(0.21)	(0.09)	(0.12)**	(0.06)
LATE	-9.09	0.70	3.14	2.03
	(14.83)	(1.31)	(1.58)	(1.34)
Panel B: *b*=0.01				
-*c*	-4.33	-3.01	2.76	-4.45
	(0.22)***	(0.12)***	(0.11)***	(0.21)***
τ_b				
ITT	-0.13	0.04	0.23	0.09
	(0.21)	(0.09)	(0.12)*	(0.06)
LATE	-9.46	0.62	3.09	1.92
	(14.90)	(1.31)	(1.59)	(1.35)
Panel C: *b*=0.1				
-*c*	-4.47	-3.15	2.61	-4.58
	(0.22)***	(0.12)***	(0.11)***	(0.21)***
τ_b				
ITT	-0.18	-0.01	0.19	0.04
	(0.22)	(0.10)	(0.12)	(0.07)
LATE	-13.03	-0.07	2.54	0.90
	(15.53)	(1.37)	(1.67)	(1.41)
N	2,009	2,009	2,009	6,027
Pr(complier)	0.01	0.07	0.07	0.05

Table reports estimates of parameters -*c*, τ_b from structural model of academic service use. ITT reports estimated coefficient on *Z*=1(assigned message for service). LATE reports ITT/Pr(complier), where Pr(complier) reported at bottom of table. Column (4) stacks all services into a student x service panel, with service fixed effects included in estimation. Discount factor δ=0.95. Standard errors in parenthesis.

12 Figures

Figure 1(a): Coaching email, with incentive

Jon Chesbro <jon.chesbro@oregonstate.edu> Thu, Jan 24, 2019 at 3:00 PM
To: todd.pugatch@oregonstate.edu

Having trouble reading this? To view this email as a web page, click here.

Oregon State University

Hi Todd,

I wanted to share important information about a free service available to refine your study strategies.

Academic Coaches are trained to improve your academic performance. Your coach will work with you to meet your goals. Coaching at the Academic Success Center (ASC) is free.

Watch a short video on Academic Coaching here.

Next step:

- Come to the ASC main office at 125 Waldo to schedule a coaching appointment or to speak with an ASC Strategist about your academic success.

or

- Set up a coaching appointment online.

Prizes!
Meet with your Academic Coach at the ASC within two weeks of this message and you will be entered into a lottery to win $250 Orange Cash.

Sincerely,
Jon Chesbro
Instructor, Economics

Figure 1(b): Tutoring email, with incentive

ECON 201: Free Economics tutoring at the Economics Undergraduate Lab
2 messages

Jon Chesbro <jon.chesbro@oregonstate.edu> Mon, Mar 4, 2019 at 3:00 PM
To: todd.pugatch@oregonstate.edu

Having trouble reading this? To view this email as a web page, click here.

Oregon State University

Hi Todd,

I wanted to share important information about a free service available to refine your study strategies.

The Economics Undergraduate Lab offers free tutoring. Your tutor will work with you to meet your goals. Tutoring at the Economics Undergraduate Lab is free.

Next step:

- Come to the Economics Undergraduate Lab at Bexell Hall, Room 100H, Monday-Thursday 10am-5pm, Friday 10am-4pm.

Note that there is no tutoring during finals week!

Prizes!
Visit the Economics Undergraduate Lab within two weeks of this message and you will be entered into a lottery to win $250 Orange Cash.

Sincerely,
Jon Chesbro
Instructor, Economics

Figure 1(c): Practice email, with incentive

ECON 201: Now available - practice problems on Canvas

2 messages

Jon Chesbro <jon.chesbro@oregonstate.edu> Thu, Jan 24, 2019 at 3:00 PM

To: todd.pugatch@oregonstate.edu

Having trouble reading this? To view this email as a web page, click here.

Oregon State University

Hi Todd,

Completing extra practice problems can be one of the most effective ways to prepare for your Economics exams.

Extra practice problems for all chapters are available on Canvas.

Next step:

- Access practice problems on Canvas. Go to Quizzes → Practice for (Midterm/Final) Exam.

Prizes!

Attempt practice problems within two weeks of this message and you will be entered into a lottery to win $250 Orange Cash.

Sincerely,
Jon Chesbro
Instructor, Economics

Figure 2(a): Coaching text, with incentive

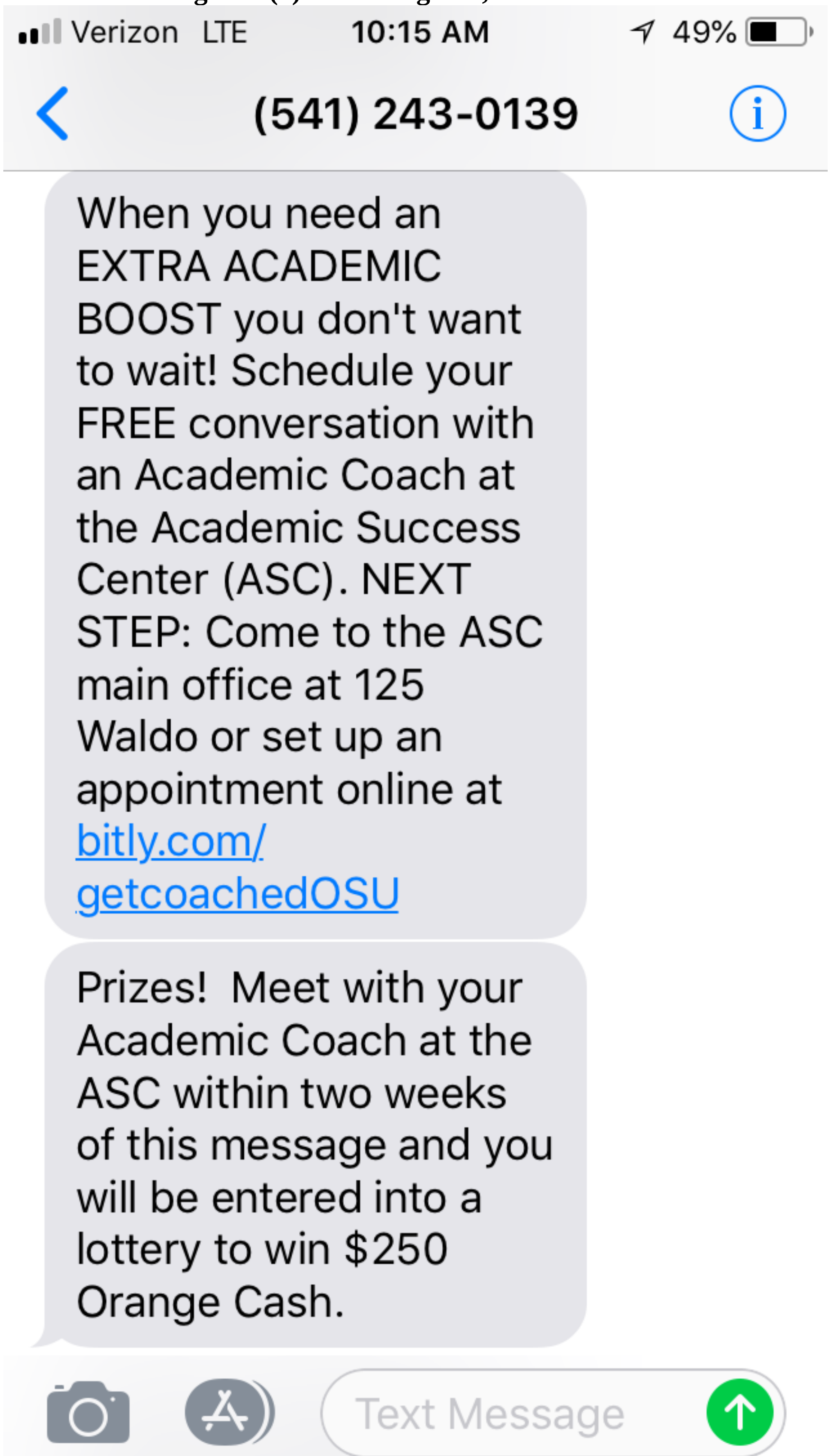

Figure 2(b): Tutoring text, with incentive

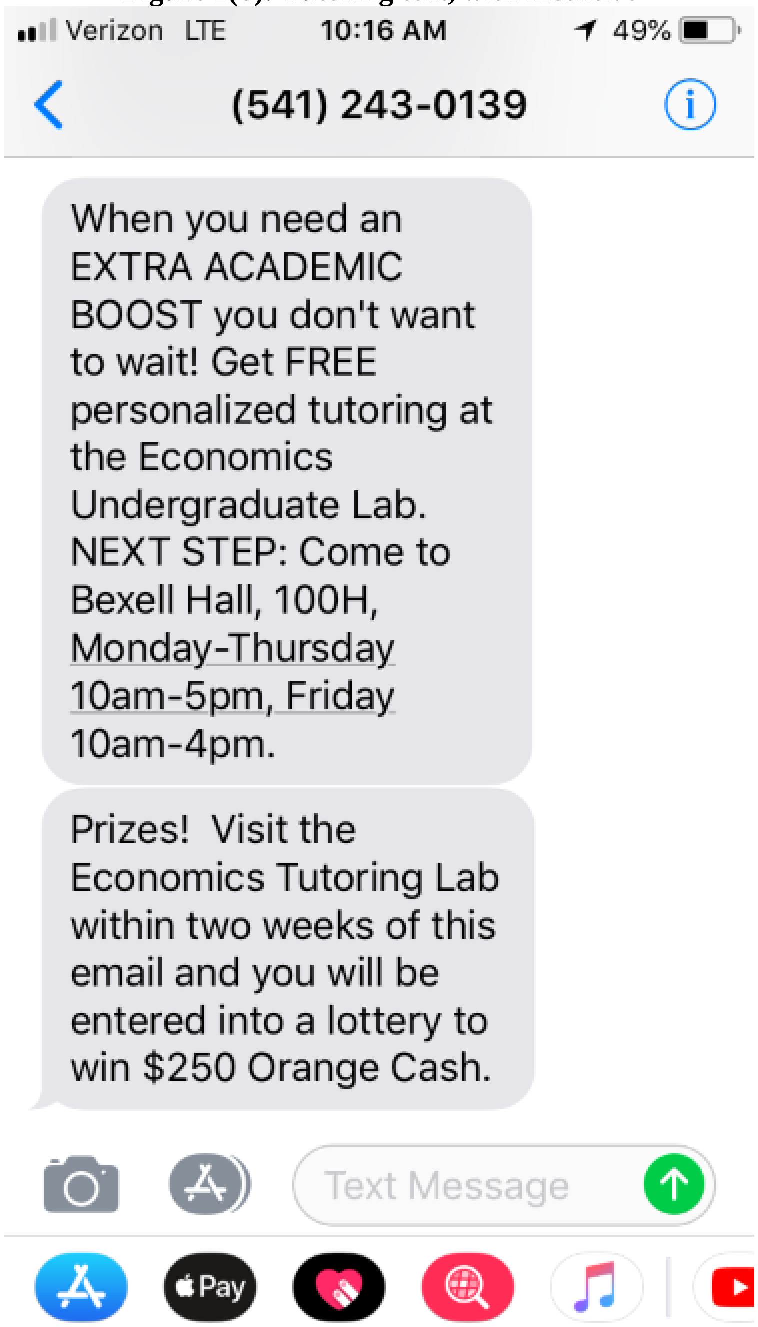

Figure 2(c): Practice text, with incentive