Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 26-1409

Scaling student support with conversational artificial intelligence

Catherine Mata
Brown University

Emily Russell
California State University, Northridge

Lindsay C. Page
Brown University

AI-enabled chatbots are increasingly used to support student success, yet evidence on their long-term sustainability and impacts remains limited. We examine the implementation of an AI-enabled text-messaging chatbot at a large, urban public university. Drawing on system observation, discussions with administrators, and a four-year randomized controlled trial, we assess institutional conditions for sustained implementation, student receptivity over time, and effects on academic outcomes. We find that centralized ownership and flexible communication are critical for sustainability. Students remain receptive over time, and impacts are concentrated in improved completion of time-sensitive administrative tasks, with no detectable effects on academic performance or persistence.

VERSION: February 2026

Suggested citation: Mata, Catherine, Emily Russell, and Lindsay C. Page. (2026). Scaling student support with conversational artificial intelligence. (EdWorkingPaper: 26-1409). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/b0dp-sn77

Scaling student support with conversational artificial intelligence

Catherine Mata[*]
Brown University

Emily Russell
California State University, Northridge

Lindsay C. Page
Brown University

January 2026

Abstract

AI-enabled chatbots are increasingly used to support student success, yet evidence on their long-term sustainability and impacts remains limited. We examine the implementation of an AI-enabled text-messaging chatbot at a large, urban public university. Drawing on system observation, discussions with administrators, and a four-year randomized controlled trial, we assess institutional conditions for sustained implementation, student receptivity over time, and effects on academic outcomes. We find that centralized ownership and flexible communication are critical for sustainability. Students remain receptive over time, and impacts are concentrated in improved completion of time-sensitive administrative tasks, with no detectable effects on academic performance or persistence.

JEL Classification: I21, I23, O33
Keywords: college success, student support, higher education, academic chatbot, artificial intelligence.

[*] Mata: Annenberg Institute for School Reform at Brown University, 164 Angell St., 2nd floor, Providence, RI 02906. Email: catherine_mata@brown.edu. Russell: emily.russell@csun.edu. Page: lindsay_page@brown.edu. This research was supported through a grant from the Kresge Foundation. We gratefully acknowledge California State University, Northridge for their support and engagement with this study. All errors are our own.

I. INTRODUCTION

Artificial Intelligence (AI) is rapidly being woven into many areas of modern society. In education, AI-based technologies entail both opportunities and risks for students' learning (Hopelab et al., 2024). On one hand, expansive reliance on AI tools raises concerns about the erosion of students' creative and analytical skills, which may not fully develop if these tools are frequently used from an early age (Doss et al., 2025; Machidon, 2025). At later stages, educators face new challenges in designing assignments and evaluative assessments that cannot readily be completed using generative AI (UNESCO, 2025; Wang et al., 2025). On the other hand, new technologies present opportunities for students and educators to offload repetitive tasks to AI tools, allowing them to dedicate time and attention to more critical work (Doss et al., 2025; Mah et al., 2025; Baek et al., 2024; Tate et al., 2024; Ferlazzo, 2023).

As AI becomes increasingly prevalent, responses within the education sector have varied. Some institutions have prohibited its use (Elsen-Rooney, 2023), while others are exploring the potential value of AI tools designed specifically to support educators and students' academic success. For example, AI has been integrated into tools that automate students' feedback to educators, enabling rapid incorporation of student input into classroom practice (Demszky et al., 2023); interactive, chat-based teacher-training tools that allow novice teachers to practice with simulated students and reduce the potential harm that teachers-in-training might cause in real classrooms (Markel et al., 2023); and supplementary supports for tutors to use during live tutoring sessions, curating expert-like guidance with the goal of reducing inequities in access to highly experienced tutors among underserved communities (Wang et al., 2024). In higher education, university student success offices are integrating AI-enabled tools into systems of support for students (EAB, 2024). For instance, some universities have successfully employed AI-enabled chatbots to communicate with students via text message about required pre-matriculation tasks to mitigate “summer melt,” whereby college-intending students fail to transition to postsecondary education (Page & Gehlbach, 2017; Nurshatayeva et al., 2021). Universities are also deploying these chatbots to help enrolled students navigate college administrative tasks (Page et al., 2025) and course-specific academic tasks (Meyer et al., 2024).

The evidence on the efficacy of such chatbot communication comes primarily from intervention studies conducted over a summer, semester or single academic year timeframe with outcomes measured during or shortly after the experimental period concludes. In contrast, little research has tracked the use and impact of AI-enabled chatbot communication over a longer timeframe. Aiming to close this evidence gap, we focus in this paper on the following research questions: (1) What does it take for an institution to sustain centralized chatbot communication over time? (2) To what extent are students receptive to chatbot communication over time? (3) What impact does chatbot communication have on student outcomes throughout their time in college?

To answer these questions, we rely on evidence from system observation, discussions with program administrators, and a randomized controlled trial that follows two cohorts of college students over four years of undergraduate study at California State University, Northridge (CSUN). We document CSUN’s use and adaptation of an AI-enabled text-messaging system for student communication over a multi-year period.

CSUN adopted the AI-enabled chatbot technology with the goal of improving students’ college persistence and completion through two potential channels. First, CSUN aimed to improve students’ awareness and navigation of financial, logistical, and administrative tasks, such as FAFSA renewal, timely course registration, and maintaining satisfactory academic progress.

Second, CSUN implemented outreach and communication aiming to foster in students a greater sense of social connectedness and belonging on campus.

To preview our key findings, based on system observation and discussions with program administrators, we conclude that a central office overseeing this communication tool is essential not only for enacting and sustaining the intended goals of chatbot-based communication but also for coordinating data flow across various university offices that hold the necessary information for the chatbot to effectively serve all students. Additionally, a certain level of flexibility within the overarching communication strategy is crucial. Such flexibility allows university administrators to adapt the tool to students' evolving needs and changing circumstances, as seen during the COVID-19 pandemic. It also helps sustain student engagement over time, both as new cohorts enter and as existing cohorts become accustomed to the tool.

From the analysis of CSUN's administrative and chatbot system records, we find that students are receptive to chatbot communication over the course of their undergraduate careers, as evidenced by their strong response rates to interactive campaigns and low rates of opting out of receiving chatbot communication. The majority of students choose to stay connected to campus communication through the chatbot. We interpret this as students finding value in the tool even when they do not always interact with it.

Finally, consistent with evidence from other universities (Page et al., 2025), our impact evaluation shows that the most tangible impact of this communication is students' improved completion of time-sensitive administrative tasks required for continued enrollment. In contrast, communication from this tool, on its own, did not change students' academic performance or persistence outcomes.

II. SETTING AND INTERVENTION

CSUN is a large, urban public institution that serves a diverse student body. Located in Los Angeles, CA, CSUN recently had the third-largest total student body (36,848) across all campuses in the California State University system (CSU, 2024). Among the students enrolled in Fall 2024, 34% were first-generation college-goers, 53% were female, 56% were Hispanic, 5% were African American, and 9% were Asian American. In recent years, about 50% of all CSUN students have received Pell Grants (CSU, 2024).

Based on prior evidence regarding the efficacy of chatbot communication for improving college student enrollment and navigation of required administrative tasks (Page & Gehlbach, 2017; Page et al., 2025), in 2018, CSUN's then-Associate Vice President of Undergraduate Studies engaged the software company Mainstay to develop a university-specific, AI-enabled text message chatbot to communicate with CSUN's undergraduates. Since then, a team of CSUN staff and faculty from multiple divisions and departments have collaboratively overseen the chatbot ("chatbot team" herein). The university aims for the chatbot, named CSUNny, to represent a friendly embodiment of the university administration.

The chatbot team plans and distributes CSUNny messages (called campaigns) from a web-based platform. CSUNny campaigns serve two primary purposes: to provide students with proactive outreach and guidance to navigate the required financial and administrative aspects of their university experience and to promote in students a sense of belonging within the university community. We have classified CSUNny campaigns into six domains, according to their focal topics, as follows:

(1) *Academic*: messages pertaining to services offered or tasks required for academic success, such as enrollment in Early Start programs or tutoring services.[1]

(2) *Financial*: alerts about upcoming financial due dates, such as FAFSA filing deadlines and billing due dates, as well as reminders about the availability of scholarships and other funds.

(3) *Administrative*: reminders about required administrative tasks. Prior to matriculation, these include submitting final high school transcripts and immunization records. For matriculated students, these include course registration, participation in mandatory advising, and other optional tasks related to topics such as parking and housing.

(4) *Social*: invitations for students to attend social activities and build community in college. These campaigns include information on pre-college orientation programs, convocation, student clubs, organizations and fairs, among others.

(5) *Informative*: campaigns about campus resources, such as those providing free food and basic essential items, as well as key campus calendar information, such as holidays and campus closing dates.

(6) *Support*: campaigns that offer additional supports to students through check-ins, surveys asking how the university can help them, and encouragement and "good luck" messages.

Each campaign has a target population, which is the set of students for whom the message is intended. A campaign is considered general if it is sent to the full set of students receiving chatbot outreach. For example, *informative* campaigns that share key campus calendar information with students are general and go to all text-eligible students. In contrast, a campaign is targeted if it is

[1] Early Start is a summer academic support program for newly admitted students who have not yet demonstrated college-level proficiency in mathematics and/or written communication based on CSU system-wide placement standards. See more here: https://www.csun.edu/prospective-students/early-start-program-0?utm

sent to a selected subset of students for whom the campaign is deemed applicable and useful. For example, a campaign regarding course registration might only be sent to those students who had not yet registered by a certain deadline.

Outreach campaigns also vary in the level of interaction CSUN expects from students in response. CSUN distinguishes between nudges and interactive campaigns. Nudges aim to prompt specific behaviors or actions from students. For example, a nudge could be a message informing students about course registration or financial aid application deadlines and prompting them to complete the required administrative tasks. Interactive campaigns are messages aiming to engage students in a text-based dialogue, for example, messages asking about whether the university can help them or asking them to respond to a survey.

Students receive and respond to CSUNny communication via text message on their cell phones. When a student responds to a message, follow-up interaction between the CSUNny system and the student can occur in one of several ways. First, a student may answer a closed-ended question on a binary (e.g., yes/no) or categorical (e.g., 1 - 5) scale. Such student responses can trigger pre-planned follow-up messages. Second, a student may text in with a question in narrative form. A core feature of the Mainstay platform used by CSUN is its knowledge base. This can be thought of as a campus-specific, pre-programmed, comprehensive set of questions and responses developed collaboratively by Mainstay staff and CSUN administrators. Incoming student questions that algorithmically match to information in the system knowledge base are answered immediately by the system's non-generative AI. When the AI cannot answer a question with a sufficiently high probability of success, the question is flagged for the human system administrator to respond. This is the third way that students receive responses to their messages. When a human administrator responds to a question, the student's question and the appropriate response are added to the system knowledge base and, in this way, the system knowledge base is expanded and improved over time. Finally, students can initiate communication with the chatbot by texting at any time of day with questions, and their messages are handled the same way as those that are prompted by a campaign. The system non-generative AI first attempts to match the question to an existing answer in the knowledge base, flagging it for a human administrator only if a suitable response is not found. As before, any new question-answer pairing from this interaction is then added to the system's knowledge base, further improving the chatbot's intelligence.

III. BUILDING THE FOUNDATIONS FOR ONGOING COMMUNICATION

Organizational resources: Project overview, collaboration, and data flow

The chatbot operates differently from a distributive communication system like email. With email, anyone within the university who has access to students' email addresses can directly contact them. In contrast, CSUN's chatbot is centrally operated, meaning that the communication channel is much more narrowly controlled and managed. Determining the optimal administrative home for the chatbot, which serves as one of the initial communication tools for incoming first-year students, was therefore crucial for the tool's success.

The Office of the Registrar was a natural home for the chatbot due to the office's role in tracking student matriculation and enrollment. At CSUN, the communications and outreach staff within the Office of the Registrar are traditionally the first to contact newly admitted students regarding registration processes, deadlines, and orientation activities. During the period of our observation and analysis, the chatbot was jointly overseen by the Office of Undergraduate Studies, within the Division of Academic Affairs, and the Office of the Registrar, which was part of the Office of Admission and Records within the Division of Student Affairs.[2]

The focus of CSUNny campaigns was determined collaboratively by the chatbot team with personnel from the two offices and was based on the perceived needs of students. This collaboration is depicted in Figure 1, with the Office of the Registrar at the center of the collaboration and data flow diagram. A designated communication specialist within the Office of the Registrar crafted all text campaigns to create consistency in CSUNny's "voice" and oversaw distribution of all campaigns through the chatbot software platform. Other members of the chatbot team monitored students' responses to interactive campaigns to inform additions to the knowledge base and future campaigns.

The chatbot team plays a crucial role in crafting and overseeing all student communications through the tool. They seek a balance between messages telling students what they need to hear (e.g., reminders about financial aid deadlines, nudges about enrolling in classes) and messages telling students what they want to hear (e.g., words of encouragement during finals week, "check in" campaigns). Campaigns were frequently tailored to students' status within the university (e.g., Freshman versus Transfer; first semester enrolled versus approaching graduation). The phrasing of messages was also carefully considered. For example, a reminder to enroll would be phrased differently within the semester, between semesters, or across academic years. This was done to humanize the chatbot and also to pique students' attention by not replicating language. The tone of the campaigns was intentionally upbeat. Threats or suggestions of negative consequences were

[2] Over time some of these offices merged and changed. However, the Office of the Registrar still exists, and the chatbot is still housed within it.

avoided. See Figure 2 for examples of messages sent to students about registration deadlines in Spring 2019 and Spring 2020.

More tailored messages require additional data customization. In these cases, the Office of the Registrar leverages its student-level administrative data to design and execute outreach campaigns. However, since student data is not centralized within the university, effective data sharing across various campus departments is essential for data-informed targeting of outreach. For instance, in preparation for a targeted campaign aiming to nudge students to file the FAFSA, the intended recipients are determined by pulling data from within the Office of the Registrar in collaboration with Financial Aid. In this scenario, the Financial Aid office provides the Office of the Registrar with the list of students who have not submitted their FAFSA, and the Office of the Registrar then passes the relevant subset of students to the chatbot team to deliver a targeted filing reminder.

Technological choices: a chatbot team, a knowledge base, and non-generative AI

CSUN's chatbot integrates a data infrastructure that guides the selection of recipients for targeted campaigns, alongside a chatbot team that carefully crafts the content of each message. The human component behind this technological tool enhances the potential of AI to benefit student populations that may be difficult to reach and engage through more traditional communication channels (e.g., email, phone calls). The AI component of the tool streamlines communication with students by allowing them to text their school at any time and receive immediate responses from a pre-approved knowledge base. This capability frees staff time from having to answer commonly asked questions repeatedly. Staff can focus on crafting targeted messages that address students' specific needs, developing language that resonates with the student body, and triaging incoming messages to the appropriate campus offices best equipped to address the questions and concerns raised through the chatbot that cannot be answered without staff input.

CSUNny's reliance on non-generative AI that draws on a pre-programmed knowledge base developed by the chatbot team ensures that students' questions are answered using a set of pre-approved responses that align with university communication standards. Whereas generative AI poses risks related to hallucinations, fabrication of information, and the inability to reliably safeguard personal data (Lee et al., 2025), the use of a non-generative AI tool allows the university to maintain tighter control over chatbot responses. This approach helps ensure that students receive accurate and up-to-date information while also protecting students' personal data.

IV. EVOLUTION OF CAMPAIGNS OVER TIME

We examine the evolution of campaigns using detailed data on all campaigns sent to students between the academic years 2018–2019 and 2022–2023. We categorize each campaign by its domain (academic, financial, administrative, social, informative, and support-related), intended level of interaction (nudge vs. interactive), and target population (targeted vs. general) (see Section II for details). Table 1 presents the total number of campaigns for each academic year, along with their distribution by type and domain.

From Fall 2018 to Spring 2023, CSUN sent an average of 78 campaigns per academic year. Campaign volume peaked during the COVID-19 pandemic in AY2020-21 with 101 distinct campaigns and was at its lowest in AY2022-23 with 61 campaigns.

While most campaigns in the first year were nudges (94%), during the pandemic, CSUN tried to communicate with students through more interactive messages (8-10% of all campaigns). However, by the end of our period of observation, interactive messages represented only 3% of all the campaigns. Campaigns also evolved from a mix of targeted and general campaigns to a high concentration of general campaigns distributed to all students. While targeted campaigns were about 36% of all campaigns in AY2018-19, this type of campaign accounted for only 6% of the total campaigns sent in AY2022-23. In absolute terms, we also observe a decline in the number of targeted campaigns, going from 30 in the first year to less than five in the last academic year that we observe. As noted above, the siloed nature of data and recordkeeping at CSUN meant that the customization and targeting of chatbot communication required ongoing data sharing and collaboration across university offices that may be burdensome and, therefore, difficult to sustain over time.

The domains covered in the campaigns that students received included academic, financial, administrative, social, informative, and support-related content. Informative campaigns make up the largest share of campaigns each year, although their overall share changes over time. For example, in AY2018-19, CSUN sent 34 informative campaigns (40%), compared to 75 in AY2020-21 (50%) and 63 in AY2022-23 (55%). As informative campaigns became more prevalent, other types of campaigns that were more frequently implemented in earlier years played a smaller role by the end of the period. For instance, in AY2018-19, administrative campaigns represented 25% with 21 campaigns sent, and academic campaigns represented 19% with 16 campaigns. In contrast, in AY2022-23, administrative campaigns represented 16% with 18 campaigns, and only 4 academic campaigns were sent (4%).

According to CSUN's chatbot team, which we consulted for this study, the variation in campaigns over time was a response to changes in students' needs. For instance, there was an increase in support for students unexpectedly transitioning to virtual learning during the COVID-19 pandemic, along with a university effort to foster a sense of connectedness through more interactive campaigns during this unprecedented time. Similarly, changes in the content of the messages (i.e., the campaign domain) were driven by the chatbot team's assessment of the most critical information students needed to monitor their academic progress and success each year. Additionally, this may help explain CSUN's shift toward more general campaigns as a way to communicate relevant information to all students during unusual periods such as the pandemic and the transition to virtual learning.

V. STUDENT ENGAGEMENT WITH CHATBOT COMMUNICATION

To assess the extent to which students were receptive to the university's chatbot communication over time, we turn to data on student engagement with and responses to the text campaigns over a four-year period. To understand students' use of chatbot communication, we evaluate their reactions to and interactions with the tool by calculating: (1) the percentage of students who opt out of receiving chatbot communication, and (2) the percentage of students who actively use the tool by sending messages to the chatbot.

Opt-out rate

The opt-out rate is the percentage of students who request to be removed from chatbot outreach, calculated as a proportion of the total number of students subscribed to chatbot communication. We measure monthly opt-out rates for each academic year. In most months, only a handful of students request removal from the chatbot. Cumulatively, as shown in Figure 3, annual opt-out rates do not exceed 4%. We also observe that most students who opt out do so early in the fall semester (September-November), with very few opting out thereafter. The only exception is AY2022-23, when opt-out levels remained lower than in previous years until a campaign in April 2023 targeting Fall 2023 admitted students. This campaign asked students about their enrollment intentions and explicitly noted to them the option to opt out of chatbot communication. This led to a couple of hundred students (about 2% of the targeted group) opting out. Nevertheless, the annual opt-out rate remained low as in the rest of the years across the study period.

Active engagement

Although less frequently used (see section IV, Evolution of Campaigns Over Time), the chatbot team also designed interactive campaigns where a response from students is sought. These campaigns, for example, might pose a question for students to answer or invite them to share how their semester is going and how the university can support them. We expect that students would be more likely to respond when asked a direct question. Students can also text the chatbot, unprompted, whenever they have a question. For this reason, we measure active engagement and aggregate it in two ways. We estimate the rate of *active engagement* as the proportion of students who respond to the chatbot after an interactive campaign, divided by the total number of students receiving the campaign. We calculate this ratio by campaign and then aggregate by (1) campaign domain and (2) common themes in the messages to which students are responding.

Among campaigns that invite some level of engagement, the average engagement rate is approximately 5% (see Figure 4). This means that 5% of students who received a given interactive campaign respond to it by texting back to the chatbot. The types of campaigns that elicited the highest response rates from students varied over time. While academic campaigns engaged about 10% of recipients on average in AY2018-19, administrative campaigns had the highest engagement in AY2022-23, with more than 20% of recipients responding on average.

Figure 4 also highlights that an individual campaign can achieve a high engagement rate. For instance, although the mean engagement rate for informative campaigns in AY2019-20 was below 10%, a campaign checking in on students before their spring semester reached nearly 46% engagement. This indicates that variation in students' response levels could be partially driven by the specific topics addressed in the messages sent to students. For instance, administrative and support-related campaigns are among those with the highest level of engagement over the years studied (see Figure 4).

With more granular data on the description of messages sent to students, we examined the main topics covered in campaigns to understand which topics elicited the most student engagement. We were able to classify most campaigns to which at least one student responded into a set of common categories or keywords, including *phone*, *check-in*, *survey*, *follow-up*, *help*, *resolutions*, *holiday*, *housing*, *immunizations*, *app*, *registration*, *orientation*, *fees*, *financial aid*, *transcripts*, and *introduction*. Once classified, we ordered the groupings by response rate and found that students were most responsive to "*check-in*" and *survey* campaigns. These campaigns made up the majority of interactive campaigns during this period, with at least 15 survey-based campaigns aimed at checking in on students. On average, these campaigns received at least a 6% response rate, with surveys that ask students about their plans to return for the fall semester achieving response rates of about 45%.

Finally, it is important to note that not all campaigns are designed to prompt an answer from students. This is particularly true for informational campaigns. For example, campaigns that remind students of upcoming deadlines are designed to encourage them to complete all necessary steps to enroll in their classes without prompting a response to the text message unless they have questions (see Figure 5 for an example of this type of campaign). In these cases, passive engagement (i.e., not opting out) serves as a useful indicator that students have chosen to remain in the system to receive reminders and general information from the university via the chatbot. So as long as opt-out rates remain low and active engagement rates are relatively high for interactive campaigns, this can be interpreted as evidence that students are receptive to communication through the tool.

VI. INTERVENTION IMPACT

CSUN's chatbot use began with experimental subsets of the entering classes of 2018 and 2019 but soon expanded beyond these groups. In Fall 2019, CSUN extended access to all incoming, first-time transfer students (FTT, n=5,700). In March 2020, as the COVID-19 pandemic unfolded, CSUN further expanded access to all continuing students, except those in the experimental study's control group and graduating seniors (n=10,000). Since Fall 2020, CSUN has communicated with all incoming classes through the chatbot.

CSUN began outreach to selected students in the entering classes of 2018 and 2019 in the spring prior to the start of their first year in college. Messages sent in the spring and summer were intended to keep newly admitted first-year students engaged with CSUN to ensure that they ultimately matriculated. CSUNny communication then continued after matriculation to provide students with ongoing support as they navigated their college experiences.

Experimental Design

We evaluated the impact of the CSUN chatbot using a randomized controlled trial design. The experiment includes the entering first-year cohorts of 2018 and 2019. Between the two cohorts, the tool was experimentally tested with approximately 4,361 undergraduates, with another 4,347 students assigned to the control group. For both cohorts, we observed and checked balance on baseline characteristics including indicators of student race/ethnicity, gender, financial aid status, first-generation status, and prior academic achievement. After randomization, we did not observe systematic differences between the two groups (see Table 2).

Across the two cohorts, about 56% of the sample was female, 69% were eligible for the Pell Grant, and 51% were the first in their family to attend college. CSUN serves a student body that is racially and ethnically diverse. In our sample, 66% of students were Hispanic, 14% were White, 9% were Asian, 6% were Black, and 4% were of another race/ethnicity. The average high school grade point average (GPA) was 3.45 out of 4, and the average combined math and verbal SAT score was 1,023.

Analytic Strategy

To assess the impact of treatment assignment on student outcomes, we estimate fixed effects regression and linear probability models with the following general form:

$$Y_{it} = \alpha_t + \beta \times Treatment_i + \boldsymbol{X\gamma} + \varepsilon_{it} \qquad (1)$$

where $Treatment_i$ is an indicator equal to 1 if $student_i$ was randomized to treatment and zero otherwise. To account for randomization being done within cohorts, we include a cohort fixed effect, α_t. $\boldsymbol{X}$ is a vector of baseline covariates that include gender, race/ethnicity, Pell eligibility, first-generation status, high school GPA, and SAT scores, and ε_{it} is the individual error term. Our key coefficient of interest, β, represents the intent-to-treat effect of being assigned to the text-communication treatment group on outcome Y_{it}.

The results presented below are intent-to-treat (ITT) effects estimated with and without baseline covariate controls. To assess whether there is any heterogeneity in effects across cohorts, we also estimate effects separately by cohort. Additionally, we explore the effects within four subgroups: female, Hispanic, Pell-eligible, and first-generation students.

Targeted campaigns and impact estimates on enrollment

Examining the effect of targeted campaigns is particularly important for understanding whether this type of communication tool can effectively reach populations that may require more tailored support. However, evaluating such campaigns requires more tailored data than the evaluation of general campaigns. For the university, this involves maintaining records of the targeting criteria and tracking the outcomes of students who received the campaigns. While this can be more challenging than evaluating general campaigns, it may be less complicated for institutions with centralized data systems. Given the organizational structure of CSUN (see Figure 1 above), it was more difficult for the communication team to access certain data elements. As a result, we are unable to report on the effects of most targeted campaigns, such as those reminding students to file the FAFSA.

We were able to examine two targeted campaigns sent during the academic year 2018-19 that were focused on enrollment-related outcomes. The first campaign was a registration reminder for students who had not yet enrolled in classes. For this campaign, we analyzed whether students who received the text reminder were more likely to enroll within the following two weeks (for the first reminder) or four weeks (for the second reminder). Results are shown on Panel A of Table 3. The first reminder was particularly effective, with treated students being 34 percentage points more likely to enroll for classes by August 16th compared to the control group. By the second reminder, most students had already enrolled; however, treated students were still 2 percentage points more likely to enroll by September 15th than those in the control group.

The second campaign we analyzed encouraged students to register for CSUN's Early Start program, a proactive summer program designed to prepare First-Time Freshman students for college-level academics, thereby improving their path to graduation. Eligibility for Early Start is determined based on multiple measures (e.g., high school coursework and GPA, SAT scores). Unless students are deemed to be exempt based on these measures, they are either strongly encouraged or required to take Early Start math and/or writing classes. We assessed whether students targeted to receive the Early Start campaign were more likely to register within two or four weeks of the outreach. These results are presented on panel B of Table 3.

For Early Start, students required to enroll received targeted campaigns on May 25th and again on June 8th. After the May 25th text message, treated students were 11 percentage points more likely to enroll in Early Start by June 7th compared to the control group. Following the June 8th reminder, treated students were 20 percentage points more likely to enroll by June 22nd compared to those who did not receive reminders.

Impact estimates on enrollment and graduation

Data on enrollment and achievement by semester is available from CSUN's Office of Institutional Records. The number of credits attempted and earned, and the term and cumulative GPA are collected for both treated and control students. Information from Fall 2018 to Spring 2023 is available, meaning that we are able to observe four complete academic years for both cohorts in our sample.

We present results based on the sequence of semesters in which students are enrolled—that is, whether they are in their first, second, or eighth semester. Additionally, we examined whether treatment group students were more likely to graduate by the end of their fourth year.

We find no statistically significant difference in enrollment and persistence between students in the treatment and control groups (Table 4). Similarly, there is no evidence of a difference in graduation rates. Furthermore, when disaggregating the analysis by subgroups, text-based outreach and communication did not lead to significant differences in enrollment across semesters for female, Hispanic, Pell-eligible, or first-generation students in the treatment group (results not shown here but available upon request).

In Table 5, we show that enrollment intensity, measured by the number of units students enroll in and earn each semester, was not affected by the treatment. Specifically, there was no significant impact on the number of credits students enrolled in, earned, or accumulated compared to the control group.

General campaigns and impact estimates on academic success

To further assess the effect of the chatbot on students' academic success, we examined both their term-level and cumulative GPA. We observe no statistically significant effect of the text-based outreach on either measure (Table 6). These results remained consistent across individual cohorts and student subgroups, including female, Hispanic, Pell-eligible, and first-generation students (results not shown here but available upon request).

Overall, the impact evaluation results for CSUN's chatbot are consistent with prior evidence from other colleges, showing that students are most responsive to this type of communication when prompted to complete time-sensitive administrative tasks required for continued enrollment (Page et al., 2025). However, communication delivered through a tool like this, on its own, may be insufficient to generate measurable changes in students' longer-run academic performance and persistence outcomes.

VII. CONCLUSIONS

As AI adoption in education expands, institutions are taking different approaches to integrate AI-enabled tools designed to support teaching, learning, and student success. In higher education in particular, student success offices are increasingly deploying AI-enabled chatbots to communicate with students about administrative and academic tasks. Existing evidence on the effectiveness of these tools is largely drawn from short-term interventions, with limited understanding of their sustainability, student receptivity, and impact over longer periods.

This study addresses that evidence gap by examining a multi-year implementation of the AI-enabled text-messaging chatbot at California State University, Northridge (CSUN). Drawing on system observation, discussions with administrators, and a four-year randomized controlled trial following two undergraduate cohorts, we assess (1) the institutional conditions required to sustain centralized chatbot communication, (2) students' responsiveness over time, and (3) the tool's effects on student outcomes throughout college.

We find that sustained chatbot implementation requires centralized institutional ownership and flexibility in communication strategies to adapt to students' evolving needs and external shocks, such as the COVID-19 pandemic. Effective implementation also depends on cross-unit coordination and data sharing, particularly when outreach is tailored to specific student needs. Because relevant student data is distributed across multiple campus offices, targeted campaigns require sustained collaboration among units across the university. This need for coordination helps explain the institution's shift toward broader, general-purpose campaigns during periods of disruption, when communicating critical information to all students became both necessary and more feasible. At the same time, the siloed nature of data and recordkeeping at CSUN meant that customizing and targeting chatbot communication required ongoing data sharing and interoffice collaboration, which can be burdensome and was difficult to sustain over time. These challenges were also evident during the evaluation of targeted campaigns, when retrieving the necessary data proved complex.

Finally, students remain receptive to chatbot communication over time, as shown by low opt-out rates and continued engagement. Consistent with prior research, the primary impacts of the chatbot are concentrated in improved completion of time-sensitive administrative tasks (Page et al., 2025), with no detectable effects on academic performance or persistence.

The lessons from studying CSUN's AI communication tool contribute to the broader understanding of AI systems by showing that these tools (1) still require significant human engagement; (2) are strengthened and made more effective through targeted outreach, highlighting that robust institution-level data systems are an important precursor to successful implementation; and (3) while they can support students in completing specific tasks, this type of communication may be insufficient on its own to improve overall student persistence and academic success. Accordingly, institutions may wish to pair such tools with complementary strategies aimed at improving broader student outcomes.

REFERENCES

Baek, C., Tate, T., & Warschauer, M. (2024). “ChatGPT seems too good to be true”: College students’ use and perceptions of generative AI. *Computers & Education: Artificial Intelligence, 7*, 100294.

Chu, T. L. (2022). Applying positive psychology to foster student engagement and classroom community amid the COVID-19 pandemic and beyond. *Scholarship of Teaching and Learning in Psychology, 8*(2), 154–163. https://doi.org/10.1037/stl0000238

California State University. (2024). *Enrollment dashboard*. https://tableau.calstate.edu/views/SelfEnrollmentDashboard/EnrollmentSummary

Demszky, D., Liu, J., Hill, H. C., Jurafsky, D., & Piech, C. (2024). Can automated feedback improve teachers’ uptake of student ideas? Evidence from a randomized controlled trial in a large-scale online course. *Educational Evaluation and Policy Analysis, 46*(3), 483–505.

Doss, C. J., Bozick, R., Schwartz, H. L., Chu, L., Rainey, L. R., Woo, A., Reich, J., & Dukes, J. (2025). *AI use in schools is quickly increasing but guidance lags behind: Findings from the RAND Survey Panels* (RR-A4180-1). RAND Corporation. https://www.rand.org/pubs/research_reports/RRA4180-1.html

EAB. (2024). *University staff say AI can help them identify and support struggling students*. https://eab.com/about/newsroom/press/survey-ai-student-success/

Elsen-Rooney, M. (2023, January 3). NYC Education Department blocks ChatGPT on school devices, networks. *Chalkbeat New York*. https://www.chalkbeat.org/newyork/2023/1/3/23537987/nyc-schools-ban-chatgpt-writing-artificial-intelligence/

Ferlazzo, L. (2023, January). ChatGPT: Teachers weigh in on how to manage the new AI chatbot. *Education Week*. https://www.edweek.org/technology/opinion-chatgpt-teachers-weigh-in-on-how-to-manage-the-new-ai-chatbot/2023/01

Hopelab, Common Sense Media, & Center for Digital Thriving at Harvard. (2024). *Teen and young adult perspectives on generative AI: Patterns of use, excitements, and concerns*. https://hopelab.org/stories/teen-and-young-adult-perspectives-on-generative-ai

Lee, C., Kim, J., Lim, J. S., & Shin, D. (2025). Generative AI risks and resilience: How users adapt to hallucination and privacy challenges. *Telematics and Informatics Reports*, 100221.

Page, L. C., & Gehlbach, H. (2017). How an artificially intelligent virtual assistant helps students navigate the road to college. *AERA Open*. https://doi.org/10.1177/2332858417749220

Page, L. C., Meyer, K., Lee, J., & Gehlbach, H. (2025). Conditions under which college students can be responsive to text-based nudging. *Journal of Research on Educational Effectiveness*, 1–29. https://doi.org/10.1080/19345747.2025.2481219

Machidon, O. M. (2025). Generative AI and childhood education: Lessons from the smartphone generation. *AI & Society*, *40*, 4295–4297.

Mah, C., Tan, M., Phalen, L., Sparks, A., & Demszky, D. (2025). *From sentence corrections to deeper dialogue: Qualitative insights from LLM and teacher feedback on student writing* (EdWorkingPaper No. 25-1193). Annenberg Institute at Brown University. https://doi.org/10.26300/p397-2p46

Markel, J. M., Opferman, S. G., Landay, J. A., & Piech, C. (2023). Gpteach: Interactive TA training with GPT-based students. In *Proceedings of the tenth ACM conference on learning at scale* (pp. 226–236). Association for Computing Machinery. https://doi.org/10.1145/3573051.3593393

Meyer, K., Page, L. C., Mata, C., Smith, E. N., Walsh, B. T., Fifield, C. L., Tyson, M., Eremionkhale, A., Evans, M., Frost, S., & Jung, E. E. (2024). *Let's chat: Leveraging chatbot outreach for improved course performance* (EdWorkingPaper No. 22-564). Annenberg Institute at Brown University. https://doi.org/10.26300/es6b-sm82

Nurshatayeva, A., Page, L. C., White, C. C., & Gehlbach, H. (2021). Are artificially intelligent conversational chatbots uniformly effective in reducing summer melt? Evidence from a randomized controlled trial. *Research in Higher Education*, *62*(3), 392–402.

Tate, T. P., Steiss, J., Bailey, D., Graham, S., Moon, Y., Ritchie, D., Tseng, W., & Warschauer, M. (2024). Can AI provide useful holistic essay scoring? *Computers & Education: Artificial Intelligence*, *7*, 100255.

UNESCO. (2025). *AI and the future of education: Disruption, dilemmas and directions*. UNESCO. https://doi.org/10.54675/KECK1261

Wang, Z., Yin, Z., Zheng, Y., Li, X., & Zhang, L. (2025). Will graduate students engage in unethical uses of GPT? An exploratory study to understand their perceptions. *Educational Technology & Society*, *28*(1), 286–300. https://doi.org/10.30191/ETS.202501_28(1).TP04

Wang, R. E., Ribeiro, A. T., Robinson, C. D., Loeb, S., & Demszky, D. (2024). *Tutor CoPilot: A human–AI approach for scaling real-time expertise* (EdWorkingPaper No. 24-1054). Annenberg Institute at Brown University. https://doi.org/10.26300/81nh-8262

FIGURES AND TABLES

Figure 1. Chatbot Data and collaboration flow

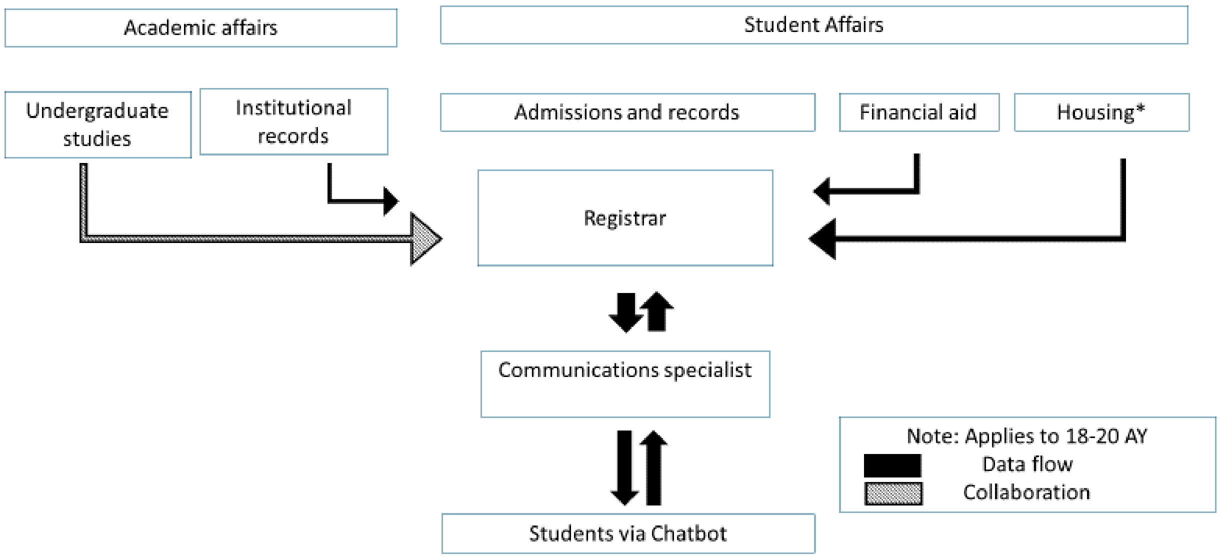
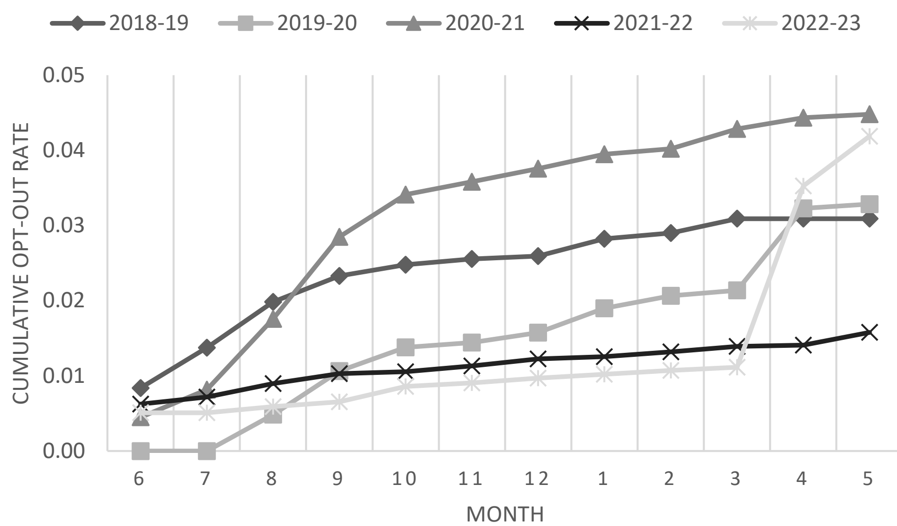


Figure 2. Selected examples of text messages sent to students via CSUNny

Spring 2019	Spring 2019 (follow-up)	Spring 2020
Welcome back from Spring Break! It may seem super early, but it's already time to start thinking about Fall 19. Be on the lookout for an email this week with your registration info. Your registration date for Fall will be here soon, so check out our Registration Guide [link] for tips, info, faqs, and links to how-to guides.	Spring registration dates are out now, so make sure to check your CSUN email for details.	Hi [first name]. I know things are kind of strange these days, and everyone is still adjusting to new challenges. But we're looking forward! Emails with Fall 20 registration info have been sent out! I'm here as always, if you have questions. My human helpers are working from home, doing their best to make sure I have the most updated information!

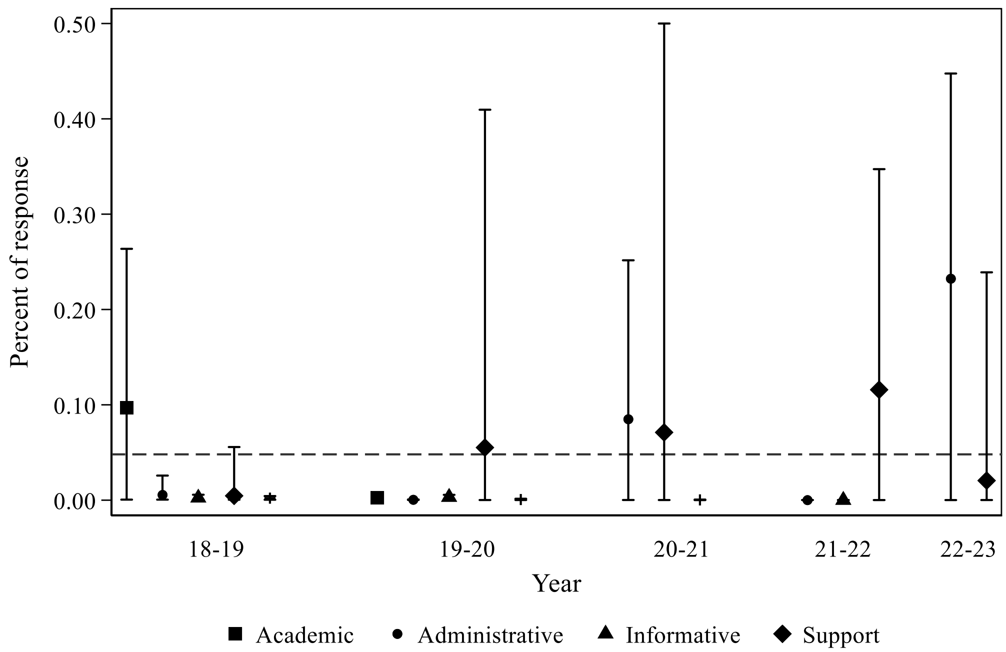
Source: CSUN

Figure 3. Cumulative opt-out rates by month and academic year



Source: Authors' calculations with CSUN data.

Figure 4. Student engagement by campaign domain and academic year
(includes only campaigns with positive rates of response)



Source: Authors' calculations with CSUN data.
Notes: Symbol denotes the mean response rate for each academic domain and the whiskers the range in the rate of response. Dash line represents the mean rate of response for all domains and years. The absence of a symbol in an academic year represents null engagement of students with that domain of campaigns. Includes only campaigns with positive rates of response as a proxy of campaigns that were drafted to prompt student's response.

Figure 5. Example of registration campaign

Hi [first name]! Just a friendly reminder that registration for freshmen begins next weekend. Be sure to make an appointment with your advisor if you haven't.

[Contacts without this profile information receive a backup text.]

Also, don't forget to complete your Title IX training before seeing your advisor. If you don't, there could be a registration hold placed on your account.

Here is a handy guide [link] on how to complete the Title IX training.

Source: CSUN.

Table 1. Distribution of chatbot campaigns by academic year, type, and domain.

Academic year	**2018-19**	**2019-20**	**2020-21**	**2021-22**	**2022-23**
Total campaigns	82	72	101	76	61
Type (%)					
Nudge	0.94	0.90	0.92	0.99	0.97
Interactive	0.06	0.10	0.08	0.01	0.03
Targeted	0.36	0.15	0.11	0.14	0.06
General	0.64	0.85	0.89	0.86	0.94
Domain (%)					
Academic	0.19	0.04	0.09	0.03	0.04
Financial	0.07	0.07	0.07	0.04	0.03
Administrative	0.25	0.12	0.15	0.16	0.16
Social	0.07	0.12	0.06	0.08	0.10
Informative	0.40	0.55	0.50	0.58	0.55
Support	0.02	0.10	0.13	0.12	0.13

Source: Authors' analysis of CSUN administrative data.

Table 2. Baseline Characteristics and Group Balance of the CSUN Analytic Sample by Entering Class

Covariates	Full Analytic Sample				Fall 2018-Cohort				Fall 2019-Cohort			
	All	Control	Treated	Diff.	All	Control	Treated	Diff.	All	Control	Treated	Diff.
Female	0.56	0.56	0.56	0.00	0.55	0.54	0.55	0.01	0.57	0.58	0.57	-0.01
White	0.14	0.14	0.13	-0.01	0.14	0.14	0.14	-0.01	0.14	0.14	0.13	-0.01
Black	0.06	0.06	0.06	0.00	0.06	0.06	0.05	0.00	0.06	0.06	0.06	0.00
Hispanic	0.66	0.66	0.66	0.00	0.65	0.65	0.66	0.01	0.66	0.66	0.66	0.00
Asian	0.09	0.09	0.09	0.00	0.09	0.09	0.09	0.00	0.08	0.08	0.08	0.00
Other ethnicity	0.04	0.04	0.04	0.00	0.04	0.04	0.03	-0.01	0.04	0.04	0.05	0.01
Pell eligible	0.69	0.69	0.68	0.00	0.59	0.59	0.59	0.01	0.77	0.77	0.76	-0.01
First-generation	0.51	0.51	0.51	0.00	0.50	0.50	0.51	0.01	0.52	0.52	0.51	-0.01
High school GPA	3.45	3.45	3.45	-0.01	3.43	3.42	3.44	0.02	3.47	3.48	3.45	-0.03
SAT	1023.49	1024.67	1022.31	-2.35	1027.98	1029.98	1026.00	-3.98	1019.61	1020.08	1019.14	-0.94
Observations	8708	4347	4361		4033	2015	2018		4675	2332	2343	

Source: Authors' calculations with CSUN data.
Notes: Statistical significance: ~ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 3. Impact estimates of enrollment and Early Start targeted campaigns, AY2018-19.

	Panel A Semester Registration					
Campaign date:	31-Jul-18			17-Aug-18		
Enrollment by:	16-Aug-18			15-Sep-18		
	Control mean	Treatment effect		Control mean	Treatment effect	
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	0.315	0.342***	0.341***	0.977	0.023**	0.022**
S.E.	(0.025)	(0.038)	(0.039)	(0.008)	(0.008)	(0.008)
Observations	621	621	621	621	621	621
R-squared	0.115	0.115	0.125	0.010	0.010	0.030
Covariates	No	No	Yes	No	No	Yes

	Panel B Early Start					
Campaign date:	25-May-18			8-Jun-18		
Enrollment by:	7-Jun-18			22-Jun-18		
	Control mean	Treatment effect		Control mean	Treatment effect	
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	0.119	0.145***	0.115***	0.204	0.246***	0.202***
S.E.	(0.010)	(0.028)	(0.028)	(0.013)	(0.032)	(0.032)
Observations	1,263	1,263	1,263	1,263	1,263	1,263
R-squared	0.029	0.029	0.081	0.056	0.056	0.136
Covariates	No	No	Yes	No	No	Yes

Note: Covariates include gender, race/ethnicity, Pell eligibility, first-generation status, high school GPA, and SAT scores. Robust standard errors in parentheses. Statistical significance: *** p<0.001, ** p<0.01, * p<0.05, ~ p<0.10

Table 4. Control Means and Treatment Effects on Enrollment by Semester Sequence and Graduation by Fourth Year

Semester sequence	Control mean	Treatment effect	
	(1)	(2)	(3)
2nd	0.945	0.001	0.001
	(0.003)	(0.005)	(0.005)
3rd	0.834	-0.009	-0.008
	(0.006)	(0.008)	(0.008)
4th	0.784	0.001	0.002
	(0.006)	(0.009)	(0.009)
5th	0.724	0.001	0.002
	(0.007)	(0.010)	(0.009)
6th	0.708	-0.002	-0.001
	(0.007)	(0.010)	(0.010)
7th	0.686	0.003	0.004
	(0.007)	(0.010)	(0.010)
8th	0.676	0.003	0.004
	(0.007)	(0.010)	(0.010)
Graduate by 4th year	0.190	0.004	0.005
	(0.006)	(0.008)	(0.008)
Observations	8,712	8,712	8,712
Covariates	No	No	Yes
Cohort FE	No	No	Yes

Source: Authors' estimation with CSUN data.
Notes: In the first semester of each cohort both treated and control were enrolled, therefore, the estimation of the impact on enrollment is not possible. The 1st academic semester of F18 cohort is Fall of 2018 (calendar semester) and its 8th is Spring of 2022. The 1st academic semester of F19 cohort is Fall of 2019 and Spring of 2023 (calendar semester) is its 8th academic semester. Robust standard errors in parentheses. Statistical significance: ~ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table 5. Control Means and Treatment Effects on Enrolled, Earned, and Cumulative Earned Units per Semester Sequence

Semester sequence	**Units Enrolled**			**Units Earned**			**Cumulative Units Earned**		
	Control mean	Treatment effect		Control mean	Treatment effect		Control mean	Treatment effect	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1st	14.373	-0.003	-0.002	12.617	-0.093	-0.080	12.617	-0.093	-0.080
	(0.024)	(0.033)	(0.032)	(0.058)	(0.083)	(0.079)	(0.058)	(0.083)	(0.079)
2nd	13.357	0.028	0.035	11.792	0.002	0.015	24.409	-0.091	-0.065
	(0.056)	(0.079)	(0.078)	(0.073)	(0.104)	(0.100)	(0.118)	(0.168)	(0.159)
3rd	11.766	-0.135	-0.122	10.729	-0.059	-0.044	35.138	-0.151	-0.108
	(0.084)	(0.120)	(0.117)	(0.090)	(0.127)	(0.123)	(0.190)	(0.271)	(0.257)
4th	10.876	-0.002	0.015	9.891	0.078	0.094	45.029	-0.072	-0.014
	(0.091)	(0.129)	(0.125)	(0.094)	(0.132)	(0.127)	(0.271)	(0.385)	(0.365)
5th	9.899	-0.046	-0.022	10.174	-0.250	-0.225	60.624	-0.414	-0.453
	(0.099)	(0.139)	(0.135)	(0.127)	(0.176)	(0.170)	(0.462)	(0.649)	(0.590)
6th	9.507	0.021	0.029	9.714	-0.138	-0.120	65.440	-0.625	-0.447
	(0.099)	(0.140)	(0.137)	(0.126)	(0.176)	(0.170)	(0.498)	(0.700)	(0.648)
7th	8.855	0.089	0.105	10.029	-0.022	0.001	75.036	-0.362	-0.300
	(0.101)	(0.143)	(0.140)	(0.146)	(0.205)	(0.200)	(0.594)	(0.840)	(0.789)
8th	8.479	-0.047	-0.036	9.64	-0.225	-0.209	79.698	-0.901	-0.770
	(0.102)	(0.144)	(0.141)	(0.145)	(0.203)	(0.199)	(0.652)	(0.913)	(0.870)
Observations	8,708	8,708	8,708	8,708	8,708	8,708	8,708	8,708	8,708
Covariates	No	No	Yes	No	No	Yes	No	No	Yes
Cohort FE	No	No	Yes	No	No	Yes	No	No	Yes

Note: Robust standard errors are in parentheses. Statistical significance: ~ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 6. Control Means and Treatment Effects on Term and Cumulative GPA by Semester Sequence

Semester sequence	Term GPA: Control mean	Term GPA: Treatment effect		Cumulative GPA: Control mean	Cumulative GPA: Treatment effect	
	(1)	(2)	(3)	(4)	(5)	(6)
1st	2.796	-0.027	-0.022	2.796	-0.027	-0.022
	(0.015)	(0.022)	(0.020)	(0.015)	(0.022)	(0.020)
2nd	2.744	-0.017	-0.010	2.746	-0.010	-0.006
	(0.017)	(0.025)	(0.023)	(0.015)	(0.022)	(0.020)
3rd	2.592	-0.008	0.001	2.602	-0.014	-0.010
	(0.020)	(0.028)	(0.026)	(0.017)	(0.025)	(0.023)
4th	2.351	0.007	0.011	2.409	0.007	0.012
	(0.023)	(0.032)	(0.031)	(0.021)	(0.029)	(0.028)
5th	2.131	0.001	0.006	2.356	-0.019	-0.013
	(0.024)	(0.034)	(0.032)	(0.021)	(0.030)	(0.029)
6th	2.086	-0.005	-0.003	2.307	-0.005	-0.001
	(0.024)	(0.034)	(0.033)	(0.022)	(0.031)	(0.029)
7th	2.054	-0.011	-0.007	2.369	0.001	0.005
	(0.024)	(0.034)	(0.033)	(0.021)	(0.030)	(0.029)
8th	1.993	-0.030	-0.027	2.301	-0.013	-0.008
	(0.024)	(0.035)	(0.033)	(0.022)	(0.031)	(0.030)
Observations	8,708	8,708	8,708	8,708	8,708	8,708
Covariates	No	No	Yes	No	No	Yes
Cohort FE	No	No	Yes	No	No	Yes

Note: The 1st academic semester of F18 cohort is Fall of 2018 (calendar semester) and its 8th is Spring of 2022. The 1st academic semester of F19 cohort is Fall of 2019 and Spring of 2023 (calendar semester) is its 8th academic semester. Robust standard errors in parentheses. Statistical significance: ~ p<0.10, * p<0.05, ** p<0.01, *** p<0.001