Keep me in Coach: The Short- and Long-Term Effects of a University Coaching Intervention *

Serena Canaan[a], Stefanie Fischer[b], Pierre Mouganie[c], Geoffrey C. Schnorr[d]

[a]Simon Fraser University and IZA
[b]Monash University and IZA
[c]Simon Fraser University and IZA
[d]California Policy Lab at UCLA and the California Employment Development Department

This Version: April 17, 2024

Abstract

To boost college graduation rates, policymakers often advocate for academic supports such as coaching or mentoring. Proactive and intensive coaching interventions are effective, but are costly and difficult to scale. We evaluate a relatively lower-cost group coaching program targeted at first-year college students placed on academic probation. Participants attend a workshop where coaches aim to normalize failure and improve self-confidence. Coaches also facilitate a process whereby participants reflect on their academic difficulties, devise solutions to address their challenges, and create an action plan. Participants then hold a one-time follow-up meeting with their coach or visit a campus resource. Using a difference-in-discontinuity design, we show that the program raises students' first-year GPA by 16.5% of a standard deviation, and decreases the probability of first-year dropout by 8.8 percentage points. Effects are concentrated among lower-income students who also experience a significant increase in the probability of graduating. Finally, using administrative data we provide the first evidence that coaching/mentoring may have substantial long-run effects as we document significant gains in lower-income students' earnings 7–9 years following entry to the university. Our mechanism analysis suggests that boosting students' self-confidence via providing them with in-person and credible encouragement can promote both their short- and long-term success.

*This paper has benefited from helpful conversations with Rachel Baker, Kelly Bedard, Laura Giuliano, Mark Hoekstra, Greg Kaplan, Philip Oreopoulos, Uros Petronijevic, David Molitor, Sarah Reber, Heather Royer, Jesse Rothstein, Bruce Sacerdote, Sally Sadoff, Wendy Stock, and seminar participants at UCSB, University of Melbourne, Hitotsubashi University, and UCSD's Rady School of Management. We thank participants at the following conferences for providing useful feedback: Society of Labor Economists Annual Meeting (2021), Association for Education, Finance and Policy Annual Meeting (2021), Montana State University's Applied Micro Workshop (2022), IZA Annual Education Workshop (2022), and Australian Annual Labour Econometrics Workshop (2022). We also thank the W.E. Upjohn Institute for Employment Research for support in conducting this research.

JEL codes: I23, I24, J16

Keywords: College Graduation, Mentoring, College advising, Academic coaching

1 Introduction

The labor market returns to a bachelor's degree are substantial and have been rising over time, with lower-income students realizing the largest economic gains (Hoekstra, 2009; Zimmerman, 2014; Goodman et al., 2020). At the same time, there is a large and growing socioeconomic gap in postsecondary attainment (Bailey and Dynarski, 2011). By age 24, degree completion rates are about five times higher for students from the highest income quartile relative to the lowest (Cahalan et al., 2021). Economic and educational disparities have been largely attributed to differences in non-cognitive skills (Heckman et al., 2006; Borghans et al., 2008). Importantly, while cognitive ability is determined in childhood, non-cognitive skills are more malleable at later ages—especially among individuals from disadvantaged backgrounds (Heckman and Rubinstein, 2001; Kautz et al., 2014).

Among these skills, self-efficacy and confidence are strongly correlated with education and labor market success (Heckman et al., 2006; Lindqvist and Vestman, 2011; Kuhnen and Melzer, 2018). Bénabou and Tirole (2005) provide a theoretical framework to explain the link between self-confidence and performance. The basis of their framework is that an individual has incomplete information about her abilities, and that effort and ability interact in the production of performance. In most settings, effort and ability are complements in producing performance, so higher self-confidence increases the motivation to exert effort towards a task and improves performance. This implies that anyone interested in boosting an individual's performance should use confidence-enhancing strategies, such as coaching, that reveal credible positive information about her ability. This model also conversely explains why individuals who anticipate failure are less likely to exert effort leading to lower performance. Importantly, the model predicts that revealing negative information about individuals' abilities without offering them confidence-enhancing incentives can be detrimental for performance.

In this paper, we ask whether a coaching program, aimed at improving self-confidence, is effective at boosting the performance of college students who receive negative information about their abilities. The program is implemented at a large, selective, public 4-year university in California and delivers mandatory in-person group coaching to all first-year students who are placed on academic probation in their first quarter at the university. Academic probation is a near-universal policy in higher education which warns and labels students who have performed inadequately (i.e., scored below a certain GPA threshold). Capable students may find themselves on probation early on in college as the transition from high school to college can be difficult. Students often struggle with the shock of more demanding course work, higher grading standards, more autonomy, and less-than-ideal living situations all while simultaneously trying to find a supportive community in a new environment. For some, these challenges interfere with academic performance. Consequently, probation is meant to serve as a wake-up call, but this label has been shown to have discouragement effects which manifest in increased first-year dropout and reduced graduation rates (Lindo et al., 2010). Such a policy will achieve its goal and be welfare-improving if it removes students who would benefit from going to a less selective institution. A concern, however, is that probation unintentionally weeds out students, especially those from non-traditional and disadvantaged backgrounds, who view the probation label as a signal that they do not belong, but who would benefit from remaining at the university. The coaching program we focus on aims to rectify this belief and correct for the potential inefficient attrition caused by academic probation.

The program's goal is to boost self-efficacy; i.e., students' confidence in their capacities to meet performance standards set by the university. It was developed based on Bandura's (1977) self-efficacy theory which—in line with the Bénabou and Tirole (2005) model—posits that when individuals have greater confidence in their capabilities to succeed in a certain task, they will exert more effort towards it.[1] Expectations of efficacy can be improved through exposing students to a trusted mentor who acts as a role model and provides encouragement, removing feelings of anxiety and stress, and setting goals (Bandura, 1977; Bandura and Schunk, 1981).

Program participants attend a two-hour workshop, where they are divided into small groups led by faculty and staff who are trained as coaches. Students are provided with role models and encouragement in several ways: they watch a video featuring very successful individuals who have previously experienced and overcome failure, and coaches share with students their own experiences with failure. The workshop is also conducted in a group setting to show students that they are not the only ones on probation. Importantly, this implies that students are not only provided with encouragement but are also given evidence that it is still possible to succeed despite failure, and that failure is normal. The program further offers students information on campus resources and academic probation in an effort to reduce feelings of anxiety and uncertainty regarding the probation label. Coaches facilitate a goal-setting and time management exercise whereby students identify the sources of their academic difficulties, find solutions to address their challenges, and create an action plan. Following the workshop, students are required to either use a campus resource once (e.g., office hours or counseling services) or have a follow-up one-on-one meeting with their coach.

We draw on rich administrative data for all first-year students entering the university in fall cohorts from 2007 to 2017. Students' academic outcomes are observed each quarter until they separate from the institution, allowing us to evaluate the impact of coaching on course performance, dropout rates, and degree attainment. We also investigate the program's effect on labor market outcomes, linking student data to administrative files from the State of California's Employment Development Department, which includes quarterly employment and earnings information for all employment covered by unemployment insurance (UI) in California for the years 2000-2022.

To estimate the causal effect of the program, we leverage the fact that first-year students are

[1] Bénabou and Tirole (2003) define self-confidence as an individual's "perceived prospects from undertaking a task". Self-efficacy is whether an individual believes that she is able to successfully execute a behavior required to produce a certain outcome (Bandura, 1977). In both frameworks, higher self-confidence or efficacy is predicted to increase individuals' motivation to undertake a task. Throughout the paper, the terms self-efficacy and self-confidence are used interchangeably.

assigned to the program if their GPA is less than 2.0 during their first quarter. Students who score below this threshold are also placed on academic probation, precluding us from using a standard regression discontinuity design. However, because the coaching program was introduced in 2009, we are able to leverage the fact that cohorts entering after that date were exposed to both academic probation and coaching, while those enrolled before that date were only exposed to academic probation. Consequently, we implement a difference-in-discontinuity (DiRD) design which compares the difference in outcomes for students on either side of the 2.0 GPA eligibility threshold in the 2009 entering cohort and later cohorts (treated cohorts) with this same discontinuity for the cohorts entering prior to 2009 (control cohorts). Intuitively, the DiRD design estimates the discontinuity in outcomes for treated cohorts and differences out any potential discontinuities in control cohorts' outcomes, thereby isolating the impact of the coaching program from that of being placed on academic probation.

We find that students achieve large academic gains from participating in the coaching program. Overall, the program increases participants' first-year GPA by 16.5% of a standard deviation, increases credits earned in the remainder of the year by 1.2, and decreases their likelihood of dropping out at the end of the first year by 8.8 percentage points (pp). We also find a positive albeit imprecisely estimated impact on 6-year graduation. Effects are more pronounced for groups of students who are expected to benefit most from such programs such as those from lower-income backgrounds, as it increases their probability of graduating in six years by a significant 14.9 pp. These findings are unsurprising given that lower-income students complete college at much lower rates than their counterparts (Bailey and Dynarski, 2011).

We next examine how the program impacts students' earnings 7-9 years following entry to the university.[2] Despite the caveat of reduced sample size and precision, our earnings analysis yields results that are consistent with the documented effects of the program on academic outcomes. In the full sample, we find that the coaching program has positive but imprecise impacts on students' earnings. However, we document large and statistically significant gains for lower-income students who realize a 31% increase in earnings. These point estimates are consistent with and contribute to a burgeoning body of work showing large earnings gains from attending a relatively more selective 4-year institution for academically marginal students (Zimmerman, 2014; Bleemer, 2018; Goodman et al., 2020; Bleemer, 2022; Black et al., 2022). Taken together, our findings indicate that providing discouraged students with group coaching largely improves their academic and early labor market outcomes.

Why would students incur such substantial benefits from participating in this program? To understand the mechanisms at play, we rely on the Bénabou and Tirole (2003) model of intrinsic and extrinsic motivation and adapt it to our setting. The model is a game between the university and the student, who has imperfect information about her ability (or her probability of success in

[2]We focus on earnings in quarters 25-36 (7-9 years) after initial enrollment, where quarter 1 is the fall quarter of the student's first year. The typical student enrolls at age 18, so this corresponds approximately to ages 24-26.

a certain task). The university, which has superiority of information about the student's ability, offers her an incentive structure in order to motivate her to remain at the university and not drop out. We argue that the different components of our coaching program can be considered as one of two incentive structures: help or encouragement with hard evidence. Help includes the assistance that coaches provide students with, such as helping them with time management and goal-setting or giving them information about probation and resources. Encouragement is when coaches convey to students that despite being on probation, they can still succeed. Importantly, coaches also provide students with hard evidence such as showing them a video of successful people who have overcome failure and sharing their own stories of success after failure. In the Bénabou and Tirole (2003) model, the student tries to infer the university's ulterior motive from choosing a certain policy. In other words, the chosen policy reveals private information that the university has about the student's ability. By analyzing the perfect Bayesian equilibria of this game, the model predicts that providing the student with "help" decreases her self-confidence, motivation and hence her likelihood of exerting effort to persist at the university. This is because the student knows that the university is incentivized to help lower-ability students, and hence the offer of help reveals negative information about the student's ability. On the other hand, encouragement with evidence is predicted to increase the student's self-confidence, motivation and effort as the university is essentially disclosing to the student "good news" about her ability.

We provide empirical support for the Bénabou and Tirole (2003) predictions by utilizing data from pre- and post-program surveys administered in select years to all program participants. A key feature of this survey is that it asks program participants questions about their motivation, self-confidence and effort i.e., the key elements for which the model provides predictions. We find that students report significant improvements in motivation, self-confidence and effort after participating in the program. While this analysis is not conclusive, it does indicate that our empirical findings are consistent with the predictions for encouragement (and not with the predictions for help). This suggests that disclosing "good news" to students about their probability of succeeding at the university (with hard evidence) is the main component of our program that is driving its positive impacts on academic and labor market outcomes. That is, disclosing such information boosts students' self-confidence, motivation and in turn effort.

Our paper is related to a broad literature which examines ways to effectively boost low college completion rates. A large body of work has considered the role of financial aid in increasing college attainment (Dynarski, 2003; Bettinger, 2015; Cohodes and Goodman, 2014; Bulman and Hoxby, 2015; Bettinger et al., 2019; Angrist et al., 2020). Our paper is however closest to recent interventions targeting non-financial barriers to students' academic success. One line of work focuses on low-touch interventions that are typically delivered online or via text messages and target students' non-cognitive skills or psychological barriers to degree attainment. Initial studies from social psychology such as one-time online goal-setting or mindset interventions showed promising effects on academic performance (Morisano et al., 2010; Yeager and Walton, 2011; Cohen and Garcia, 2014).

However, recent work from economics that replicates these interventions on a large number of students failed to produce similar results (Oreopoulos and Petronijevic, 2018, 2019; Oreopoulos et al., 2022).[3] Similarly, other low-touch interventions provided on a more continuous basis such as virtual or in-person but non-proactive coaching tend to be ineffective (Oreopoulos and Petronijevic, 2018, 2019; Dobronyi et al., 2019; Oreopoulos et al., 2020).[4] On the other end of the spectrum are high-touch intensive programs that are delivered in person and provide students with academic supports. These include one-on-one, continuous and proactive coaching and advising (Bettinger and Baker, 2014; Barr and Castleman, 2018; Canaan et al., 2022), as well as comprehensive programs which offer an array of structured student supports such as one-on-one advising, financial aid, weekly tutoring and referral to social services (Clotfelter et al., 2018; Page et al., 2019; Weiss et al., 2019; Evans et al., 2020). These programs are quite effective at producing long-lasting improvements in academic performance and degree attainment, but are typically expensive and difficult to scale.

Within this literature, a number of studies target students on academic probation and reach similar conclusions. Low-touch interventions are either ineffective or their benefits cannot be replicated in large samples. Moss and Yeaton (2015) show that including a message encouraging students to use academic support services as part of the probation letter has no significant effect on GPA. Brady (2017) sent students a psychologically attuned probation letter instead of the standard letter. She finds that for one cohort of students, this improved academic standing and reduced dropout rates after one year, but she could not replicate her findings using a larger sample of students from the same university.[5] Weiss et al. (2011) had students on probation attend a one-semester course to help them with study skills and setting goals, and required them to visit an academic support center. Students were in better academic standing while the program was offered, but this initial improvement faded afterwards. Albert and Wozny (2022) evaluate a more intensive program, which requires U.S. Air Force Academy probation students to work on goal-setting, attend weekly supervised study times and restricts their ability to leave campus. The program improved their GPA and graduation from certain majors.

Our paper contributes to this literature in several ways. First, our coaching program is unique in that it targets non-cognitive skills but is administered in person. Indeed, while previous low-touch interventions target non-cognitive skills and incorporate elements that are included in our program (such as goal-setting, time management and mindset), they are typically administered virtually

[3] Morisano et al. (2010) show that a one-time 2.5 hours long online goal-setting exercise improved college students' academic performance after 4 months. Oreopoulos and Petronijevic (2018) find that a similar exercise has no significant effects on performance in a larger sample of students. Further, Oreopoulos and Petronijevic (2019) show that several online "mindset" interventions, aimed at promoting a positive mindset, are ineffective among college students. For example, in their social-belonging mindset intervention, students are asked to read short stories by upper-year students describing their challenges during their transition to college and emphasizing that these difficulties are normal.

[4] Coaching seems to be most effective when it is proactive. Previous work finds that non-proactive coaching has limited positive effects on students' academic performance, with effects dissipating once the intervention ends (Angrist et al., 2009; Scrivener and Weiss, 2009; Angrist et al., 2014)

[5] Waltenbury et al. (2018) implement a similar intervention at Mohawk college and show that it increased students' GPA in the subsequent term but caution that their findings are not causal.

and lack the "personal" element—which is potentially why they are ineffective. On the other hand, high-touch interventions target students' academic skills, but do not focus on non-cognitive skills. While effective, these programs are intensive, expensive and often difficult to scale (Oreopoulos and Petronijevic, 2019). Our program is relatively lower touch and is delivered in a shorter time span (one quarter) than other successful interventions. Students in our setting are required to attend only one group coaching workshop, and hold one individualized meeting with their coach or visit an on-campus resource. In contrast, a common feature of all the aforementioned effective programs is that coaches or advisors are proactive, and they regularly initiate contact and schedule one-on-one meetings with their students. Furthermore, most of these interventions are implemented for at least one academic year, with some lasting multiple years.[6] Instead, our findings highlight that a coaching program targeting non-cognitive skills (instead of academic skills) but delivered in person (rather than virtually) can be a lower-cost and potentially more scalable way to effectively improve academically-struggling students' outcomes. Understanding which student supports are most cost-effective is important as per-student resources at U.S. postsecondary institutions have been declining over time (Bound et al., 2010; Denning et al., 2021).

Second, beyond the nature of the program, we add to this literature by providing the first estimates of the impact of coaching on later earnings. The existing literature focuses solely on academic outcomes, but understanding if treatment effects persist in the labor market provides a more comprehensive view of the benefits of coaching. While we cannot conclusively establish that *all* program participants realize labor market gains, we nonetheless show that our coaching program substantially increases earnings of students who benefit the most from it academically.

Finally, our findings provide new insights into the impacts of academic probation. Several previous studies document that academic probation increases first-year dropout rates and reduces college graduation rates (Lindo et al., 2010; Fletcher and Tokmouline, 2018; Casey et al., 2018; Dong, 2019). Our results validate the concern that an unintended consequence of academic probation is inefficient attrition. Indeed, probation appears to push out some students who would benefit from remaining in the more selective university. We find this is particularly the case for students from lower income backgrounds. They incur the largest penalties from probation and also experience large long-term gains from the coaching program. This is unsurprising given that high-achieving low-income students are often susceptible to non-financial barriers which impede their degree completion. While academic probation is likely here to stay as universities need a way to warn students who are performing inadequately, the coaching program under study is an effective

[6] Specifically, Barr and Castleman (2018) report increases in college enrollment and persistence in students enrolled in Bottom Line (BL), a program which offers intensive counseling to students starting their senior year in high school and up to 6 years after high school. BL counselors meet one-on-one with first-year college students around three to four times per semester. Bettinger and Baker (2014) document a rise in persistence from InsideTrack, a for-profit coaching service offered to non-traditional college students where coaches regularly initiate contact via phone to keep in touch with students. Oreopoulos and Petronijevic (2018) implement a one-year coaching intervention at a Canadian University in which they instruct coaches to be proactive and offer personalized regular support. Their coaching intervention substantially increased academic performance.

and low-cost way to remedy this inefficient attrition. The intervention comes at a pivotal time, and we find it greatly changes the long-term trajectories of these students.

The rest of this paper is organized as follows: Section 2 details the institutional background. Section 3 provides a theoretical framework. Section 4 describes the data. Section 5 outlines the empirical framework. Section 6 presents results. Section 7 discusses the plausibility of the identifying assumptions. Section 8 presents labor market results and section 9 highlights the mechanisms. Section 10 concludes.

2 Background

The setting for this study is a large, selective, public 4-year university located on the Central Coast of California. The university serves approximately 21,000 students, with a focus on undergraduate education, particularly in engineering and agriculture fields. To provide context, for the 2019–20 academic year, the undergraduate acceptance rate was 28%, and tuition and fees (excluding books, supplies, room/board, etc.) totaled nearly \$10,000/year for California residents and \$25,000 for out-of-state and international students. The 2019 entering cohort of first-time freshmen had an average high school GPA of 4.1 (on a 5.0 scale), an average SAT score of 1,375 (among the top 20% nationally), and an average ACT score of 29.

2.1 Success Program

Like many institutions, this university is concerned with retention and completion rates. Not only are these factors important inputs in university rankings, but administrators are also aware that it is costly to students both directly and indirectly to begin and not complete a degree, as they are unable to realize the associated wage premium and often accumulate student loans. In an attempt to improve these outcomes, in 2009 the university introduced the Success Program (SP), a mandatory academic coaching program for first-time freshmen who are placed on academic probation at the end of the fall quarter of their first year.[7] At this university, it has always been the case that students are placed on academic probation if their term GPA or cumulative GPA falls below 2.0.[8] While the intent of academic probation is to serve as a warning or motivate students to improve their performance, there is a concern that it is ineffective or may even discourage some students. In light of this concern, SP was intentionally designed to take a different approach from the standard academic probation corrective measure. SP curriculum is designed based on psychologist Dr. Albert Bandura's Theory of Self-Efficacy (1977). The theory posits that individuals with sufficient levels of self-efficacy have confidence in their ability to exert control over their motivation and behavior and, consequently, are able to achieve specific performance benchmarks. As such, SP aims to improve self-efficacy.

[7] To preserve institution anonymity, we have modified the name of the program to "Success Program".

[8] Students on probation are subject to dismissal if their cumulative GPA does not exceed the 2.0 threshold by the end of their first year.

The program was first implemented as a pilot in the fall of 2009 for four of the university's six colleges, and was extended to the entire university the following year.[9] This institution operates on a quarter calendar where three eleven-week terms make up the academic year: fall quarter, winter quarter and spring quarter (Q1–Q3). Students who earn below a 2.0 GPA in Q1—their first fall quarter at the institution—are required to complete SP during Q2 (winter quarter). The program consists of two parts: (1) a two-hour workshop led by trained faculty coaches that is held during the first two weeks of Q2, and (2) a mandatory campus engagement assignment to be completed by week 5 of that quarter. For this assignment, students choose between either visiting a campus resource and completing a reflection assignment, or attending a one-on-one follow-up meeting with their workshop coach and completing a reflection assignment. The university typically offers two workshop dates during the first weeks of Q2 to accommodate student and faculty schedules. Roughly 320 students qualify for the program each fall quarter, and almost all of them participate in one of the two workshop sessions. A small share of students are unable to attend the session due to a scheduling conflict and complete the requirement one-on-one with a trained coach. To enforce participation, students are unable to enroll in Q3 (spring quarter) courses until they have completed all parts of the program.

Bandura (1977) posits that expectations of efficacy are affected by four main sources of information: mastery of experiences, vicarious experiences, physiological arousal, and verbal persuasion. Mastery of experience refers to the idea that individuals can increase their self-efficacy if they undertake and succeed in a task repeatedly. This can start by setting goals then working through and achieving those goals. Physiological arousal implies that negative feelings, such as being anxious about undertaking a task, can reduce self-efficacy. Vicarious experiences implies that seeing other people with similar characteristics succeed in a task, will boost an individual's confidence to undertake the task. Verbal persuasion implies that if individuals are told by someone they trust that they are able to succeed in a certain task, they will be more willing to undertake it. Vicarious experiences and verbal persuasion imply that a trusted mentor can boost self-efficacy through being a role model and providing encouragement.

SP was designed to expose students to these information sources in a way that boosts self-efficacy. The two-hour workshop is broken into two parts: a thirty-minute meeting with all session participants (typically about 150 students), followed by a ninety-minute breakout session with a smaller group of 6 to 8 students led by one or two trained faculty or staff coaches. One of the goals of the large group portion of the workshop is to normalize failure and academic probation, and show students that they are not the only ones who have experienced failure and that it is possible to succeed despite failing. To do so, students watch a video featuring various high-profile people in Silicon Valley (e.g., Elon Musk) who have overcome challenges and failure. Within the self-efficacy framework, this is intended to expose students to vicarious experiences. In other

[9] One might be concerned that the pilot colleges in 2009 are non-randomly chosen and thus are driving our results. We show in Table B8 that the results are robust to the exclusion of the 2009 cohort.

words, when students see other individuals—with whom they share similarities i.e., in this case failure—succeed, this will improve their beliefs about their own probability of success. Another goal of the large group portion of the workshop is to remove some of the anxiety surrounding academic probation. This is based on the idea that negative physiological arousal such as feelings of anxiety or stress can decrease self-efficacy. To do so, the SP leadership team presents students with information on campus resources, including tutoring services, health and well-being services, counseling services, and cross-cultural services such as the Gender Equity Center, Pride Center and Multicultural Center (see Appendix C). They also outline the rules of academic probation in an attempt to remove some of the anxiety surrounding this term.

The second part of the workshop is meant to be more interactive and discussion-based. It involves a breakout session with a smaller group, typically 6–8 students, led by one or two trained coach. Coaches are faculty and staff from across the university who have undergone a two-hour training led by the SP leadership team.[10] In an attempt to reiterate the message that failure is common among successful people and expose students to vicarious experiences, the session opens with coaches sharing a time when they experienced failure. The coaches then facilitate a reflection and goal-setting exercise. Students are allotted time to reflect on factors that may have contributed to their academic struggles in Q1, such as academic challenges (e.g., problems with time management, study skills, class attendance, or school/life balance); college adjustment difficulties (e.g., roommate issues, homesickness, difficulty finding resources, or difficulty fitting in); and personal hardships (e.g., mental or physical health issues, personal or family crises, or identity-based isolation). To guide this process, participants work through a worksheet, pausing to discuss their responses with the group. The worksheet, presented in Appendix C, is set up in three steps: identifying weaknesses, identifying solutions or resources that can aid in overcoming these weaknesses, and goal-setting for the current quarter. The intent of this exercise is for students to leave the workshop with a tangible plan to change their academic trajectory going forward. During this time, participants are also required to fill in a time management worksheet where they are encouraged to allocate time for classes, studying and social activities (Appendix C).

An indirect goal of the time management and goal-setting exercises is to initiate mastery of experiences. Having students set goals gives them standards against which they can evaluate their performance and hence learn whether they are mastering tasks. Furthermore, achieving small goals may provide students with proof of mastery of tasks which in turn boosts self-efficacy (Bandura and Schunk, 1981). Finally, throughout SP, students are exposed to verbal persuasion; that is, encouragement and positive feedback from coaches. Importantly, coaches not only tell students that it is possible to succeed despite failing but they also present them with evidence of this (through showing them videos and sharing their own personal stories).

[10] Faculty and staff who become coaches select into this service role. Faculty receive service credit from their home departments for their participation, and staff's time spent on coaching is considered part of their typical work day.

2.2 Academic Probation

Academic probation is a near universal policy in higher education institutions in the US. Such a policy notifies students who have fallen below a minimum GPA threshold to remain in good academic standing and is meant to serve as a warning. Once on probation, students who do not improve their GPA in a set number of terms are then typically eligible for dismissal from the university. The policy is well-intentioned as it is meant to serve as a wake-up call for capable students who have been underperforming, and to weed out students who are "mismatched" and would be better off going to a lower-ranked university. Indeed, probation has been shown to increase some students' chances of dropping out of college and to reduce graduation rates (Lindo et al., 2010). This is efficient if it only removes students who would be better off going to a lower-ranked institution. The concern is that this label may unintentionally weed out capable students from lower income backgrounds as they often face barriers to completion that their higher-income counterparts do not.

At this institution, the academic probation threshold is determined by the minimum GPA of 2.0, which is a "C" average. At the end of each quarter, students who have a cumulative GPA or a term GPA below 2.0 are placed on academic probation and are promptly notified of this status via email. Students remain on probation until their term and cumulative GPA are above the 2.0 threshold. First-year students are granted a probationary period of the first year to improve their GPA above this threshold. If they fail to meet this mark by the end of their first year (Q3), they are subject to dismissal and are not eligible to return for their second year.[11]

3 Theoretical Predictions

The coaching program has multiple components such as conveying to students that it is possible to succeed despite failing, requiring them to use a campus resource and helping them with goal-setting and time management. A priori however, it is unclear which of these components—if any—is most effective at improving student outcomes. In this section, we aim to *theoretically* understand which elements of the program have the potential to reduce dropout rates amongst students on academic probation. To do so, we rely on Bénabou and Tirole (2003)'s model of extrinsic and intrinsic motivation. The model is a game between two players: a principal and an agent. In our setting, the principal is the university and the agent is the student. The model delineates the interplay between an agent (student) who has imperfect information about her ability and a principal (university) who offers the agent an incentive structure (such as help or encouragement) in order for her to undertake a task (such as increasing her grades or persisting at the university). The model allows us to derive predictions regarding how different incentive structures affect the student's self-confidence and motivation, and subsequently her effort and performance.

[11] Students who are placed on academic probation in any other quarter than Q1, are granted only one quarter to improve their academic standing above the 2.0 threshold before being eligible for dismissal.

We first set up the general model and discuss how it applies to our setting. We classify the different elements of our coaching program into two types of incentive structures: help and encouragement. We then review the predictions of Bénabou and Tirole (2003) regarding how these two incentive structures affect students.

3.1 Model Setup

The model is set up as a game between the university (principal) and the student (agent). A student who scores just below the 2.0 GPA cutoff is on academic probation and faces two choices: she can either decide to improve her academic performance and not drop out from the university, or drop out from the university. In other words, she can choose whether or not to exert effort $e \in \{0;1\}$ in order to not drop out from the university. The cost of exerting effort or undertaking the task of not dropping out is $c > 0$. The student and university's payoffs are respectively $V > 0$ and $W > 0$ if the student is successful and does not drop out. If the student drops out, both get 0.

Let $\theta \in [0;1]$ denote the student's probability of success if she decides to exert effort. θ can alternatively be viewed as the student's ability. A key assumption in the model is that the university has more information than the student about her probability of success or ability. Specifically, the university is assumed to perfectly know θ, while the student knows that θ is drawn from a cumulative distribution function $F(\theta)$ with a density $f(\theta)$.

In our setting, both the university and the student know that the student is on academic probation, but the university has more information about the student's probability of success. It can derive this information from having seen the outcomes of other students on academic probation in previous years. For example, the university might think that it is still possible for the student to succeed despite her initial failure, as previous probation students have persisted at the university. On the other hand, it is possible that the university is less trusting of the student or more pessimistic about her chances of success because they have seen previous students be discouraged by academic probation. In both cases, the university is assumed to have superiority of information.

The student may receive a signal $\sigma \in [0;1]$ about her probability of success θ, with conditional cumulative distribution $G(\sigma|\theta)$ and positive conditional density $g(\sigma|\theta)$. Another key assumption is that the monotone likelihood ratio property (MLRP) holds:

$$\text{for all } \sigma_1 \text{ and } \sigma_2 \text{ with } \sigma_1 > \sigma_2, \ \frac{g(\sigma_1|\theta)}{g(\sigma_2|\theta)} \text{ is increasing in } \theta.$$

Intuitively, the MLRP guarantees that a higher σ is "good news" for the student about her probability of success θ.

The timing of the game is as follows:

Stage 1: The university learns the parameter θ and selects a policy p.[12]

[12] For example, if the university's incentive structure is to offer the student help, the policy p will be the level of help provided.

Stage 2: After observing the policy chosen by the university and learning σ, the student chooses action e.

The student's expected payoff from exerting effort is $\hat{\theta}(\sigma, p)\mathrm{V}$ which can also be viewed as student's intrinsic motivation. Her expected payoff depends on $\hat{\theta}(\sigma, p)$, which is the conditional expectation of her probability of success or her interim self-confidence. The conditioning of $\hat{\theta}$ on p, referred to as the "looking-glass-self", implies that the student tries to see through the university's ulterior motive from choosing p. In other words, the choice of policy may reveal private information that the university has about the student's ability.

Bénabou and Tirole (2003) analyze the perfect Bayesian equilibria of this two–stage game for different types of incentive structures (such as offering rewards, bonuses, help, delegation and encouragement). The intuition behind solving for the equilibrium is as follows:

For a given policy p, the student will only exert effort e if her expected payoff from it is greater than its cost i.e., if $\hat{\theta}(\sigma, p)\mathrm{V} \geq c$. Therefore, there is a cutoff $\sigma^*(p)$ such that the student exerts effort if and only if she receives a signal $\sigma \geq \sigma^*(p)$—which depends on p due to the "looking-glass-self" phenomenon. In other words, $\sigma^*(p)$ is the threshold signal for which the student's expected payoff from exerting effort is equal to its cost, so that any $\sigma \geq \sigma^*(p)$ implies that the expected benefits from effort are higher than its costs.

The university then chooses p to maximize its own expected payoff. The university's choice of p depends on θ and $\sigma^*(p)$. The university might therefore choose to offer different policies to students of different ability levels.[13,14] Observing the university's equilibrium strategy, the student will update her beliefs about $\hat{\theta}$ as well as her expected benefit from exerting effort and subsequently her threshold for exerting effort $\sigma^*(p)$. In other words, the student will infer from the chosen policy information about her true ability, which affects her self-confidence $\hat{\theta}$ and intrinsic motivation (i.e., $\hat{\theta}\mathrm{V}$). This will subsequently determine her optimal action e^* (i.e., whether or not she exerts effort). In most settings, performance is produced by the interaction between effort e and ability θ which are complements. This implies that increasing (lowering) self-confidence should increase (decrease) effort.

The coaching program in our setting includes a bundle of interventions, which can be classified under one of two incentive structures: help and encouragement. Specifically, the program offers students "help" by providing them with information about probation and campus resources, having coaches assist them with time management and goal-setting exercises, and requiring them to meet once individually with their coach or use a campus resource. The other key component of the program is to show students that failure is normal and that it is still possible to succeed despite failure.

[13]Formally, this follows from the university's maximization problem, the MLRP and additional assumptions that can be imposed depending on the type of incentive structure. For example, as discussed below, when the university is deciding on the level of help to offer students, an additional sorting condition will be imposed. Along with the MLRP, this will make it such that the university offers more help to lower-ability students.

[14]Another assumption is that the university is uncertain about how the student will respond to their choice of policy or how it will affect his intrinsic motivation. Otherwise, they can perfectly predict the student's response to any chosen policy.

In what follows, we argue that the university's incentive structure in this case is encouragement or disclosing to the student "good news" about her probability of success. Following Bénabou and Tirole (2003), we reframe the model and derive its predictions under these two incentive structures.

3.2 Help

Assume that the university wants to provide the student with a level of help h (at a private cost h) in case the student decides to exert effort to remain at the university. The probability of success is now defined as the function $P(\theta, h)$, which is increasing in h and θ i.e., $P_h > 0$ and $P_\theta > 0$. As discussed earlier, the student only exerts effort if and only if she receives a signal $\sigma \geq \sigma^*(h)$ where $E[P(\theta, h)|\sigma^*(h), h]V = c$, with $\sigma^*(h)$ being a decreasing function. The university's payoff is $U_P = [1 - G(\sigma^*(h)|\theta)][P(\theta, h)W - h]$.[15]

In the case of help, Bénabou and Tirole (2003) impose a sorting condition:

$$\frac{\partial^2 ln(P(\theta,h)W-h)}{\partial\theta\partial h} < 0$$

The sorting condition means that the percentage increase in the university's payoff due to a higher level of help is lower when the student's ability or probability of success is higher. Bénabou and Tirole (2000) argue that this sorting condition is suitable when we have a zero-one task such as not dropping out of the university. This is because the university's payoff would be independent of the student's margin of success.[16]

In equilibrium, one can show that the university's optimal policy is to offer a higher level of help for students with lower probability of success.[17] Because the university helps less when it is more trusting of the student, a higher level of help is perceived by the student to be "bad news" about her probability of success. This lowers her self-confidence $\hat{\theta}$ and intrinsic motivation $\hat{\theta}V$, which in turn increases $\sigma^*(h)$. This higher threshold effectively decreases the likelihood of exerting effort to not drop out from the university. In summary, help lowers student's self-confidence, motivation and effort, as the offer of help reveals to her that she is low-ability.

3.3 Encouragement

A key element of the coaching program is to communicate to students on probation that they are not the only ones struggling academically, failure is normal and it is still possible to succeed despite probation. Initially, students only see the probation label and as a result, they may think that their probability of success is low. The university however knows that it is still possible to succeed despite probation. In other words, the university has private positive information about

[15] $[1 - G(\sigma^*(h)|\theta)]$ is the probability that the student exerts effort.

[16] A reversal of the sorting condition would occur in cases where the principal's payoff increases with the margin of success and as a result with the agent's ability. For example, if a professor is coauthoring papers with students, she would be more tempted to help high-ability students as her payoff depends on the margin of success (i.e., the paper's publication outcome).

[17] This follows from the sorting condition and the MLRP.

student's probability of success. Under the Bénabou and Tirole (2003) model, the university's policy p here is whether or not to disclose to the student this private information (i.e., whether or not to provide students with encouragement). Importantly, the information to be disclosed by the university is "hard" meaning that students are presented with *evidence* supporting the claim that they can still succeed despite failure.[18] Indeed, the program offers students three pieces of evidence. First, at the beginning of the workshop, students watch a video of various high profile individuals who have succeeded despite previously failing. Second, during the breakout sessions, coaches share their own experience with failure. Third, conducting the workshop in a group setting also provides students with evidence that they are not the only ones on probation.

To understand the implications of encouragement, we look to the model in Section 3.1. Recall that the university's policy is whether or not to disclose positive private information about student's probability of success θ. It can be shown that because this private information is "good news" about the student's probability of success, the university's optimal policy would be to disclose the information or provide encouragement.[19] Formally, the university's policy here would be to release a signal that covaries positively with θ. Because the student is receiving a positive signal about her probability of success, her self-confidence and intrinsic motivation increase. The boost in self-confidence will in turn increase the likelihood of exerting effort towards persisting at the university.

To summarize, this model predicts that in our context, disclosing "good" news with hard evidence about the student's probability of succeeding (or encouragement with evidence) increases her motivation and self-confidence. This in turn pushes the student to work harder towards the goal of persisting at the university. On the other hand, offering a student help lowers her self-confidence and motivation, which induces her to work less. The coaching program we focus on offers both help and disclosure of information. As a result, it is unclear whether, ex-ante, the program is expected to improve student performance and persistence.

4 Data

All data in the main analysis are student level and come from three sources: (1) administrative records from the university office of Institutional Research, (2) SP participation files, and (3) pre- and post-SP survey responses from the Success Program office, which are discussed in detail in Section 9.2. The full sample includes eleven entering cohorts of first-time freshmen who enrolled at the university in a fall quarter between 2007 and 2017 (45,864 students) and tracks them by quarter through graduation or until they separate from the university. In an auxiliary analysis, these data are then augmented with labor market outcomes from the California Employment Development

[18] This is important because if the information is "soft" (for example, encouragement or praise without evidence), the student will see through the university's ulterior motive of simply wanting to boost her self-confidence and may not believe the information (i.e., cheap talk).

[19] Conversely, the university would want to conceal "bad news" about the student's probability of success.

Department and are described in Section 8.

The administrative transcript files provide a detailed view of students' academic progress, allowing for examination of a rich set of outcomes. By student-quarter, we observe enrolled courses, course grades, cumulative GPA, academic major, probation status, and the timing of separation from the university either as a dropout or graduate. As noted above, Q1 is defined as a student's first fall quarter at the university, and Q2 and Q3 correspond to their first winter and spring quarters, respectively. Understanding student performance in year one is of primary interest as students qualify for the coaching program in Q1 and complete it in Q2. As such, the main outcome of interest is dropping out at the end of Q3 (i.e., year one retention). Q3 dropout takes the value of 1 if a student does not appear in the data the following academic year (Q5–Q7), and 0 otherwise.[20] Q1 and Q2 dropout are coded in a similar fashion. Other outcomes of interest include graduating: whether a student graduates in 4 years (on-time graduation) or 6 years (a proxy for ever graduating). We construct two additional outcomes, Q2 + Q3 GPA and Q2 + Q3 total credits earned, to capture a student's academic performance in the rest of their first year.[21] These files also contain a rich set of time-invariant background characteristics including a student's high school GPA, gender, race/ethnicity, whether they are required to enroll in remedial math and English courses, eligibility for the Federal Pell Grant program, their expected family contribution (EFC) as determined by the Free Application for Federal Student Aid (FAFSA), and parental education.

Probation status is observed in the administrative files and takes the value of 1 if a student is placed on academic probation in Q1, and 0 otherwise. Based on the probation policy assignment rule, students who score below a 2.0 GPA in a given quarter are placed on probation. For this analysis, Q1 probation status is of primary interest. Figure A1 confirms that the probability of Q1 probation is solely determined by Q1 GPA for both treated and control cohorts. The probability of probation changes sharply at the 2.0 threshold.[22]

The administrative files are merged with SP participation records to identify which students complete the program. Recall that SP assignment is determined by a Q1 GPA of less than 2.0. Figure 1 shows the likelihood of program participation by Q1 GPA for all entering cohorts from fall 2010 to fall 2018 (i.e., since the inception of the coaching program).[23] Indeed, there is a sharp jump at the 2.0 GPA threshold in the likelihood of SP participation. No student with a Q1 GPA above the threshold of 2.0 participates in SP while virtually all students scoring below the 2.0 cutoff

[20]This coding of Q3 dropout allows a student to take several quarters away, perhaps to study abroad or for employment reasons, and not be coded as a dropout. The results, however, are robust to alternative ways of defining Q3 dropout including coding Q3 equal to 1 if a student never appears in the data again following Q3 or if they don't appear in the fall quarter of their second year (Q5).

[21]While GPA is a common measure of performance used in the literature, it is only defined for students who are enrolled in Q2 and Q3. As a way to circumvent this selection issue, we also analyze total credits earned in Q2 and Q3, as that is defined for all students in the sample.

[22]Unfortunately, the probation variable indicator is missing for the 2010–2016 entering fall cohorts. As such, our first-stage analysis is based on the freshmen 2007, 2008, 2009, and 2017 fall cohorts.

[23]We exclude from this figure students from the few faculties who were selected for the SP pilot in 2009 as we do not have program participation data for them.

participate.[24] Together, Figure A1 and Figure 1 confirm that probation status and SP participation are binding in practice and are solely a function of Q1 GPA.

Summary statistics are presented in Table 1.[25] While column 1 presents means for the full sample to provide context, column 2 reports summary statistics for the 22,225 students who are part of the analysis sample (i.e., the marginal students with GPAs between 1 and 3 in their first quarter at the university). Relative to the full sample, the analysis sample has more men (56% vs. 52%), more non-white students (42% vs. 38%), and is lower income. The share of Pell Grant-eligible students is 19% compared with 16% in the full sample. Moreover, this group experiences relatively worse academic outcomes. More than half of the students in the analysis sample are placed on academic probation at some point during their college career, and 9% dropout after their first year. Consequently, the 4- and 6-year graduation rates for the students around the 2.0 GPA cutoff are 35% and 80%, respectively—much lower than the full sample.

Columns 3 and 4 report summary statistics for the group of students in the control cohorts. This constitutes a sample of 4,628 students enrolled in all colleges in 2007 and 2008 and the 2 colleges in 2009 that did not participate in the pilot program.[26] Column 3 includes the students who are placed on probation, GPA $\in [1, 2)$, and column 4 includes students scoring just above the 2.0 GPA cutoff, GPA $\in [2, 3]$, and thus not on probation. Columns 5 and 6 present summary statistics for the 17,597 students in the treated cohorts. We split this sample into those eligible for both programs, GPA $\in [1, 2)$, as presented in column 5, and those who are barely ineligible, GPA $\in [2, 3]$, as presented in column 6.

5 Empirical Methods

5.1 Visual Motivation for DiRD Design

We begin by presenting graphical motivation for the DiRD design. Figure 2 presents regression discontinuity (RD) plots for several outcomes separately for treated and control cohorts, with first quarter GPA as the running variable. All figures take similar forms, in that circles represent local averages over a 0.1 GPA score range. All figures are drawn over a bandwidth of 1 GPA point on either side of the cutoff using a linear fit. Figures in the left panel summarize effects at the 2.0 GPA cutoff for students in control cohorts (exposed to probation-only policy), while those on the right present effects for those in treated cohorts (exposed to probation + SP).

Figure 2 provides visual evidence of meaningful differences between students exposed exclusively to the probation policy (control cohorts) and those exposed to probation and SP (treated cohorts).

[24]Each fall quarter there are a few students (less than 10) who qualify for the SP but who are granted a waiver by the dean of their college and are thus excused from participating. These waivers are typically reserved for extenuating circumstances such as health shocks or family emergencies.

[25]Summary statistics by various subgroups are reported in Table B1.

[26]The following colleges participated in the 2009 pilot: College of Agriculture, Food, and Environmental Sciences; College of Business; College of Engineering; and College of Architecture and Environmental Design. The College of Science and Mathematics and College of Liberal Arts did not participate.

Figure 2a and Figure 2b highlight significant first-year dropout differences between control and treated cohorts at the cutoff. Students who just qualify for probation are 10.9 pp more likely to drop out by the end of the first year compared to those who just avoid probation. This large and meaningful gap at the cutoff is greatly reduced to a statistically insignificant 2 pp for cohorts exposed to probation and SP. Under the assumption that the negative effects of probation are similar for all cohorts, this suggests that students who qualified for the coaching program experienced significantly lower dropout rates compared with those not exposed.

While there is not strong visual evidence that SP changes students' probability of graduating on time, comparing Figure 2c and Figure 2d, there is for 6-year graduation as shown in Figure 2e and Figure 2f. We find that students in control cohorts with GPAs just below the threshold are 9.1 pp less likely to graduate in 6 years compared with treated cohorts who are only 4.7 pp less likely. Finally, Figure 2g and Figure 2h show that while students exposed only to probation were not significantly affected in terms of first year GPA, those exposed to probation and SP experienced a large and significant 16% of a standard deviation increase in performance at the cutoff. That is, the coaching program seems to have positive grade impacts on marginal students. Figure A2 shows the results are robust to using a quadratic fit on either side of the cutoff. Additionally, Figures A3 and A4 present results for subgroups of students by gender, major and socioeconomic status. While this exercise provides suggestive evidence of positive impacts of the program, we next turn to an econometric framework using a DiRD design to more rigorously probe this possibility.

5.2 Difference-in-Discontinuity Design

To identify the causal effect of coaching, we draw on variation in exposure to the coaching program within cohorts and across cohorts. First, in the spirit of an RD design, we leverage variation in SP participation that arises from the strictly enforced SP assignment rule which is a function of first quarter GPA at the university. Students who score below a 2.0 GPA in Q1 are required to complete SP and those above the threshold are excluded from the program. In a standard RD framework, if this cutoff is orthogonal to student characteristics, any observed discontinuity in outcomes around the threshold can be attributed to SP. However, because the SP assignment rule is identical to the academic probation assignment rule, the interpretation of the standard RD estimate will capture both the effect of SP and probation. To isolate the effect of SP net of the confounding probation policy, we further leverage the fact that some cohorts were exposed to SP and others were not. The SP pilot was introduced in the 2009 academic year for a subset of colleges at the university, and for all cohorts in all colleges from 2010 to the present. Thus, the 2007 and 2008 cohorts and some students in the 2009 cohort (control cohorts) were exposed to the probation rule but not to the SP assignment rule. All other cohorts were exposed to SP and probation (treated cohorts).

Formally, we implement a DiRD design. Intuitively, this design estimates the discontinuity in outcomes across the 2.0 GPA cutoff for cohorts exposed to both SP and probation and then purges the effects of probation by differencing out any discontinuity in outcomes for cohorts exposed to the probation policy only.

The estimation equation is as follows:

$$\begin{aligned} Y_i = {} & \beta_1 + \beta_2 GPA_i + \beta_3 Treat_i + \beta_4 Below_i + \beta_5 (Treat_i * GPA_i) \\ & + \beta_6 (Below_i * GPA_i) + \beta_7 (Below_i * Treat_i) + \beta_8 (Below_i * Treat_i * GPA_i) \\ & + \rho_c + \delta_k + \gamma X_i + \epsilon_i, \end{aligned} \tag{1}$$

Y_i is the outcome of interest for student i. GPA_i is the running variable and represents student i's normalized first quarter GPA relative to the cutoff of 2.0. $Treat_i$ is a binary variable that takes the value of 1 for treated cohorts, corresponding to all students exposed to both the probation policy and SP, and 0 for control cohorts exposed only to the probation policy. *Below* is a binary variable that takes the value of 1 for students scoring below the GPA cutoff of 2.0, and 0 otherwise. The interactions with GPA_i allow slopes to vary on either side of the GPA cutoff as well as across treated and control cohorts.

The parameter of interest is β_7, which represents the difference between treated and control cohorts in the discontinuous jump in the outcome at the 2.0 GPA cutoff.[27] X_i is a vector of controls composed of students' predetermined characteristics—high school GPA, gender, race, required enrollment in remedial math and English courses, Pell Grant eligibility, EFC and parental education—and is included to improve precision by reducing residual variation in the outcome variable. ρ_c is cohort fixed effects which control for any common shocks and overall trends in the outcome. δ_k is college fixed effects. Finally, ϵ_i represents the error term. We report robust standard errors rather than clustering. Clustering with a discrete running variable yields confidence intervals with worse coverage properties and does not resolve specification bias issues (Kolesár and Rothe, 2018).

Unfortunately, the standard data-driven bandwidth selectors often implemented with a RD design do not extend to a DiRD research design. As such, in the main tables of results, we report estimates using bandwidths for the running variable of 0.75 and 1.0 GPA points on either side of the cutoff. To further probe the sensitivity of the results to bandwidth choice, following Jackson (2021) Figure A5 reports point estimates across a variety of bandwidths ranging from 0.25 to 2 GPA points. Additionally, we compute the optimal bandwidth separately for treated and untreated cohorts using the CCT procedure described in Calonico et al. (2014), and then estimate Equation (1) using the average of the two optimal bandwidths, as is done in Grembi et al. (2016). Figure A5 denotes this bandwidth for each outcome with a vertical line.

As formalized in Grembi et al. (2016), when both policies induce sharp RDs the validity of the DiRD estimate, β_7, requires that the following two identifying assumptions hold.

A1. Potential outcomes are smooth across the threshold (standard RD assumption).

[27] The parameter β_3 summarizes the average difference in outcomes for students scoring above the 2.0 cutoff in the treated versus control cohorts. The parameter β_4 represents the average difference in outcomes for students scoring below versus those scoring above the cutoff in the control cohorts.

A2. The effect of the confounding policy, probation, is constant over time (akin to the DiD parallel trends assumption).

Section 7 presents several pieces of evidence in strong support of A1 and A2.

Finally, Grembi et al. (2016) presents a third assumption, LATEs of both polices (probation and SP) are additively separable. That is, the main policy does not interact with the confounding policy. This assumption is not required to obtain an internally valid DiRD estimate, but rather it shapes its interpretation. An intuitive way to assess this requirement in our setting is to ask, would the coaching program generate the same effect if it were mandatory for students around the 2.0 cutoff but who were not also on probation? We are unable to test this empirically as we do not observe coaching absent of probation. Consequently, we interpret our findings as the effect of coaching in the presence of the probation policy. Given that some form of academic probation is universal policy in the US, our results have broad implications.

6 Results

6.1 Academic Results

The main results come from estimating Equation (1). Table 2 reports the point estimates using two different bandwidths, 0.75 and 1 GPA points. All estimates are reported with and without controls to ensure results are robust to the inclusion of predetermined student characteristics. Columns 1 through 3 present DiRD estimates for Quarter 1 (Q1), Quarter 2 (Q2) and Quarter 3 (Q3) dropout, respectively. As shown in columns 1 and 2, we find no significant treatment effects for Q1 or Q2 dropout indicating that SP had no effect on dropout *during* the first year. The null effect on Q1 and Q2 dropout is reassuring given the timing of the program and the probationary period for first-year students; program participation is in Q2 and university policy states that a student is not subject to dismissal for poor academic performance until the end of the first year (Q3). On the other hand, we find large and statistically significant treatment estimates on dropout directly following first year (column 3). SP decreases marginal students' Q3 dropout by 7.3–8.8 pp, an approximate 37% decrease from the baseline dropout rate of 23% for students placed on probation in pre-program cohorts with a Q1 GPA $\in [1.75, 2.0)$.

Consistent with the improved first-year retention, SP also positively impacts academic performance in the rest of the first year. Column 4 shows standardized Q2 + Q3 GPA improves by a large and significant 14.6–16.8% of a standard deviation, depending on bandwidth choice.[28] Column 5 reveals that the program increases total earned credits in Q2 and Q3 by approximately 1.2, which is about 5% from a baseline mean of 22.77.

[28] For comparison purposes, we standardize Q2 + Q3 GPA within cohort to mean zero, standard deviation 1.

Columns 6 and 7 report treatment effects for graduation outcomes. The program does not appear to improve on-time graduation, as the point estimates for 4-year graduation are small and indistinguishable from zero (column 6). There is suggestive evidence, however, that 6-year graduation (as defined by graduating within 6 years from entry) is affected (column 7). While the point estimates are imprecise, they are positive and relatively large in magnitude. Given that the program targets students lower down in the grade distribution, it is not surprising that it has no effect on the likelihood of on-time graduation. A more plausible story, one in line with our findings, is that the program reduces dropout at the end of the first year (as shown in column 3), and students who remain because of this, are more likely to eventually graduate than those who just missed out on the program. Finally, all estimates are robust to the inclusion of predetermined student characteristics and bandwidth choice (see Figure A5).[29]

6.2 Heterogeneity Analysis

Next, we consider the possibility that the SP may be particularly beneficial to students from lower-income backgrounds. Indeed, it has been shown in other higher educational settings that lower-SES students tend to incur long-term benefits from academic support interventions (Evans et al., 2020; Weiss et al., 2019; Barr and Castleman, 2018; Clotfelter et al., 2018). In Table 3 we report treatment effects by students' socioeconomic status (SES). Lower SES takes the value of 1 if a student files the FAFSA and is eligible for federal financial aid (i.e., they have an EFC<30,000), and 0 otherwise. Higher SES is the complement. We find that the program effects are largely driven by students from lower-income backgrounds (i.e., those who ex ante may be more likely to benefit from advising or coaching). Students from lower SES backgrounds who are marginally exposed to coaching are 12.5 pp less likely to dropout at the end of the first year, and are 14.9–16.9 pp more likely to graduate within six years, depending on the bandwidth. In line with the 6-year graduation findings, these students also experience large improvements in their Q2 + Q3 GPA and Q2 + Q3 total earned credits. Overall, the heterogeneity analysis reveals that coaching for students on the margin of the 2.0 GPA threshold has substantial impact for those who come from lower-income backgrounds.[30]

[29]Our focus is on year 1 dropout because most attrition occurs in year 1. For the full sample, year 1 dropout is 6%; year 2, 4%; year 3, 1.8% and year 4, 1.7%. Bostwick et al. (2022) find a similar dropout pattern among students attending Ohio public institutions. Moreover, Table B2 shows no detectible effects of the coaching program on dropout in years 2 and 3 suggesting that the program averts dropout among this discouraged group of students rather immediately.

[30]Table B3 and Table B4 report heterogeneity results by gender and field of study. We find that the SP is also particularly beneficial to marginal men and STEM majors. Note, at this university, students are admitted to and enroll in a specific major prior to entry, ruling out endogenous student sorting across majors. See policy here: https://www.calpoly.edu/admissions/first-year-student/selection-criteria. Also, to the extent that STEM students around the threshold have different academic ability than non-STEM students, their responsiveness to SP may differ.

7 Validity of the Research Design

Standard RD assumption (A1). The identifying assumption required for a valid RD design is that individuals are not able to manipulate the running variable. If individuals can influence which side of the cutoff they are on, it will call into question the causal interpretation of the point estimates as it will be difficult to distinguish between student sorting and the true effect of the intervention. In our setting, this could occur if instructors or students strategically manipulate grades in such a manner that the distribution of observable and/or unobservable characteristics of students are discontinuous at the 2.0 GPA cutoff. Although this is a fundamentally untestable assumption, we provide several indirect tests that support its plausibility.

First, it is unlikely that instructors would be able to strategically manipulate a student's entire quarterly GPA since they are generally responsible for only one of three or four course grades. Second, we check for manipulation around the 2.0 threshold by plotting the distribution of student GPAs for all cohorts, as shown in Figure A6a. While there are two large density spikes at the GPA cutoffs of 2.0 and 3.0, these heaping patterns are similar across treatment (SP-eligible) cohorts and control (pre-SP) cohorts, as shown in Figure A6b and Figure A6c. Since overall GPA patterns did not change with the implementation of SP, any heaping will be differenced out with the DiRD research design, mitigating concerns that heaping is biasing the DiRD estimate.

More generally, heaping at round GPA points is not necessarily indicative of manipulation. It is possible that discontinuities in the GPA distribution are linked to other exogenous factors such as grade rounding. Natural, non-strategic, institutional grade bumps are common in many U.S. institutions and have been documented in GPA-based RD settings such as Zimmerman (2014) and Ost et al. (2018). To further alleviate concerns over grade heaping, Table B5 presents results from a donut DiRD design which involves dropping the heaping points at the 2.0 and 3.0 GPA cutoff following Barreca et al. (2016). This exercise yields similar results to those obtained from our main specification. We conclude that the heaping observed in our data is most likely non-strategic and due to natural grade rounding, as observed in previous studies.

Finally, we examine whether observable student characteristics evolve similarly around the 2.0 GPA threshold. If individuals are unable to manipulate the side of the threshold they fall on, we should observe no differences in predetermined characteristics across the cutoff. To implement this, we estimate a series of balance tests using Equation (1). Indeed, for the three different bandwidth windows of 0.5, 0.75, and 1 GPA points on either side of the cutoff, we find no evidence of differences in discontinuities for any of the nine observable predetermined student characteristics. Results are reported in Table 4. To summarize these effects, we construct predicted dropout and predicted first-year GPA outcomes for each student based on these nine baseline covariates and estimate our main specification. If no GPA manipulation is present, the estimates should not be statistically different from zero. The DiRD treatment estimates for these predicted outcomes are presented in Table B6 and, in fact, are statistically insignificant at the cutoff. In summary, the fact that observable student characteristics appear to be smooth across the threshold further alleviates concerns over GPA manipulation. Altogether, the findings from these empirical tests indicate that the DiRD design should purge our estimates of any such unobservable bias—assuming the unobservable differences are also constant across cohorts.

The confounding policy is constant over time (A2). For the DiRD estimate to be valid, it is necessary that the effect of the confounding policy (here, academic probation) has the same effect before and after the introduction of the policy of interest (here, SP). First, the probation policy and its implementation did not change over the sample period. We verified this by consulting each year's Student Handbook. We also spoke with administrators from the University's retention office who further confirmed this information. Second, Figure A1 provides empirical evidence that the assignment rule is the same before and after program implementation. The first stage for the probation policy is sharp around the 2.0 cutoff for treatment and control cohorts.

Another way to empirically assess the plausibility of this assumption is to analyze how the effect of probation evolves over time. If the effect is similar across the different control cohorts (i.e., no preexisting trends), then it suggests that the effect of probation is constant over time. To test for this, we separately estimate the RD coefficient for each cohort for the two main outcomes: Q3 dropout and Q2 + Q3 standardized GPA. We focus on these outcomes as this is where we document significant SP impacts. All estimates rely on a bandwidth of 1 GPA point, with the treatment defined as scoring below a 2.0 GPA.

To most easily assess the dynamics of the effect of probation, Figure 3 plots these RD estimates by cohort. Robust standard errors are reported in bars. The first three estimates, those displayed before the dashed vertical line, are probation effects for students enrolled in our two control cohorts (2007, 2008) and the two colleges that were not exposed to the pilot program in 2009.[31] All estimates after the dashed line represent probation effects for the treated cohorts (i.e., the four pilot colleges in 2009 and the 2010–2017 cohorts).

For both outcomes, the effect of probation is quite similar over time as evident among the control cohorts. In Figure 3a, the first three estimates show positive and mostly significant effects on dropout rates for students just eligible for probation alone, and these probation effects are similar across the three control cohorts. Once SP is introduced, the positive dropout effects dissipate, suggesting that SP likely has a moderating effect on probation. Figure 3b displays a similar pattern. Q2 + Q3 standardized GPA is unaffected for the three control cohorts only faced with probation, while the cohorts also exposed to SP are positively affected. Importantly, the "pre-trend" patterns in both figures seem to indicate that probation had a similar effect on outcomes regardless of cohort.

Perhaps the most compelling evidence is comparing the two different RD estimates from 2009 (2009-1 and 2009-2). Here the year is held constant, but two colleges were exposed only to probation

[31] Data for cohorts entering prior to 2007 is unavailable as the university only began digitally tracking student administrative data in 2007.

while the other four were exposed to both probation and SP. It is unlikely that the probation policy would differ within the same year. As such, it is reasonable to interpret the difference in the two 2009 estimates as the impact of SP.[32]

Moreover, as a placebo test, we estimate Equation (1) but use Q5 GPA (first quarter of the second year) as the running variable where the outcome is year 2 dropout. Recall that scoring below at 2.0 GPA in Q5 still places a student on probation, but importantly there is no SP in year 2. As such, if something else is driving our main result; e.g., the effect of probation changing at the same time as SP implementation or a trend in dropout, then we should expect the DiRD estimates to be economically and statistically significant. We do not find this to be the case. The point estimates from this exercise are reported in Table B7 and are indeed small and insignificant. Overall, the weight of the evidence suggests that the confounding probation policy had the same effect before and after the introduction of SP.

Finally, one might be concerned that trends in the outcomes—e.g., caused by grade inflation, economic downturns, or improvements in student quality—are driving our results. However, we do not find evidence of this. The best way to assess this is to turn to Figure 3. If a trend in year 1 dropout, for example, was driving our DiRD estimate, it wouldn't show up as an abrupt change at the exact time of program adoption as we observe in Figure 3, rather we would expect to see the RD estimates in that figure gradually change across cohorts due to trends, as the above described trends are slow moving.[33] We would also expect our placebo test presented in Table B7 to yield statistically significant results if we were confounding the program effects with trends in the outcomes. We do not find this to be the case.

8 Labor Market Analysis

8.1 Data

To estimate the impact of the coaching program on labor market outcomes, we link the student-level education files to administrative data from the California Employment Development Department. Specifically, we combine two data sources used to administer the state UI program: quarterly earnings records and the Quarterly Census of Employment and Wages (QCEW).[34]

The quarterly earnings records include total earnings in the relevant quarter for each employer–employee (firm) pair. The QCEW data contain earnings and employment data at the

[32] Figure 3c reports RD estimates by cohort for the outcome 6-year graduation. We find similar patterns to the results reported in Figure 3a and Figure 3b, though the estimates are less precise.

[33] Additionally, we probe the robustness of these trends to the population most affected by the policy, low-income students. These results are presented in Figure A7 where we find no strong evidence of pre-trends. We also probe robustness to the use of various bandwidth methods, the addition of controls as well as functional forms. These are summarized in Figure A8, Figure A9, Figure A10 and Figure A11 in the Appendix.

[34] A subset of these data has been used in a series of policy briefs on UI in California during the pandemic (Bell et al., 2022, 2020). Similar and/or related data has also been used in other post-secondary education contexts (e.g., Bleemer and Mehta, 2022; Gurantz, 2019; Hoekstra, 2009; Ost et al., 2018; Zimmerman, 2014).

establishment-quarter level, which we aggregate to the firm level (summing across establishments in California) before linking to the earnings data. Both datasets include the universe of UI-covered employment in the state for the years 2000–2022.[35] As such, we will not observe labor market outcomes for the small share of students who work outside the state of California, are self-employed, or who work for the Federal government.[36] The labor market data are linked to the education files at the student level via social security number.[37] The linked data allow us to construct several labor market outcomes of interest for each student-quarter: log of total earnings, firm pay premium, an indicator for employment, and the cumulative quarters of covered work experience since entering the university. The firm-specific pay premia are estimated as fixed effects from a regression of log earnings on worker and firm identifiers, following Abowd et al. (1999), and are interpretable as the firm-specific component of log earnings received by all workers at the firm.[38] We use the Consumer Price Index for All Urban Consumers to adjust dollar amounts to 2019.

For our main analysis, we limit the data to quarters 25–36 (7–9 years) from initial enrollment (where Q1 is the fall quarter of the student's first year) to ensure that earnings are measured at similar ages for treated and control cohorts. Consequently, we restrict the sample to cohorts 2007 through 2017. Despite this restriction, we cannot observe labor outcomes in all years (i.e., in every year of years 7 to 9) for students who enrolled at the university after 2013 because our data ends in 2022 Q4. Specifically, control cohorts are always observed in each year 7–9 relative to entry, while treatment cohorts entering after 2013 are observed in only some of these years (e.g., the 2014 cohort is observed only in years 6–8.). To alleviate concerns over an unbalanced panel in our main analysis, we present additional estimates in Table B10 using only cohorts that enrolled at the university in 2013 or earlier. This creates a balanced panel across treatment and control groups but reduces the sample size. Overall, results are consistent across balanced and unbalanced samples.

Since the typical student is 18 at enrollment, we observe labor market outcomes at approximately ages 24–26 for this restricted sample. The final dataset used in our labor market analyses includes 204,436 student-quarters that meet these criteria and have GPAs within 1 point of the 2.0 cutoff. Summary statistics for the analysis sample are reported in Table B11. Average quarterly earnings are \$11,660 (approximately \$46,640 annually) and 71% of students are employed.

[35] In Table B9 we exclude Covid-19 pandemic years using 2000–2019 only, and find similar results.

[36] Per Gurantz (2019), the Employment Development Department has estimated that 92% of employed Californians are included in the data.

[37] Of the 45,864 students in our main sample, 43,081 (94%) were employed in at least one quarter following entry to the university.

[38] Specifically, we run the following regression: $y_{it} = \alpha_i + \psi_{\mathbf{J}(i,t)} + \mathbf{X}'_{it}\beta + \epsilon_{it}$, where y_{it} is the log of worker i's earnings in year t, $\mathbf{X}$ contains controls for experience, and $\phi_{\mathbf{J}(i,t)}$ are the firm effects. The estimation sample consists of the largest "connected set" of workers and firms in the CA earnings data between 1997 and 2021. This connected set includes 99% of workers and worker-years in the CA earnings data, 77.4% of firms in the data, and 99.6% of employed student-quarters in the main sample used to estimate the effect of coaching on labor market outcomes. Additional details on sample restrictions, estimation, identification assumptions, and specification checks can be found in Appendix E and Flamang and Kancherla (2024).

8.2 Results

The program we study has the potential to affect student's earnings in their mid-20s in different ways. Suppose, in general, that earnings increase in tandem with level of education and work experience. On the one hand, the program may increase participants' level of education. Indeed, we find it increases the likelihood of obtaining a bachelor's degree for certain groups. On the other hand, the program likely reduces work experience as remaining in school delays entering the labor force, especially relative to the control group who may have dropped out and started working right away. As these two channels are countervailing, the effect of the program on earnings ex ante is unclear.[39]

Our labor market estimates are taken from a variant of Equation (1), but the unit of observation is now the student-quarter level. We cluster standard errors at the student level, and each student-quarter is weighted by the inverse of the number of quarters the student is present in the sample. We report DiRD estimates for various labor market outcomes in Table 5. Row 1 reveals that for the full sample, the coaching program does not appear to produce significant labor market effects. There is suggestive evidence that treated students experience higher earnings—the program leads to about a 12% earnings gain (column 1), but the estimate does not attain statistical significance at conventional levels.[40]

When we focus on the marginal lower-income students (row 2)—the group that experienced the largest academic gains from the program—we find large and statistically significant effects on earnings. Low income students experience about a 30% gain in wages 7-9 years from entry due to the program, while the program appears to have little effect on higher income students' earnings (row 3).[41] Consistent with this finding, the program also leads to low income students being employed at firms with higher pay premia (column 2). Treated low-income students work at firms with a 5.7 percentage point higher firm pay premia compared to their counterparts who just missed the program. While large, the magnitude of these earnings estimates are comparable to estimates of the returns to attending a 4-year university or a relatively higher-quality university (Hoekstra, 2009; Zimmerman, 2014).[42]

There are several caveats to consider when interpreting these results. First, the sample is restricted to earnings in a students' mid-twenties to maintain a nearly balanced panel. Consequently, some students may still be completing their bachelor's degree or pursuing graduate studies at this age. To assess whether the observed earnings benefits persist to an age when wages are more stable, albeit using an unbalanced panel, Figure 4 plots the DiRD estimates for log earnings separately for years 1 to 12 from university entry for low income students. We observe a significant divergence in wages between treated and untreated students emerging in year 7 which corresponds to the year

[39]For visual motivation, Appendix Figure A12 and Figure A13 show earnings effects for treatment and control groups separately.

[40]Appendix Figure A14 shows how labor market DiRD estimates varies by bandwidth choice.

[41]We report DiRD estimates in Table B12 for other subgroups.

[42]We obtain similar results when we exclude Covid years, see Table B9.

that most of the coached students are in the labor market with a bachelor's degree. Importantly, this positive wage effect persists through year 12 from entry, roughly corresponding to age 30, providing further evidence that low-income program participants incur large long-term benefits.

Second, a potential concern, one that would threaten the causal interpretation of the earnings estimates, is if the program induces out-of-state migration. If this is the case, because we do not observe out-of-state earnings, we may systematically code observations as unemployed when they are employed in another state. To address this concern, we follow several recommendations from Foote and Stange (2022). First, our preferred outcome (log earnings) limits the sample to student-quarters with positive earnings. While estimates for this outcome could be affected by differential sample selection, Foote and Stange (2022) note that this is often preferable to incorrectly assuming that the portion of zero-earning observations who have migrated are not working. Additionally, in column 3 of Table 5, we empirically assess if the program changes the likelihood of in-state employment (non-zero earnings). We find no evidence that the program changes the likelihood of being employed 7-9 years from entry for the full sample or for low-income students, where the positive earnings effects of the coaching program are concentrated. This suggests that out-of-state migration bias is of little concern in this setting.

Finally, it is important to recognize that the confidence intervals around the labor market estimates are often large, which is a consequence of a smaller sample. Nevertheless, the weight of the evidence, as presented in Table 5 and Figure 4, suggests the program has large positive effects on the earnings of the impacted students. Moreover, the heterogeneity findings further add to this evidence: as would be expected ex-ante, the earnings benefits are concentrated among those students who experienced the largest academic gains, lower-income students. We view our analysis as an important step in understanding the full benefits of academic coaching, given the dearth of evidence on the effects of such interventions on labor market outcomes.

8.3 Interpreting the Results

To help inform the interpretation of the long-term effects of coaching, we consider the sequential nature of the decisions faced by marginal students. Marginal students who were exposed to and just missed the coaching program (and probation) face a series of decisions summarized by Figure 5. We provide suggestive evidence for which path is more likely for the full sample and for students from lower-income backgrounds, the group that is most responsive to the coaching intervention.

All students begin at Node 0 (Figure 5), their first quarter at the university. We show empirically that students who are exposed to coaching (while also on academic probation) are more likely to progress to Node 1 (i.e. remain enrolled and progress to year 2), while those who just miss the intervention are more likely to end up at Node D, i.e. dropout at the end of year 1 as shown in Table 2 and Table 3.

First, consider the dynamic choices faced by the group that progresses to Node 1. From here, students either graduate from the university and obtain a bachelors degree (Node 2), or dropout at some point in a future period, and obtain "some college". To understand which of these pathways is more likely, we consider marginal students in cohorts that we can observe for at least six years from entry (cohorts 2007-2013). We then construct a series of persistence outcomes by year from entry up to year 6, which take the value of 1 if the student was enrolled in that year or graduated, and 0 otherwise. DiRD estimates for these outcomes are presented in Table 6 and reveal that persistence is quite stable across years in school, and this pattern is particularly strong for students from low-income backgrounds (row 2).[43] These results suggest that students induced to remain at the university due to coaching following year one, especially low-income students, end up successfully graduating and completing their degree at the same university (i.e., successfully reach node 2). In other words, we do not find evidence that coaching is merely delaying dropout. This conclusion is in line with our findings for low-income students who, as a result of the program, are 12.5 percentage points more likely to remain at university after their first year and 13.3 percentage points more likely to graduate within six years.

Next, we focus on dynamic selection for marginal students who drop out at the end of year 1 and thus, end up at Node D. These students face three broad choices: 1) return to the same institution through a readmission path, 2) permanently enter the labor market, or 3) enroll in a lower quality 4-year or 2-year college. In what follows, we provide evidence against the first two counterfactual pathways and argue that the third choice is the most likely counterfactual decision being undertaken by first-year dropouts.

To evaluate the first possibility, that students eventually return to the same university, we obtain DiRD estimates for a re-enrollment outcome.[44] These estimates are presented in Table B14 and show that the coaching treatment is unrelated to re-enrollment for the full sample and for low-income students. We conclude that students induced to dropout are, on average, not re-enrolling at the same institution.

We also show that the second possibility, i.e., permanently entering the labor market after leaving the university at the end of year 1, is not consistent with our data. To asses this possibility, we leverage the fact that we can observe earnings trajectories from university entry through 12 years post entry. Figure 4 summarizes these findings by plotting DiRD estimates by year from entry for students from low-income backgrounds using a bandwidth of 0.75 and 1 GPA points.[45] Each estimate is from a separate DiRD regression where log earnings is the outcome. 95% confidence intervals are represented by the dashed lines and the sample is balanced up until year 9 post entry.[46]

The results highlight that there is no discernible earnings difference between low-income students

[43] The sample for this analysis differs slightly from that in our main table of results, as we restrict to the cohorts for whom we can observe year-by-year enrollment status for up to six years. This restrictions holds constant the sample accross the columns in Table 6.

[44] We define the outcome variable ("re-enroll") to take the value of 1 if a student earns no credits in year 2 and then earns credits in each quarter of year 3, and 0 otherwise.

[45] Figure A15 summarizes findings from this analysis for all students using a bandwidth of 0.75 and 1 GPA points. Although less pronounced compared to the low-income sample, the wage dynamics look similar.

[46] We restrict the sample to cohorts enrolling between 2007 and 2013. We are able to observe earnings up until the 9th year for students enrolled in 2013 which occurs at the end of 2022.

exposed to coaching, compared to those who were not, in the first 6 years post university entry. If anything, we observe a slight dip in earnings for coached students in their second year following university entry. This suggests that the marginal students who were not exposed to coaching, who were more likely to dropout following year 1, experience a transitional period of higher earnings (relative to those who are retained by coaching). However, this difference is temporary as there is a rapid shift back to comparable earnings between the two groups in years 3 through 6 post entry. This pattern provides evidence against the proposed counterfactual channel that students, who just miss coaching and drop out, are permanently entering the labor market in non-degree type jobs. Moreover, Figure 4 reveals a significant divergence in wages between the treated and untreated beginning in year 7 from entry; which corresponds to the first year that most students exposed to coaching would be in the labor market with a bachelors degree (i.e., directly after 6-year graduation). This positive pattern is remarkably stable and persists through 12 years post entry i.e., around 30 years of age.

Taken together, these additional results help link the contemporaneous impacts of coaching and the more downstream outcomes. While we cannot provide conclusive evidence, the interpretation most consistent with our findings is that students who are exposed to coaching are more likely to remain at the university following year 1, and eventually graduate (move directly from node 1 to 2). The most likely counterfactual for these students; i.e., those who would have dropped out and transitioned to note D in Figure 5, is enrolling in a lower quality 4-year or 2-year institution. That said, we caution interpreting the earnings results as solely the returns to graduating from a high-quality university. This is because the exclusion restriction is likely not met as our coaching treatment also changes important student observable and unobservable traits such as effort, motivation and confidence which may directly affect labor market outcomes—a point we emphasize in the next section.[47]

9 Mechanisms

9.1 Linking the Model's Predictions to our Results

Our results indicate that SP improves students' academic performance and retention. Next, we seek to understand the channel underlying these positive findings. While the program is designed as a coaching intervention, it includes several components (i.e., encouragement, information, goal-setting, and time management skills), which all have the potential to individually boost students' academic success. It is difficult to provide conclusive evidence for which of these components drive our effects. Nonetheless, the model we introduced in Section 3 offers some interesting insights.

In the model, we classify the different components of the program into one of two incentive structures: help and encouragement with evidence. Help includes all components of the program

[47] Put differently, instrumenting university graduation with the DiRD threshold would not yield a strict instrumental variable interpretation since other intermediary student outcomes change discontinuously at the GPA cutoff.

that aim to assist students in some way. These comprise giving students information about campus resources and academic probation, coaches walking them through the time management and goal-setting exercises and initiating students' use of campus resources (through requiring them to individually meet with their coach or use a campus resource). On the other hand, components which aim to communicate to students the message that it is still possible to succeed despite failure are classified as encouragement. These include showing a video of successful individuals who previously failed, coaches sharing with students their own experiences with failure, and conducting the workshop in a group setting to show students that they are not the only ones struggling academically.

The model predicts that help reduces students' self-confidence and motivation which leads them to exert less effort to persist at the university. This is because the university will only help students who are low ability. Offering help will thus reveal to the student "bad news" about her probability of success at the university. For students in our setting who have already received negative information about their ability via the probation label, the offer of help will compound any discouragement effects due to probation. An additional assumption of the model is that effort and ability are complements in the production of performance, and so help is predicted to lower both effort and performance. On the other hand, encouragement has the opposite effect. This is because providing encouragement (with evidence) is analogous to revealing to the student "good news" about her probability of success at the university. Encouragement is therefore predicted to boost self-confidence and motivation, which in turn induces students to work harder to persist at the university and translates into improved performance.

The results from our empirical analysis show that the program increases student's academic performance and persistence. These results are consistent with the model's predictions for encouragement, but go in the opposite direction to the predictions of help. This suggests that the program is effective because it provides encouragement with evidence.

9.2 Survey Evidence: Self-Confidence, Motivation, Effort

To further probe the encouragement channel, we explicitly test whether program participation improves self-confidence, motivation and effort—in line with the Bénabou and Tirole (2003) predictions for encouragement. To do so, we utilize data from surveys conducted by the university's Success Program Office. These surveys were administered to SP participants pre- and post-program for eight cohorts of students: those qualifying in the fall quarters of 2013 and 2015–2018 and those qualifying in the winter quarters of 2017–2019.[48] All surveys were administered through the online platform SurveyMonkey and are included in Appendix D. Responses to the survey questions allow

[48]SP was first implemented university-wide in Fall 2010 (though, there was a pilot in Fall 2009 for a subset of faculties) and has been in operation each fall since. For a subset of years, the university also operated a second program in the year for students qualifying in the winter quarter (Winter quarters 2014, 2015 and 2017–2019). While we do not use Winter quarter participation in our main analysis, we do use survey responses from these cohorts of students in the mechanism analysis.

us to evaluate whether the program achieved its direct goals such as improving time management skills and knowledge of campus resources. Importantly, it also allows us to see its impacts on students' self-reported effort, self-confidence and motivation. To implement the analysis, we first use a principal component analysis to create several indices. These indices are standardized weighted-averages of the corresponding survey responses and are used as outcomes in our individual fixed effects model presented below. A complete description of the questions included in each index is provided in Table 7. For example, the effort index includes whether the student attends classes, completes her assigned readings, her number of study hours per week, number of instructor office hours attended per week, whether she edits her written work, studies in advance for exams, prioritizes studying, and reviews incorrect exam answers following the exam. The self-confidence index includes questions about whether the student feels confident in undertaking certain academic tasks. The motivation index comprises questions about the student's motivation to work and earn good grades.

Given the structure of the data, this exercise is more descriptive in nature as the analysis relies on within-student comparisons of outcomes before and after SP participation. We estimate the following model:

$$Y_{it} = \beta_1 + \beta_2 Post_t + \delta_i + \epsilon_{it}, \tag{2}$$

where Y_{it} is the outcome for individual i in period t, $Post_t$ takes the value of 1 to indicate the post-program period, and 0 otherwise; and δ_i is an individual fixed effect. All standard errors are clustered at the individual level.

Estimates from this analysis are reported in Table 8. Appendix Tables B15 and B16 further report estimates for each individual question included in our indices. Column 1 of Table 8 measures students' knowledge and access to helpful resources. The index includes whether students know about campus resources, whether they feel connected to a community within the university and whether they seek help from family and friends. Compared to pre-SP, students report significant improvements in this index. Column 4 further shows that students see an improvement in their ability to manage their time well. This is unsurprising given that the program explicitly targets students' knowledge of campus resources and time management skills. In our model, these are the components that we classify as help and are not predicted to improve performance, as the offer of help reduces self-confidence and effort. Fortunately, our survey allows us to explicitly measure whether the program affects self-confidence, motivation and effort. Indeed, our analysis indicates that students report significant increases in their self-confidence, motivation and study effort and habits (Columns 2-3 and 5-6). These findings are consistent with the predictions of the Bénabou and Tirole (2003) model regarding encouragement, but contradict the theoretical predictions for help. While suggestive, these results provide further evidence that the encouragement component of SP may be driving an important part of our academic performance and persistence effects. Finally, we see no significant effects on outcomes that were not necessarily targeted by the program such as students' reported time spent on leisure and sleep (Column 7 of Table 8).[49]

9.3 Is Encouragement Enough to Boost Performance?

Besides providing encouragement, our coaching program has several other features that may make it particularly effective: it is targeted at academically-struggling students, mandatory and delivered in person. A key question is whether any or a combination of these features is what makes it successful at improving student outcomes. Previous studies indicate that while some programs with these features are effective, not all are. To see this clearly, we compare our coaching program to a previous in-person mandatory intervention targeted at students on academic probation. Weiss et al. (2011) randomly assign probation students to attend a one-semester course and a workshop designed to help students with their study skills, setting goals and learn about their college's rules and regulations. Students were also required to visit a "Success Center", five times during the semester, where they received additional help with study skills, time management and use of campus resources. The intervention improved students' GPA and academic standing during the time it was offered, but effects did not persist after it ended.

The Weiss et al. (2011) intervention is particularly interesting for our purposes since it shares many elements of our program. In addition to being targeted, mandatory and in person, it also includes other components of our program such as helping students with goal-setting, time management, providing them with information on college rules and regulations and requiring them to visit a campus resource. It is arguably more intensive than our program since students attend a course for a full semester (as opposed to a one-time workshop) and are required to visit a campus resource multiple times (versus one time). Yet, the intervention is ineffective at boosting long-term outcomes. We argue that one potential reason for this is that despite it offering all the components of our program that we classify as "help" under the Bénabou and Tirole (2003) model, it fails to provide a key feature of our setting which is "encouragement with evidence".

Is encouragement then enough to boost struggling students' long-term outcomes? Surveying previous low-touch interventions that provide encouragement, it does not seem to be the case. For example, Brady (2017) sends students a psychologically attuned letter instead of the standard academic probation letter. The letter is meant to be encouraging and relays to students that it is still possible to succeed despite probation and that failure is normal. The letter also provides evidence of this as it features the story of a previous academic probation student who ended up overcoming this setback. While this intervention improved the short-term academic performance of a small cohort of students, it had no significant effects on several later cohorts at the same university.

This suggests that our program is particularly effective not just because it focuses on non-cognitive skills but also because it incorporates a "personal" element. Specifically, our program

[49] Looking at individual outcomes in Table B16, there is a decrease in time spent on tv and social media (albeit only significant at the 10% level) in favor of an increase in study hours.

targets non-cognitive skills as it aims to boost self-efficacy. The model and survey evidence suggest that this is mainly achieved through providing encouragement with hard evidence. Importantly however, the program is also delivered in person. Compared to the broader literature on college interventions, this combination of non-cognitive focus and in-person delivery is unique to our setting. Indeed, previous non-cognitive interventions are not typically delivered in person, but they are also ineffective (Brady, 2017; Oreopoulos and Petronijevic, 2018, 2019). On the other hand, in-person interventions focus on helping students with their academic skills but are not typically aimed at improving non-cognitives (Weiss et al., 2011; Barr and Castleman, 2018; Oreopoulos and Petronijevic, 2018; Weiss et al., 2019). Some of these programs are effective at boosting long-term outcomes but they are delivered over a long period of time and are very intensive—as students are offered an array of supports including proactive in-person individualized advising.

Instead, our study highlights an effective and more scalable strategy to improve academically-struggling students' performance. Specifically, our findings indicate that boosting students' self-confidence via providing them with *in-person* and *credible* encouragement can promote both their short- and long-term success.

10 Conclusion

This paper evaluates a coaching program aimed at boosting academically-struggling students' self-efficacy. Students placed on academic probation in their first year at a 4-year US university are required to attend an in-person group workshop. The workshop is designed to improve self-efficacy by exposing students to role models and providing them with encouragement, removing feelings of anxiety and stress, incentivizing the use of campus resources, and helping them with goal-setting and time management.

We find that the coaching program largely improves targeted students' academic outcomes. Overall, program participants experience substantial improvements in their academic performance and retention. Effects are concentrated among lower-income students who also experience an increase in the probability of graduating. For these students, we find that the program increases 6-year graduation rates and earnings 7–9 years following initial enrollment at the university. We posit that the potential reason why our program is effective is because it targets non-cognitive skills (self-efficacy) and is also delivered in person. Based on the predictions of Bénabou and Tirole's (2003) model of intrinsic and extrinsic motivation as well as our own survey evidence, we speculate that this is mainly achieved through providing credible in-person encouragement—and not necessarily through offering help (such as use of campus resources or goal-setting).

Our findings provide evidence for the positive reduced form effects (LATE) of academic coaching on both short- and long-run outcomes. The most likely interpretation of our long-term effects is that the program causes students to shift away from a lower quality 4-year or 2-year college towards graduating from a high quality 4-year university. Estimates from the literature comparing the outcomes of students attending a 4-year as opposed to a lower quality 4-year or 2-year institution, suggest a large and significant 25 to 35 percent earnings premium (Goodman et al., 2020; Zimmerman, 2014; Hoekstra, 2009). Moreover, the magnitude of our earnings estimates are not surprising given that graduates from this institution experience relatively high earnings. According to Grade Reports, which draws on data from the U.S. Department of Education's College Score Card, the institution in our study scores a 90 out of 100 on the "best colleges by earnings one year out of college".[50] However, we caution that the nature of our treatment precludes us from interpreting our results as solely the returns to graduating from a high-quality university. That is, the exclusion restriction is likely not met as our coaching treatment also increases other important student observables and unobservables such as intrinsic motivation and confidence which in itself directly affect labor market outcomes.

From a policy perspective, our findings contribute to the ongoing debate of how to address the college "completion crisis" in the US. Importantly, we show that targeting non-cognitive skills via a short-term, low-touch in-person coaching program can be an effective and inexpensive way to increase academically-struggling students' college retention and long-run success. While the degree to which our findings can be replicated at scale remains an open question, results from this coaching program remain quite promising. Furthermore, our program's lower-cost and less complex structure makes it potentially easy to implement and scale.

[50] The ranking is reported here: https://www.gradreports.com/rankings/california-polytechnic-state-university-san-luis-obispo and is based on median alumni earnings by program in the year after graduating compared to that of other schools.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High wage workers and high wage firms," *Econometrica*, 67, 251–333. (Cited on pages 26 and 100.)

Albert, A. and N. Wozny (2022): "The Impact of Academic Probation: Do Intensive Interventions Help?" *Journal of Human Resources*. (Cited on page 7.)

Angrist, J., D. Autor, and A. Pallais (2020): "Marginal effects of merit aid for low-income students," Tech. rep., National Bureau of Economic Research. (Cited on page 6.)

Angrist, J., D. Lang, and P. Oreopoulos (2009): "Incentives and services for college achievement: Evidence from a randomized trial," *American Economic Journal: Applied Economics*, 1, 136–63. (Cited on page 7.)

Angrist, J., P. Oreopoulos, and T. Williams (2014): "When opportunity knocks, who answers? New evidence on college achievement awards," *Journal of Human Resources*, 49, 572–610. (Cited on page 7.)

Bailey, M. J. and S. M. Dynarski (2011): "Gains and gaps: Changing inequality in US college entry and completion," Tech. rep., National Bureau of Economic Research. (Cited on pages 3 and 5.)

Bandura, A. (1977): "Self-efficacy: toward a unifying theory of behavioral change," *Psychological Review*, 84, 191. (Cited on pages 4 and 10.)

Bandura, A. and D. H. Schunk (1981): "Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation," *Journal of Personality and Social Psychology*, 41, 586. (Cited on pages 4 and 11.)

Barr, A. and B. Castleman (2018): "An engine of economic opportunity: Intensive advising, college success, and social mobility," *Texas A&M Working Paper*. (Cited on pages 7, 8, 22, and 34.)

Barreca, A. I., J. M. Lindo, and G. R. Waddell (2016): "Heaping-induced bias in regression-discontinuity designs," *Economic Inquiry*, 54, 268–293. (Cited on page 23.)

Bell, A., T. Hedin, P. Mannino, R. Moghadam, C. Romer, G. C. Schnorr, and T. von Wachter (2022): "Estimating the disparate cumulative impact of the pandemic in administrative unemployment insurance data," *AEA Papers and Proceedings*, 112, 78–84. (Cited on page 25.)

Bell, A., T. J. Hedin, G. Schnorr, and T. Von Wachter (2020): "An analysis of unemployment insurance claims in California during the COVID-19 Pandemic," *California Policy Lab policy brief*. (Cited on page 25.)

Bénabou, R. and J. Tirole (2000): "Self-confidence and social interactions," . (Cited on page 15.)

——— (2003): "Intrinsic and extrinsic motivation," *The Review of Economic Studies*, 70, 489–520. (Cited on pages 4, 5, 6, 12, 13, 14, 15, 16, 31, 32, and 33.)

——— (2005): "Self-confidence and personal motivation," *Psychology, Rationality and Economic Behaviour: Challenging Standard Assumptions*, 19–57. (Cited on pages 3 and 4.)

Bettinger, E. (2015): "Need-based aid and college persistence: The effects of the Ohio College Opportunity Grant," *Educational Evaluation and Policy Analysis*, 37, 102S–119S. (Cited on page 6.)

Bettinger, E., O. Gurantz, L. Kawano, B. Sacerdote, and M. Stevens (2019): "The long-run impacts of financial aid: Evidence from California's Cal Grant," *American Economic Journal: Economic Policy*, 11, 64–94. (Cited on page 6.)

Bettinger, E. P. and R. B. Baker (2014): "The effects of student coaching: An evaluation of a randomized experiment in student advising," *Educational Evaluation and Policy Analysis*, 36, 3–19. (Cited on pages 7 and 8.)

Black, S. E., J. T. Denning, and J. Rothstein (2022): "Winners and losers? the effect of gaining and losing access to selective colleges on education and labor market outcomes," *American Economic Journal: Applied Economics*. (Cited on page 5.)

Bleemer, Z. (2018): "Top percent policies and the return to postsecondary selectivity," *Available at SSRN*, 3272618. (Cited on page 5.)

——— (2022): "Affirmative action, mismatch, and economic mobility after California's Proposition 209," *The Quarterly Journal of Economics*, 137, 115–160. (Cited on page 5.)

Bleemer, Z. and A. Mehta (2022): "Will studying economics make you rich? a regression discontinuity analysis of the returns to college major," *American Economic Journal: Applied Economics*. (Cited on page 25.)

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. Ter Weel (2008): "The economics and psychology of cognitive and non-cognitive traits," *Journal of Human Resources*, 43, 972–1059. (Cited on page 3.)

Bostwick, V., S. Fischer, and M. Lang (2022): "Semesters or quarters? The effect of the academic calendar on postsecondary student outcomes," *American Economic Journal: Economic Policy*, 14, 40–80. (Cited on page 22.)

Bound, J., M. F. Lovenheim, and S. Turner (2010): "Why have college completion rates declined? An analysis of changing student preparation and collegiate resources," *American Economic Journal: Applied Economics*, 2, 129–57. (Cited on page 8.)

Brady, S. T. (2017): "A Scarlet letter?: Institutional messages about academic probation can, but need not, elicit shame and stigma," Ph.D. thesis, Stanford University. (Cited on pages 7, 33, and 34.)

Bulman, G. B. and C. M. Hoxby (2015): "The returns to the federal tax credits for higher education," *Tax Policy and Economy*, 29, 13–88. (Cited on page 6.)

Cahalan, M. W., M. Addison, N. Brunt, P. R. Patel, and L. W. Perna (2021): "Indicators of higher education equity in the United States: 2021 historical trend report." *Pell Institute for the Study of Opportunity in Higher Education*. (Cited on page 3.)

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014): "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 82, 2295–2326. (Cited on pages 20, 58, and 67.)

Canaan, S., A. Deeb, and P. Mouganie (2022): "Advisor value-dded and student outcomes: evidence from randomly assigned college advisors," *American Economic Journal: Economic Policy*. (Cited on page 7.)

Card, D., J. Heining, and P. Kline (2013): "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly Journal of Economics*, 128, 967–1015. (Cited on page 100.)

Casey, M. D., J. Cline, B. Ost, and J. A. Qureshi (2018): "Academic probation, student performance, and strategic course-taking," *Economic Inquiry*, 56, 1646–1677. (Cited on page 8.)

Clotfelter, C. T., S. W. Hemelt, and H. F. Ladd (2018): "Multifaceted aid for low-income students and college outcomes: Evidence from North Carolina," *Economic Inquiry*, 56, 278–303. (Cited on pages 7 and 22.)

Cohen, G. L. and J. Garcia (2014): "Educational theory, practice, and policy and the wisdom of social psychology," *Policy Insights from the Behavioral and Brain Sciences*, 1, 13–20. (Cited on page 6.)

Cohodes, S. R. and J. S. Goodman (2014): "Merit aid, college quality, and college completion: Massachusetts' Adams scholarship as an in-kind subsidy," *American Economic Journal: Applied Economics*, 6, 251–85. (Cited on page 6.)

Denning, J. T., E. R. Eide, K. Mumford, R. W. Patterson, and M. Warnick (2021): "Why have college completion rates increased? An analysis of rising grades," Tech. rep., National Bureau of Economic Research. (Cited on page 8.)

Dobronyi, C. R., P. Oreopoulos, and U. Petronijevic (2019): "Goal setting, academic reminders, and college success: A large-scale field experiment," *Journal of Research on Educational Effectiveness*, 12, 38–66. (Cited on page 7.)

Dong, Y. (2019): "Regression discontinuity designs with sample selection," *Journal of Business & Economic Statistics*, 37, 171–186. (Cited on page 8.)

Dynarski, S. M. (2003): "Does aid matter? Measuring the effect of student aid on college attendance and completion," *American Economic Review*, 93, 279–288. (Cited on page 6.)

Evans, W. N., M. S. Kearney, B. Perry, and J. X. Sullivan (2020): "Increasing community college completion rates among low-income students: Evidence from a randomized controlled trial evaluation of a case-management intervention," *Journal of Policy Analysis and Management*, 39, 930–965. (Cited on pages 7 and 22.)

Flamang, N. and S. Kancherla (2024): "Temporary shocks, permanent impacts: The effects of liquidity on job search and reemployment," Tech. rep., Working Paper. (Cited on pages 26 and 100.)

Fletcher, J. M. and M. Tokmouline (2018): "The effects of academic probation on college success: Regression discontinuity evidence from four texas universities," . (Cited on page 8.)

Foote, A. and K. M. Stange (2022): "Attrition from administrative data: Problems and solutions with an application to postsecondary education," Tech. rep., National Bureau of Economic Research. (Cited on page 28.)

Goodman, J., M. Hurwitz, and J. Smith (2020): "The economic impact of access to public four-year colleges," *NBER Working Paper No. w27177*, 35. (Cited on pages 3, 5, and 35.)

Grembi, V., T. Nannicini, and U. Troiano (2016): "Do fiscal rules matter?" *American Economic Journal: Applied Economics*, 1–30. (Cited on pages 20 and 21.)

Gurantz, O. (2019): "Impacts of state aid for non-traditional students on educational and labor market outcomes," *Journal of Human Resources*, 0618–9570R2. (Cited on pages 25 and 26.)

Heckman, J. J. and Y. Rubinstein (2001): "The importance of noncognitive skills: Lessons from the GED testing program," *American Economic Review*, 91, 145–149. (Cited on page 3.)

Heckman, J. J., J. Stixrud, and S. Urzua (2006): "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior," *Journal of Labor economics*, 24, 411–482. (Cited on page 3.)

Hoekstra, M. (2009): "The effect of attending the flagship state university on earnings: A discontinuity-based approach," *The Review of Economics and Statistics*, 91, 717–724. (Cited on pages 3, 25, 27, and 35.)

Jackson, C. K. (2021): "Can introducing single-sex education into low-performing schools improve academics, arrests, and teen motherhood?" *Journal of Human Resources*, 56, 1–39. (Cited on page 20.)

Kautz, T., J. J. Heckman, R. Diris, B. Ter Weel, and L. Borghans (2014): "Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success," . (Cited on page 3.)

Kolesár, M. and C. Rothe (2018): "Inference in regression discontinuity designs with a discrete running variable," *American Economic Review*, 108, 2277–2304. (Cited on page 20.)

Kuhnen, C. M. and B. T. Melzer (2018): "Noncognitive abilities and financial delinquency: The role of self-efficacy in avoiding financial distress," *The Journal of Finance*, 73, 2837–2869. (Cited on page 3.)

Lindo, J. M., N. J. Sanders, and P. Oreopoulos (2010): "Ability, gender, and performance standards: Evidence from academic probation," *American Economic Journal: Applied Economics*, 2, 95–117. (Cited on pages 3, 8, and 12.)

Lindqvist, E. and R. Vestman (2011): "The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment," *American Economic Journal: Applied Economics*, 3, 101–128. (Cited on page 3.)

Morisano, D., J. B. Hirsh, J. B. Peterson, R. O. Pihl, and B. M. Shore (2010): "Setting, elaborating, and reflecting on personal goals improves academic performance." *Journal of Applied Psychology*, 95, 255. (Cited on pages 6 and 7.)

Moss, B. G. and W. H. Yeaton (2015): "Failed warnings: Evaluating the impact of academic probation warning letters on student achievement," *Evaluation Review*, 39, 501–524. (Cited on page 7.)

Oreopoulos, P., R. W. Patterson, U. Petronijevic, and N. G. Pope (2022): "Low-touch attempts to improve time management among traditional and online college students," *Journal of Human Resources*, 57, 1–43. (Cited on page 7.)

Oreopoulos, P. and U. Petronijevic (2018): "Student coaching: How far can technology go?" *Journal of Human Resources*, 53, 299–329. (Cited on pages 7, 8, and 34.)

——— (2019): "The remarkable unresponsiveness of college students to nudging and what we can learn from it," Tech. rep., National Bureau of Economic Research. (Cited on pages 7, 8, and 34.)

Oreopoulos, P., U. Petronijevic, C. Logel, and G. Beattie (2020): "Improving non-academic student outcomes using online and text-message coaching," *Journal of Economic Behavior & Organization*, 171, 342–360. (Cited on page 7.)

Ost, B., W. Pan, and D. Webber (2018): "The returns to college persistence for marginal students: Regression discontinuity evidence from university dismissal policies," *Journal of Labor Economics*, 36, 779–805. (Cited on pages 23 and 25.)

Page, L. C., S. S. Kehoe, B. L. Castleman, and G. A. Sahadewo (2019): "More than dollars for scholars the impact of the dell scholars program on college access, persistence, and degree attainment," *Journal of Human Resources*, 54, 683–725. (Cited on page 7.)

Scrivener, S. and M. J. Weiss (2009): "More guidance, better results? Three-year effects of an enhanced student services program at two community colleges," *Three-Year Effects of an Enhanced Student Services Program at Two Community Colleges (August 1, 2009). New York: MDRC*. (Cited on page 7.)

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter (2019): "Firming up inequality," *The Quarterly Journal of Economics*, 134, 1–50. (Cited on page 100.)

Waltenbury, M., S. Brady, M. Gallo, N. Redmond, S. Draper, and T. Fricker (2018): "Academic probation: Evaluating the impact of academic standing notification letters on students," *Higher Education Quality Council of Ontario*. (Cited on page 7.)

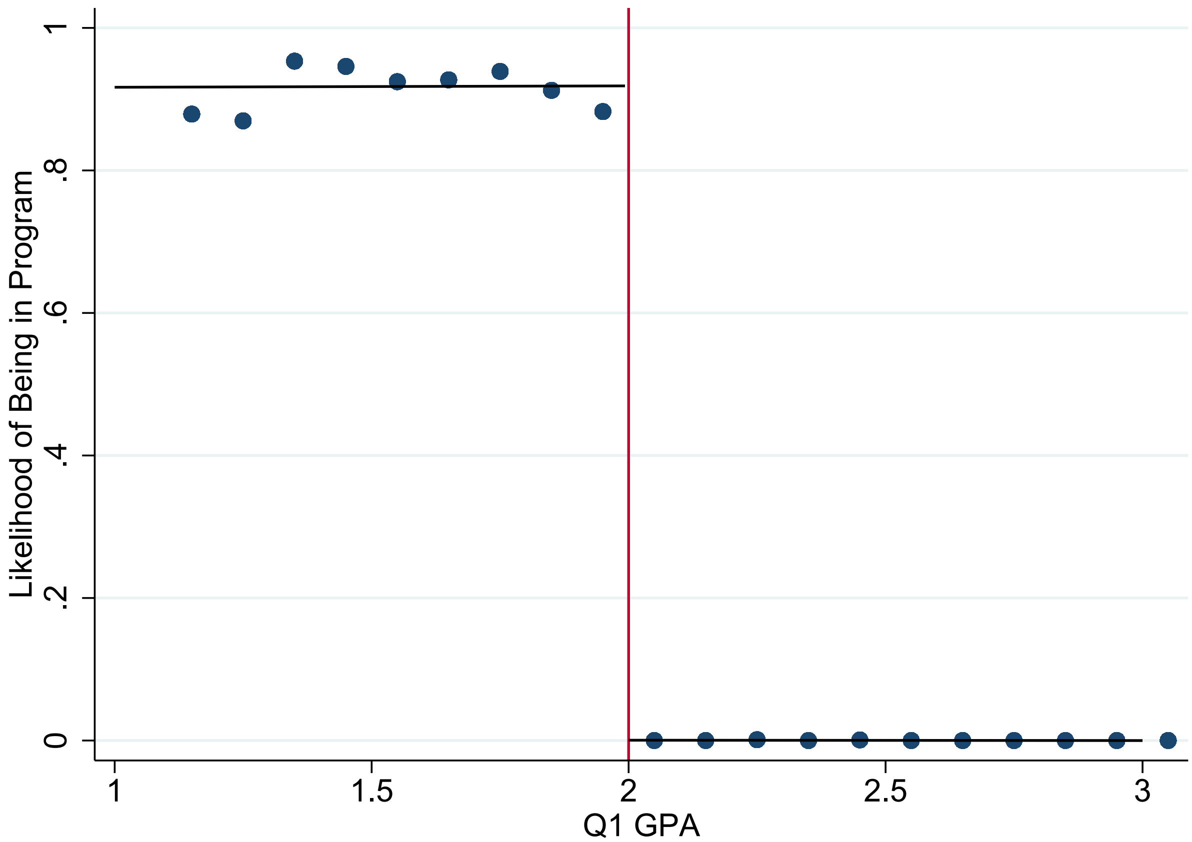
Weiss, M., T. Brock, C. Sommo, T. Rudd, and M. C. Turner (2011): "Serving community college students on probation: Four-year findings from Chaffey College's opening doors program." *MDRC*. (Cited on pages 7, 33, and 34.)

Weiss, M. J., A. Ratledge, C. Sommo, and H. Gupta (2019): "Supporting community college students from start to degree completion: Long-term evidence from a randomized trial of CUNY's ASAP," *American Economic Journal: Applied Economics*, 11, 253–97. (Cited on pages 7, 22, and 34.)

Yeager, D. S. and G. M. Walton (2011): "Social-psychological interventions in education: They're not magic," *Review of Educational Research*, 81, 267–301. (Cited on page 6.)

Zimmerman, S. D. (2014): "The returns to college admission for academically marginal students," *Journal of Labor Economics*, 32, 711–754. (Cited on pages 3, 5, 23, 25, 27, and 35.)

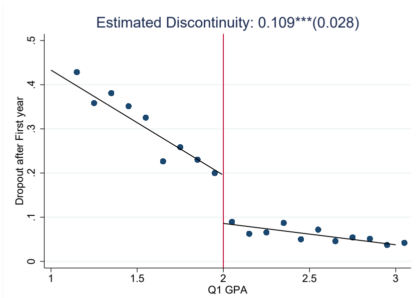
Figure 1: First Stage–Likelihood of Attending Coaching Program



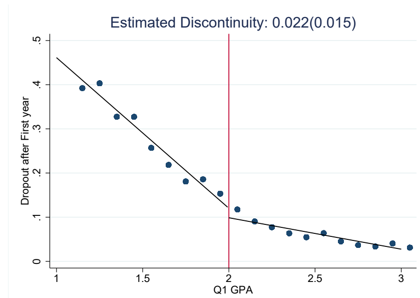
Notes: The sample includes students enrolled at the university after the implementation of the coaching "Success Program" (SP). This includes all first-year students entering the university in the fall cohorts 2010-2017. Circles represent local averages over a 0.1 GPA range. The figure is drawn using a linear fit on either side of the cutoff.

Figure 2: RD Figures for Academic Outcomes by Control and Treated Cohorts

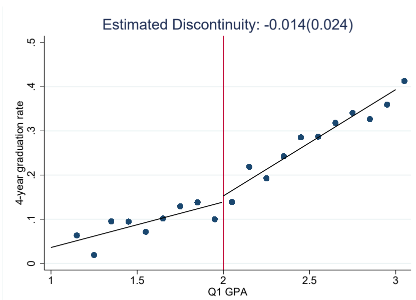
(a) Year 1 Dropout (Control Cohorts)



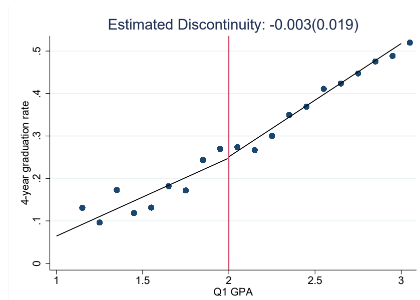
(b) Year 1 Dropout (Treatment Cohorts)



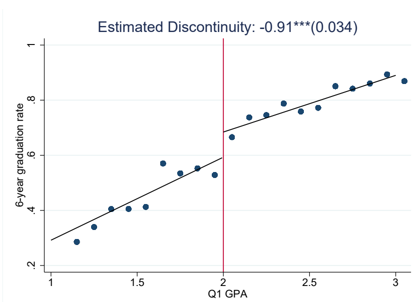
(c) 4-Year Graduation (Control Cohorts)



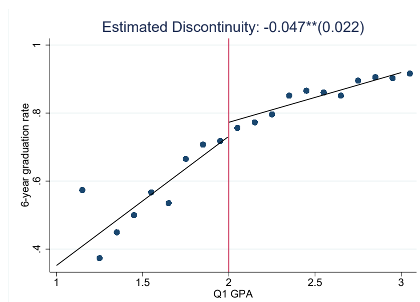
(d) 4-Year Graduation(Treatment Cohorts)



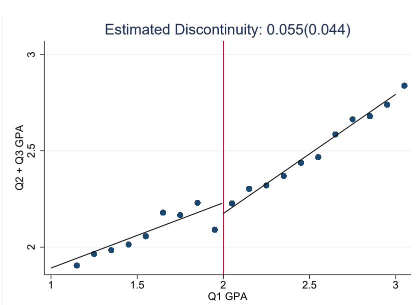
(e) 6-Year Graduation (Control Cohorts)



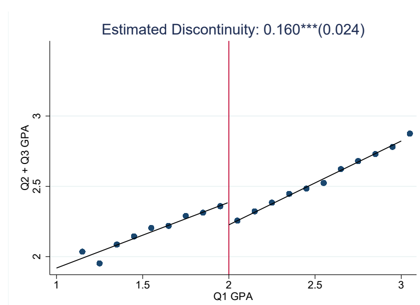
(f) 6-Year Graduation (Treatment Cohorts)



(g) Q2 + Q3 GPA (Control Cohorts)



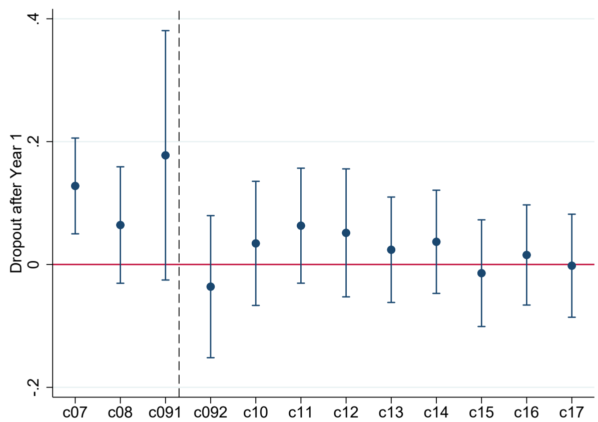
(h) Q2 + Q3 GPA (Treatment Cohorts)



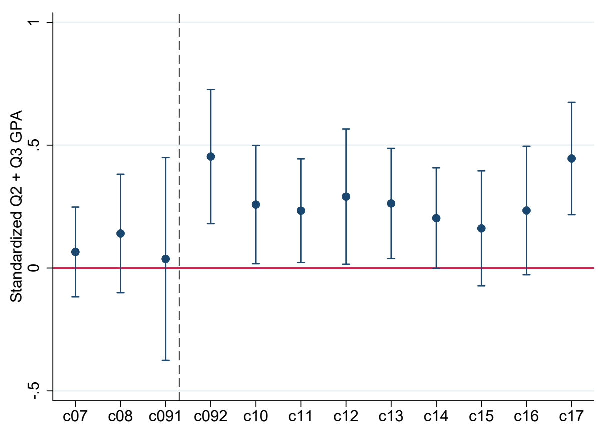
Notes: Control cohort figures include the sample of students enrolled at the university in 2007, 2008 and a subset of colleges in 2009. Treated cohort figures include the sample of students enrolled at the university in select colleges in 2009 and all cohorts from 2010 to 2017 (except for graduation outcomes which include up to the 2014 cohort). Circles represent local averages over a 0.1 GPA range. Figures are drawn using a linear fit on either side of the cutoff. Figure A2 shows these figures are robust to a parametric specification. Estimates and standard errors (in parentheses) are reported above each figure.

Figure 3: RD Estimates by Cohort

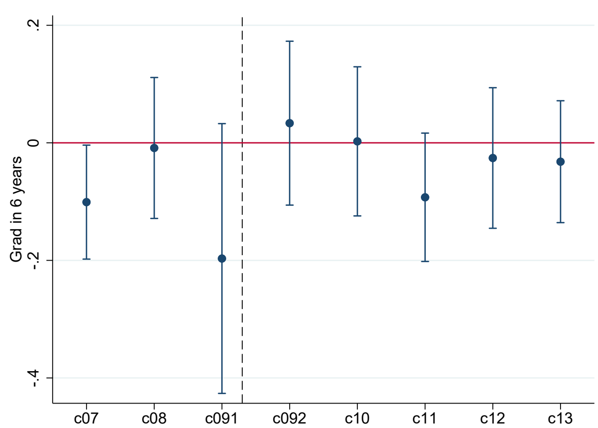
(a) Likelihood of Year 1 Dropout



(b) Standardized Q2 + Q3 GPA



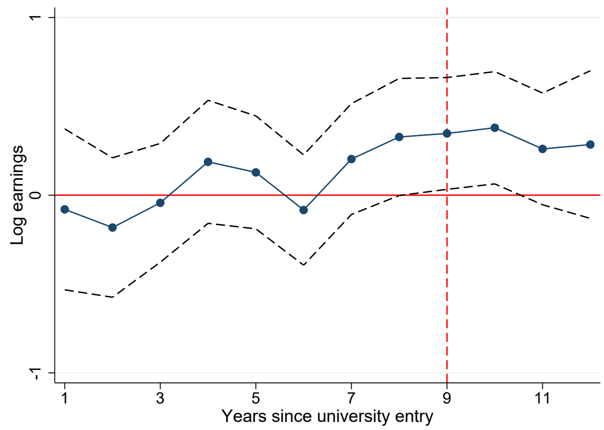
(c) Graduate in 6 Yrs



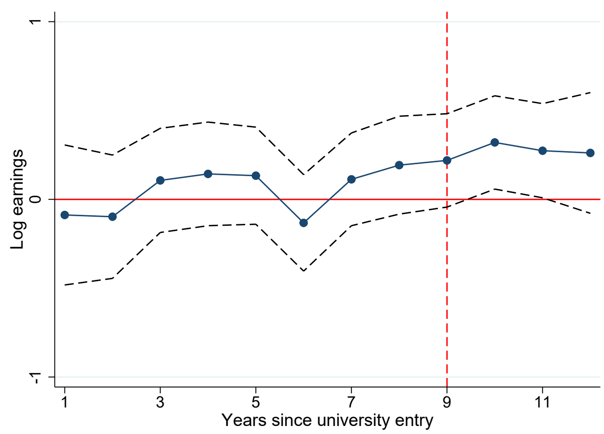
Notes: Figures include all first-year students enrolled at the university in entering fall cohorts 2007-2017. Each point estimate is derived from a separate RD regression, using a bandwidth of 1 GPA point, for each cohort where treatment is defined as scoring below a 2.0 GPA. All point estimates to the left of the dotted line represent control cohorts (first-time freshman students enrolled in 2007, 2008 and two colleges in 2009, i.e. prior to the introduction of the SP). All point estimates to the right of the dotted line represent treated cohorts (first time freshman students enrolled in four colleges in 2009 and the 2010 to 2017 cohorts, i.e. after the introduction of SP). Bars represent upper and lower 95% confidence intervals for each point estimate.

Figure 4: DiRD Wage Estimates Relative to University Enrollment Year for Low-SES Students

(a) Bandwidth=0.75 GPA points



(b) Bandwidth=1 GPA point



Notes: The sample is limited to quarters 7-9 years since the student enrolled at the university and to students who enrolled in or before 2013–where the unit of observation is the student-quarter. The last calendar quarter included in the sample is 2022, Quarter 4. The dashed vertical line reported at 9 years represents the last year in which the sample is balanced. All longer run wage estimates from years 10 through 12 are from unbalanced regressions. Lower SES is defined as students with an expected family contribution (EFC) less than \$30,000. Dollar amounts have been adjusted to 2019 dollars using the Consumer Price Index for All Urban Consumers. Control variables include high school GPA, gender, non-white, Math and English remedial status, indicators for whether parents attended college, and cohort and college fixed effects. All regressions use a bandwidth of 0.75 or 1 grade point on either side of the cutoff. Standard errors are cluster-robust at the student level, and regressions are weighted by one over the number of quarters in which a given student is present in the regression sample. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figure 5: Nodes Summarizing Sequential Nature of University Students' Decisions

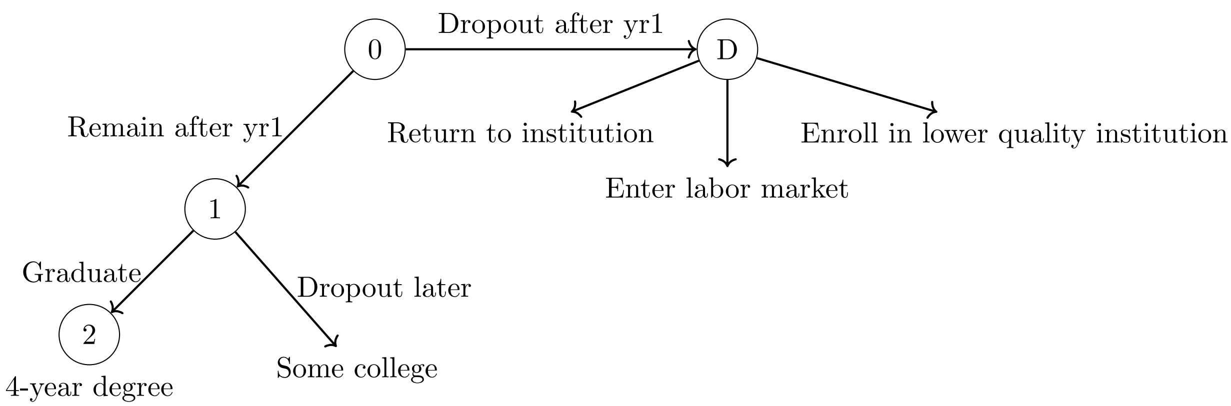


Table 1: Summary Statistics

	(1) Full Sample	(2) Bandwidth=1 Q1 GPA$\in[1,3]$	(3) Pre-program yrs. Q1 GPA $\in[1,2)$ Probation	(4) Pre-program yrs. Q1 GPA $\in[2,3]$ No Probation	(5) Program yrs. Q1 GPA $\in[1,2)$ Probation + SP	(6) Program yrs. Q1 GPA $\in[2,3]$ Neither
Covariates						
HS GPA	3.84	3.73	3.50	3.67	3.61	3.78
	[0.45]	[0.44]	[0.38]	[0.42]	[0.47]	[0.43]
Female	0.48	0.44	0.30	0.46	0.37	0.46
	[0.50]	[0.50]	[0.46]	[0.50]	[0.48]	[0.50]
Non-White	0.38	0.42	0.47	0.35	0.50	0.43
	[0.49]	[0.49]	[0.50]	[0.48]	[0.50]	[0.49]
Remedial Math	0.02	0.03	0.05	0.05	0.04	0.02
	[0.14]	[0.17]	[0.22]	[0.23]	[0.19]	[0.15]
Remedial English	0.04	0.07	0.18	0.12	0.09	0.04
	[0.21]	[0.25]	[0.39]	[0.33]	[0.28]	[0.20]
Pell Grant Eligible	0.16	0.19	0.21	0.13	0.26	0.19
	[0.37]	[0.39]	[0.41]	[0.34]	[0.44]	[0.39]
EFC $<\$30,000$	0.40	0.43	0.43	0.34	0.49	0.44
	[0.49]	[0.49]	[0.50]	[0.47]	[0.50]	[0.50]
Father College +	0.80	0.77	0.70	0.81	0.71	0.77
	[0.40]	[0.42]	[0.46]	[0.39]	[0.45]	[0.42]
Mother College +	0.83	0.80	0.74	0.82	0.74	0.81
	[0.38]	[0.40]	[0.44]	[0.39]	[0.44]	[0.40]
Obs.	45,864	22,225	920	3,708	2,430	15,167
Outcomes						
Dropout Q1	0.01	0.01	0.04	0.01	0.03	0.01
	[0.10]	[0.11]	[0.20]	[0.09]	[0.18]	[0.10]
Dropout Q2	0.02	0.02	0.07	0.02	0.08	0.02
	[0.14]	[0.16]	[0.25]	[0.13]	[0.27]	[0.13]
Dropout Year 1	0.06	0.09	0.30	0.06	0.25	0.05
	[0.24]	[0.28]	[0.46]	[0.23]	[0.43]	[0.23]
Obs.	45,864	22,225	920	3,708	2,430	15,167
4-Yr Grad Rate	0.44	0.35	0.09	0.30	0.18	0.41
	[0.50]	[0.48]	[0.29]	[0.46]	[0.38]	[0.49]
Obs.	36,523	18,281	920	3,708	1,940	11,713
6-Yr Grad Rate	0.86	0.80	0.46	0.81	0.59	0.86
	[0.35]	[0.40]	[0.50]	[0.40]	[0.49]	[0.34]
Obs.	33,440	16,529	920	3,708	1,575	10,326
Q2 + Q3 GPA	2.87	2.53	2.09	2.54	2.21	2.59
	[0.64]	[0.57]	[0.63]	[0.55]	[0.61]	[0.53]
Obs.	44,523	21,421	842	3,619	2,178	14,782
Q2 + Q3 Total Credits	27.77	26.20	21.90	26.43	22.68	26.97
	[5.97]	[6.22]	[7.86]	[5.45]	[8.00]	[5.68]
Obs.	45,864	22,225	920	3,708	2,430	15,167
Treatment						
Probation Ever*	0.38	0.57	1.00	0.53	0.99	0.43
	[0.48]	[0.50]	[0.05]	[0.50]	[0.08]	[0.50]
Probation Yr 1*	0.25	0.41	1.00	0.31	0.99	0.28
	[0.43]	[0.49]	[0.06]	[0.46]	[0.09]	[0.45]
Probation Q1*	0.11	0.17	0.99	–	0.98	–
	[0.31]	[0.38]	[0.09]		[0.12]	
Obs.	45,864	22,225	920	3,708	2,430	15,167
SP Participant Fall**	0.06	0.12	–	–	0.84	–
	[0.24]	[0.32]			[0.37]	
Obs.	37,244	17,597			2,430	

Notes: The sample includes all first-time freshmen enrolled at the university in the entering fall cohorts 2007-2017. Standard deviations are in brackets. *The summary statistics reported for the three probation variables (Probation Ever, Probation Yr 1 and Probation Q1) are based only on the 2007-2009 and 2017 entering fall cohorts. The probation variable is not available for the other years. **The reported means for the SP Participant Fall variable are based only on the program years (entering cohorts 2010-2017 and the subset of colleges that participated in the pilot in 2009) as this variable is undefined for the other years.

Table 2: DiRD Estimates for Academic Outcomes

	(1) Dropout Q1	(2) Dropout Q2	(3) Dropout Q3	(4) Std. GPA (Q2 + Q3)	(5) Earned Credits (Q2 + Q3)	(6) Grad. 4 yr.	(7) Grad. 6 yr.
Panel A: Bandwidth = 0.75							
No Controls	-0.011	0.000	-0.073**	0.168*	1.305*	0.022	0.084*
	(0.016)	(0.021)	(0.037)	(0.089)	(0.682)	(0.036)	(0.048)
With Controls	-0.011	0.000	-0.069*	0.149*	1.266*	0.004	0.073
	(0.016)	(0.021)	(0.037)	(0.085)	(0.679)	(0.034)	(0.046)
Obs.	14,407	14,407	14,407	13,821	14,407	11,895	10,647
Control Mean	0.04	0.05	0.23	-0.84	22.77	0.12	0.55
Panel B: Bandwidth = 1							
No Controls	-0.008	0.001	-0.088***	0.165**	1.241**	0.011	0.050
	(0.014)	(0.018)	(0.032)	(0.076)	(0.589)	(0.031)	(0.041)
With Controls	-0.008	0.001	-0.085***	0.146**	1.189**	-0.012	0.044
	(0.014)	(0.018)	(0.032)	(0.073)	(0.584)	(0.029)	(0.040)
Obs.	22,225	22,225	22,225	21,421	22,225	18,281	16,529
Control Mean	0.04	0.05	0.23	-0.84	22.77	0.12	0.55

Notes: The sample includes all first-time freshmen enrolled at the university in the entering fall cohorts 2007-2017. Control variables include high school GPA, gender, non-white, Math and English remedial status, Pell Grant eligibility status, whether EFC is less than $30,000, indicators for whether parents attended college, and cohort and college fixed effects. Each point estimate is from a separate regression. DiRD estimates are equivalent to differencing two local linear RD regressions. Figure A5 reports the DiRD estimate for each outcome for a range of bandwidths. The reported control mean is the average outcome of the pre-period cohorts with Q1 GPA $\in [1.75, 2.0)$. Robust standard errors reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: DiRD Estimates for Academic Outcomes by SES

	(1) Dropout Q1	(2) Dropout Q2	(3) Dropout Q3	(4) Std. GPA (Q2 + Q3)	(5) Earned Credits (Q2 + Q3)	(6) Grad. 4 yr.	(7) Grad. 6 yr.
Panel A: Bandwidth = 0.75							
Low SES	-0.012	-0.006	-0.091	0.267**	1.694	-0.076	0.169**
	(0.025)	(0.031)	(0.058)	(0.131)	(1.037)	(0.050)	(0.071)
High SES	-0.011	-0.002	-0.048	0.050	0.942	0.069	0.004
	(0.021)	(0.028)	(0.048)	(0.112)	(0.909)	(0.046)	(0.061)
Obs. (Low SES)	6,255	6,255	6,255	6,014	6,255	5,121	4,526
Control Mean	0.04	0.05	0.26	-0.95	22.31	0.13	0.42
Obs. (High SES)	8,152	8,152	8,152	7,807	8,152	6,774	6,121
Control Mean	0.04	0.06	0.21	-0.76	23.09	0.12	0.64
Panel B: Bandwidth = 1							
Low SES	-0.012	-0.015	-0.125**	0.200*	1.912**	-0.064	0.149**
	(0.020)	(0.026)	(0.050)	(0.113)	(0.879)	(0.043)	(0.060)
High SES	-0.006	0.013	-0.057	0.111	0.586	0.032	-0.024
	(0.019)	(0.025)	(0.042)	(0.096)	(0.791)	(0.040)	(0.052)
Obs. (Low SES)	9,491	9,491	9,491	9,158	9,491	7,772	6,954
Control Mean	0.04	0.05	0.26	-0.95	22.31	0.13	0.42
Obs. (High SES)	12,734	12,734	12,734	12,263	12,734	10,509	9,575
Control Mean	0.04	0.06	0.21	-0.76	23.09	0.12	0.64

Notes: The sample includes all first-time freshmen enrolled at the university in the entering fall cohorts 2007-2017. Control variables include high school GPA, gender, non-white, Math and English remedial status, Pell Grant eligibility status, whether EFC is less than $30,000, indicators for whether parents attended college, and cohort and college fixed effects. Each point estimate is from a separate regression. DiRD estimates are equivalent to differencing two local linear RD regressions. Table B8 shows that these results are robust to dropping the 2009 cohort, the year of the pilot program. The reported control mean is the average outcome of the pre-period cohorts with Q1 GPA $\in [1.75, 2.0)$. Robust standard errors reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Baseline Covariates Balance Check for DiRD Research Design

	(1) HS GPA	(2) Female	(3) Non-White	(4) Remedial Math	(5) Remedial English	(6) Pell elig.	(7) EFC < $30K	(8) Father college	(9) Mother college
Bandwidth= 0.5	-0.070	-0.010	0.047	-0.006	0.022	-0.002	0.001	0.017	0.028
	(0.045)	(0.051)	(0.056)	(0.027)	(0.041)	(0.046)	(0.057)	(0.052)	(0.049)
Bandwidth= 0.75	-0.030	-0.022	0.029	-0.029	-0.032	-0.044	-0.055	0.031	0.019
	(0.037)	(0.043)	(0.047)	(0.022)	(0.034)	(0.038)	(0.047)	(0.043)	(0.041)
Bandwidth= 1	0.005	0.010	0.050	-0.020	-0.008	-0.015	-0.020	0.013	-0.028
	(0.032)	(0.037)	(0.040)	(0.018)	(0.029)	(0.033)	(0.040)	(0.037)	(0.035)
Obs. (BW=0.5)	8,973	8,973	8,973	8,973	8,973	8,973	8,973	8,973	8,973
Obs. (BW=0.75)	14,407	14,407	14,407	14,407	14,407	14,407	14,407	14,407	14,407
Obs. (BW=1)	22,225	22,225	22,225	22,225	22,225	22,225	22,225	22,225	22,225

Notes: The sample includes all first-time freshmen enrolled at the university in entering fall cohorts 2007-2017. Each point estimate is from a separate regression. The estimation equation is presented in Equation (1). DiRD estimates are equivalent to differencing two local linear RD regressions. Robust standard errors reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: DiRD Estimates for Labor Market Outcomes

	(1)	(2)	(3)	(4)
	Log Earnings	Firm Pay Premia	Employed	Cumulative Experience (Qtrs)
All Students	0.121	0.0481	–0.0166	–0.929
	(0.0904)	(0.0315)	(0.0315)	(0.662)
Low SES	0.310**	0.0569	0.0133	0.0772
	(0.144)	(0.0475)	(0.0478)	(1.03)
High SES	–0.0168	0.0420	–0.0345	–1.58*
	(0.115)	(0.0421)	(0.0420)	(0.859)
Obs. (All)	144,297	143,614	204,436	204,436
Ctrl Mean	13,030	0.574	0.729	16.7
Obs. (Low SES)	61,715	61,431	86,868	86,868
Ctrl Mean	12,213	0.522	0.686	16.3
Obs. (High SES)	82,582	82,183	117,568	117,568
Ctrl Mean	13,553	0.608	0.760	17.1

Notes: The sample is limited to quarters 7-9 years since the student enrolled at the university where the unit of observation is the student-quarter. The last calendar quarter included in the sample is 2022, Quarter 4. Dollar amounts have been adjusted to 2019 dollars using the Consumer Price Index for All Urban Consumers. Control means are calculated among students in the pre-coaching group, with GPAs just below the probation cutoff (Q1 GPA $\in [1.75,2)$). Control means for log earnings outcomes are in dollars. All regressions use a bandwidth of 1 grade points on either side of the cutoff. Standard errors are cluster-robust at the student level, and regressions are weighted by one over the number of quarters in which a given student is present in the regression sample. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: DiRD Estimates: Persistence by Year

	(1) Complete Yr 2 (or Graduate)	(2) Complete Yr 3 (or Graduate)	(3) Complete Yr 4 (or Graduate)	(4) Complete Yr 5 (or Graduate)	(5) Complete Yr 6 (or Graduate)
All Students	0.051	0.047	0.033	0.036	0.048
	(0.039)	(0.041)	(0.041)	(0.042)	(0.042)
Low-SES	0.104*	0.138**	0.153**	0.146**	0.163**
	(0.061)	(0.063)	(0.063)	(0.063)	(0.063)
High-SES	0.016	-0.012	-0.050	-0.040	-0.031
	(0.051)	(0.053)	(0.054)	(0.055)	(0.055)
Obs. (All)	13,927	13,927	13,927	13,927	13,927
Control Mean	0.69	0.62	0.59	0.56	0.56
Obs. (Low SES)	5,841	5,841	5,841	5,841	5,841
Control Mean	0.62	0.50	0.47	0.44	0.43
Obs. (High SES)	8,086	8,086	8,086	8,086	8,086
Control Mean	0.74	0.70	0.68	0.65	0.65

Notes: The sample includes all first-time freshmen enrolled at the university in the entering fall cohorts 2007-2013. Control variables include high school GPA, gender, non-white, Math and English remedial status, Pell Grant eligibility status, whether EFC is less than $\$30,000$, indicators for whether parents attended college, and cohort and college fixed effects. Each point estimate is from a separate regression. DiRD estimates are equivalent to differencing two local linear RD regressions. The reported control mean is the average outcome of the pre-period cohorts with Q1 GPA $\in [1.75, 2.0)$. Robust standard errors reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Indices for Testing Mechanisms

Panel A: 2013 Surveys

1. Effort	2. Self Confidence	3. Motivation
Attend class	Feel confident in writing ability	Motivated to focus on school
Post-exam, review material that got wrong	Feel confident taking exams	Motivated to earn good grade in classes that are uninteresting
Number of study hours		Study even when less important things are distracting
Number office hours attended		
Complete assigned reading		
Edit written work		
Study in advance for exams		
Prioritize studying		
4. Resources/Community	5. Leisure	6. Study Habits
Ask family/friends for help	Av. hours sleep per night	Retain info from assignments
Know about student services	Av. hours socializing/extracurriculars per day	Class notes help to prepare for exams
Connected to a community at university	Av. hours of TV/gaming/social media per day	Easily/effectivly communicate thoughts
		Ask questions to clarify when don't understand instructor
		Easily remember things learned in class
		At end of class, can summarize lecture material
		Adjust learning style to instructor teaching style
7. Time Management		
Manages time well		

Panel B: 2015-18 Surveys

1. Resources/Community	2. Self Confidence	3. Motivation
Not only one on academic probation	Confident in time management skills	Motivated to focus on academics
Can identify faculty/staff who cares	Confident in decision to attend the university	
Know of campus resources to help get back on track	Confident will get off probation by Q2	
Feel connected to the university	Confident will graduate from the university	

Notes: The outcomes in Panel A come from pre- and post-SP surveys for students who qualified for SP in Fall 2013. The outcomes in Panel B come from pre- and post-program surveys for students who qualified for SP in Fall 2015, 2016, 2017 and 2018 as well as Winter 2017, 2018 and 2019. Note that the survey questions changed following the 2013 cohort so it is not possible to aggregate survey responses across all years.

Table 8: Mechanisms: Individual FE Estimates with Indexed (PCA) Outcomes

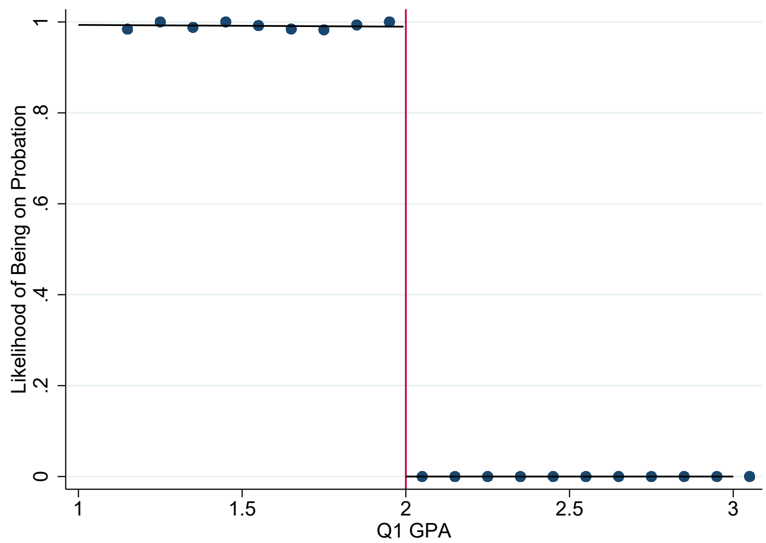
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2013 Surveys							
	Resources	Self-Confidence	Motivation	Time Management	Study Effort	Study Habits	Leisure
Post	1.012***	0.331***	0.332***	0.620***	0.796***	0.603***	-0.022
	(0.069)	(0.082)	(0.072)	(0.080)	(0.068)	(0.079)	(0.078)
Obs.	430	430	430	430	430	430	430
Panel B: 2015-18 Surveys							
	Resources	Self-Confidence	Motivation				
Post	0.694***	0.110***	-0.475***				
	(0.040)	(0.039)	(0.042)				
Obs.	1,842	1,842	1,842				

Notes: The data come from pre- and post-SP program survey responses. An observation is a student-survey, where there are two survey waves (pre and post SP). All regressions include a post program indicator "Post" and individual fixed effects. The sample for Panel A is a balanced panel of 430 observations for those participating in SP in 2013. The response rate is about 91%, as 316 students from this cohort completed the SP and 288 responded to at least one of the surveys. The sample for Panel B is a balanced panel of 1,842 observations from those participating in SP in 2015-2018. Outcomes are a standardized index; mean 0, standard deviation 1. The survey administered in 2103 includes a different set of questions from that administered in the following years (2015-2018), as such, the responses cannot be aggregated across these two periods. Table B15 reports estimates for each outcome in the PCA Index. Standard errors are reported in parentheses and are clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1

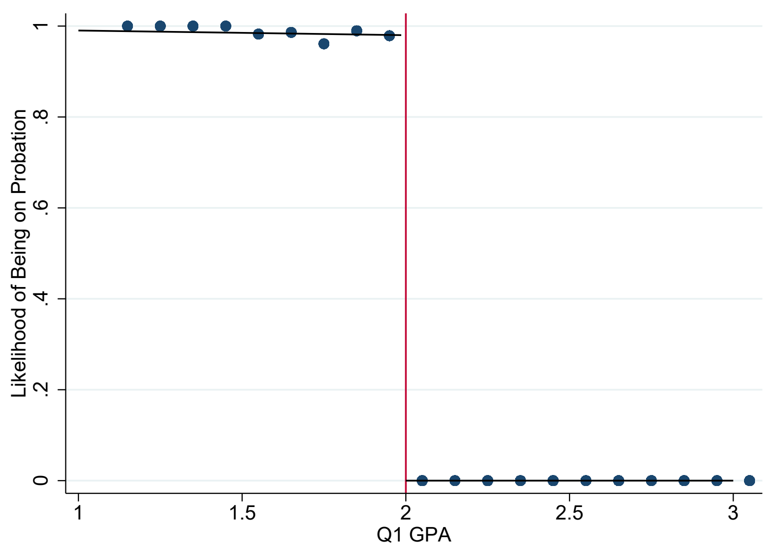
A Appendix Figures

Figure A1: Likelihood of Probation Following Q1

(a) Pre-Program (Control) Cohorts



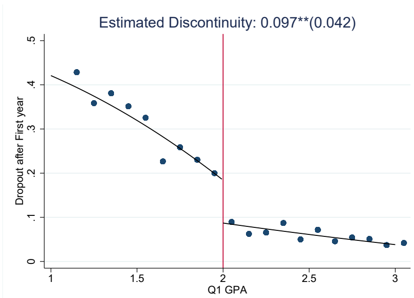
(b) Program (Treated) Cohorts



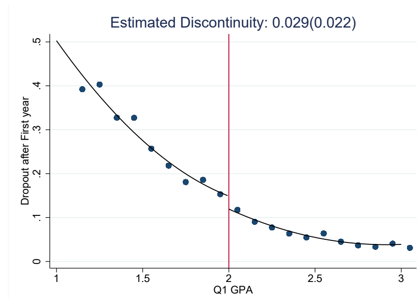
Notes: The sample includes all first-year students enrolled at the university in entering fall cohorts 2007, 2008, 2009 and 2017. 2010-2016 cohorts are excluded because the probation variable is missing. The running variable is first quarter GPA in both figures. Circles represent local averages over a 0.1 GPA range. Figures are drawn using a linear fit on either side of the cutoff.

Figure A2: RD Figures for Academic Outcomes by Control and Treated Cohorts—Parametric Estimation

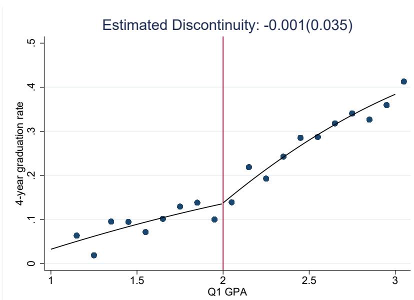
(a) Year 1 Dropout (Control Cohorts)



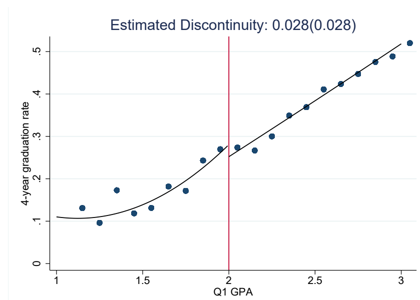
(b) Year 1 Dropout (Treatment Cohorts)



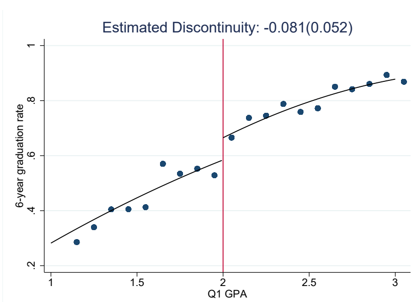
(c) 4-Year Graduation (Control Cohorts)



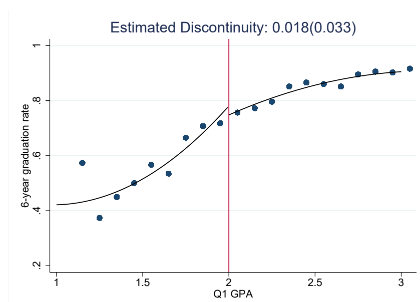
(d) 4-Year Graduation(Treatment Cohorts)



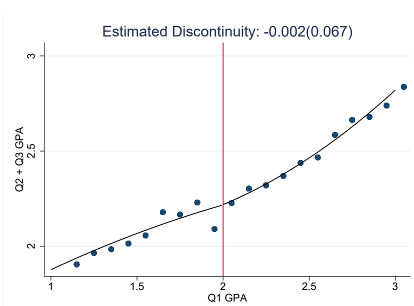
(e) 6-Year Graduation (Control Cohorts)



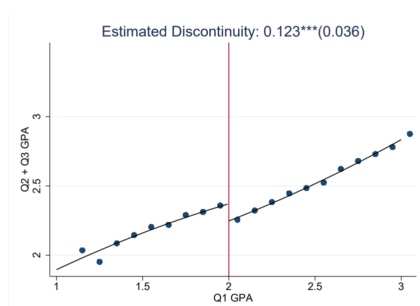
(f) 6-Year Graduation (Treatment Cohorts)



(g) Q2 + Q3 GPA (Control Cohorts)



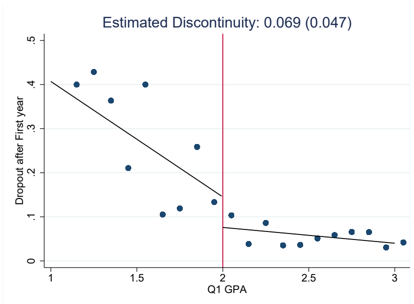
(h) Q2 + Q3 GPA (Treatment Cohorts)



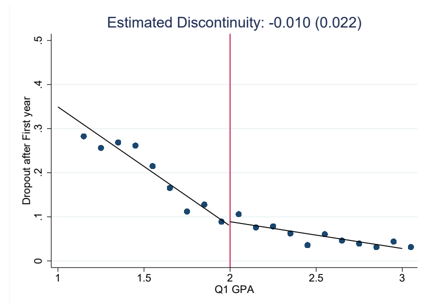
Notes: Control cohort figures include the sample of students enrolled at the university in 2007, 2008 and a subset of colleges in 2009. Treated cohort figures include the sample of students enrolled at the university in select colleges in 2009 and all cohorts from 2010 to 2017 (except for graduation outcomes which include up to the 2014 cohort). Circles represent local averages over a 0.1 GPA range. Figures are drawn using a quadratic fit on either side of the cutoff. Estimates and standard errors (in parentheses) are reported above each figure.

Figure A3: RD Figures for Yr 1 Dropout by Control and Treated Cohorts for Subgroups

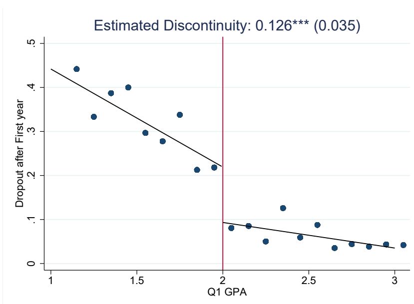
(a) Females (Control Cohorts)



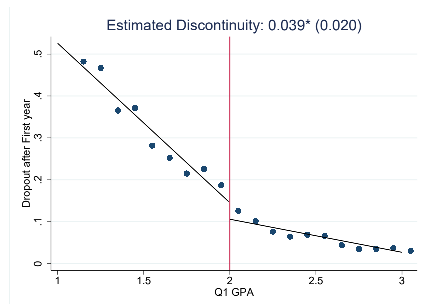
(b) Females (Treatment Cohorts)



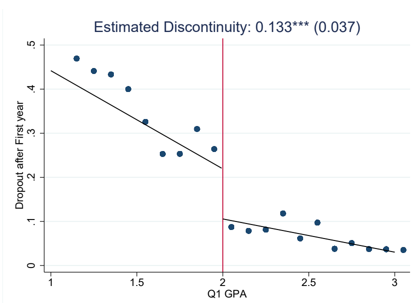
(c) Males (Control Cohorts)



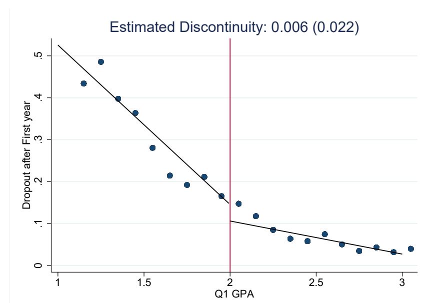
(d) Males (Treatment Cohorts)



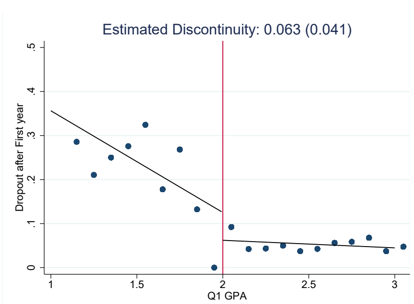
(e) STEM (Control Cohorts)



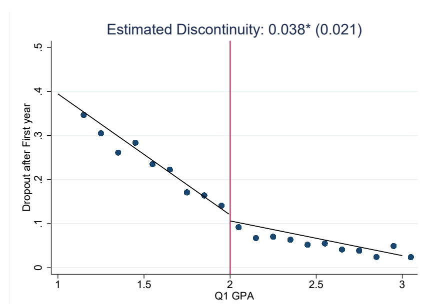
(f) STEM (Treatment Cohorts)



(g) Non-STEM (Control Cohorts)



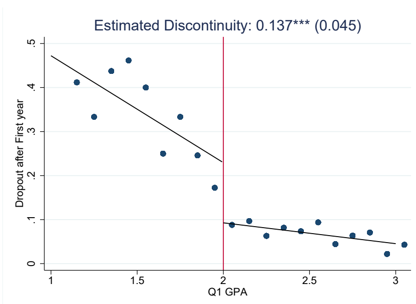
(h) Non-STEM (Treatment Cohorts)



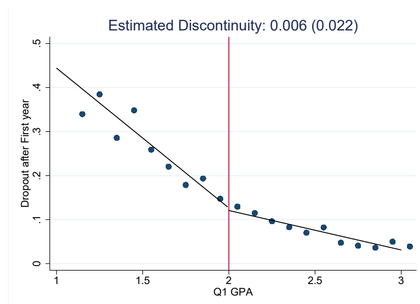
Notes: Control cohort figures include the sample of students enrolled at the university in 2007, 2008 and a subset of colleges in 2009. Treated cohort figures include the sample of students enrolled at the university in select colleges in 2009 and all cohorts from 2010 to 2017 (except for graduation outcomes which include up to the 2014 cohort). The STEM subsample includes students in the college of engineering, architecture and sciences. The non-STEM subsample includes students who are in the college of agriculture, business and liberal arts. Circles represent local averages over a 0.1 GPA range. Figures are drawn using a linear fit on either side of the cutoff. Estimates and standard errors (in parentheses) are reported above each figure.

Figure A4: RD Figures for Yr 1 Dropout by Control and Treated Cohorts for Subgroups

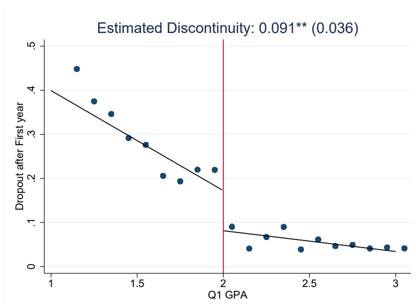
(a) Low SES (Control Cohorts)



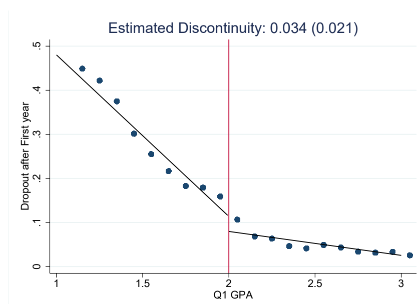
(b) Low SES (Treatment Cohorts)



(c) High SES (Control Cohorts)

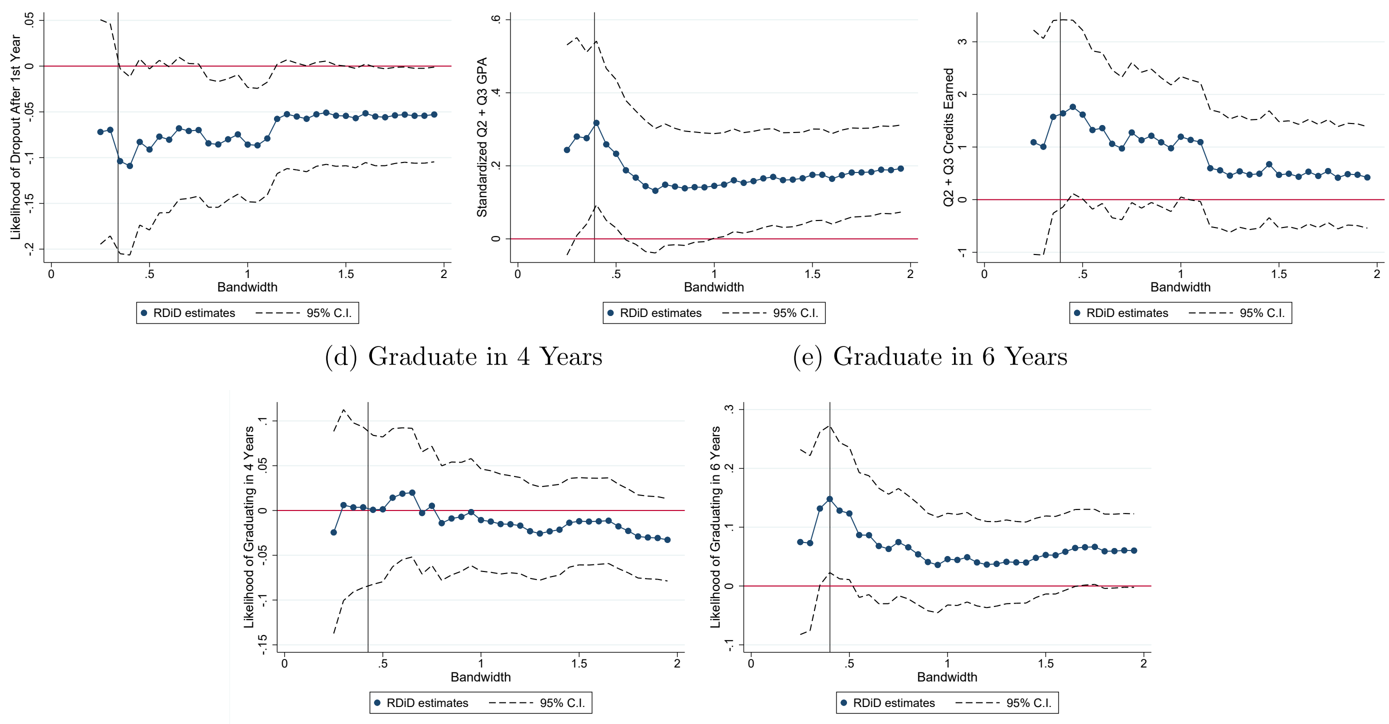


(d) High SES (Treatment Cohorts)



Notes: Control cohort figures include the sample of students enrolled at the university in 2007, 2008 and a subset of colleges in 2009. Treated cohort figures include the sample of students enrolled at the university in select colleges in 2009 and all cohorts from 2010 to 2017 (except for graduation outcomes which include up to the 2014 cohort). Lower SES is defined as students with an expected family contribution (EFC) less than \$30,000. Higher SES are those with an EFC greater than or equal to \$30,000, and includes those who have a missing EFC. Circles represent local averages over a 0.1 GPA range. Figures are drawn using a linear fit on either side of the cutoff. Estimates and standard errors (in parentheses) are reported above each figure.

Figure A5: DiRD Estimates for Academic Outcomes by Bandwidth



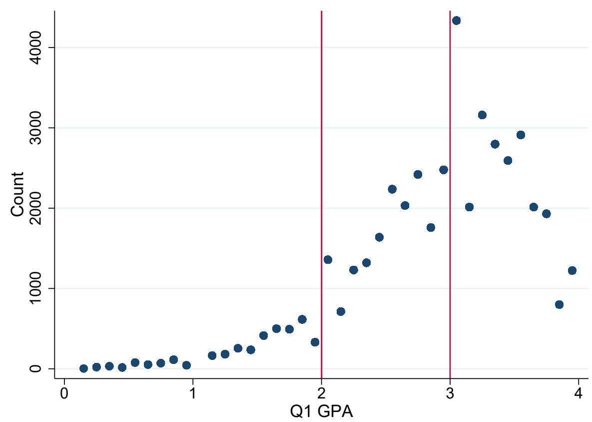
(a) Year 1 Dropout (b) Standardized Q2 + Q3 GPA (c) Q2 + Q3 Earned Credits

(d) Graduate in 4 Years (e) Graduate in 6 Years

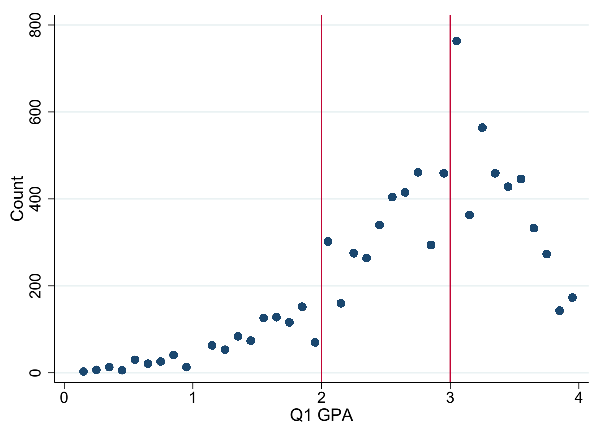
Notes: The sample includes all first-time freshmen enrolled at the university in the entering fall cohorts 2007-2017. Control variables include high school GPA, gender, non-white, Math and English remedial status, Pell Grant eligibility status, whether EFC is less than $\$30,000$, indicators for whether parents attended college, and cohort and college fixed effects. Dashed lines represent 95% confidence intervals. DiRD estimates are equivalent to differencing two local linear RD regressions. The vertical black lines indicate the optimal bandwidth calculated using the CCT procedure as described in Calonico et al. (2014).

Figure A6: Bunching at Whole GPA Cutoffs

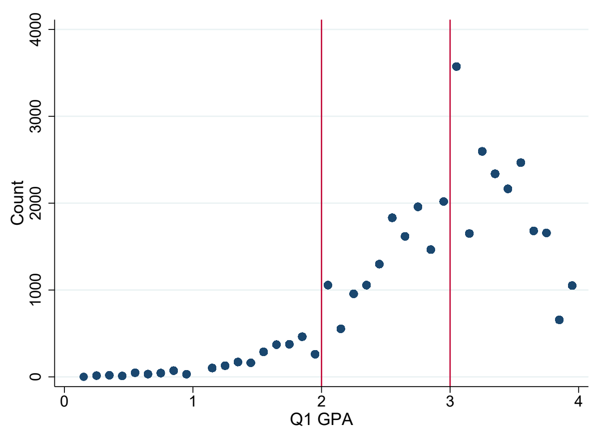
(a) Distribution of Q1 GPA (All Cohorts)



(b) Distribution of Q1 GPA (Control Cohorts)



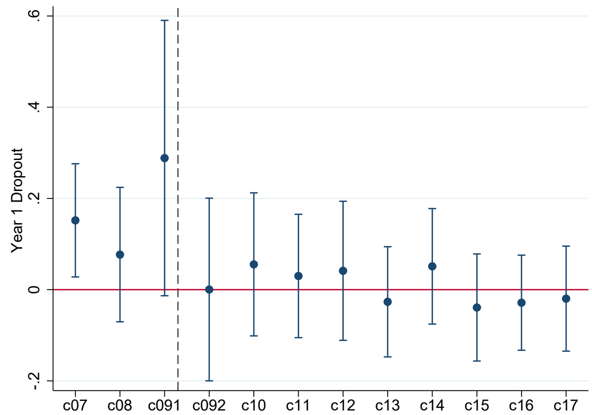
(c) Distribution of Q1 GPA (Treated Cohorts)



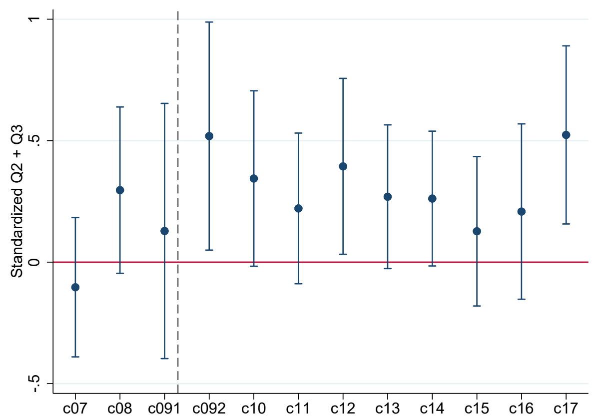
Notes: The sample used in Panel A includes all first-year students entering the university in fall cohorts 2007-2017. The sample used in Panel B includes cohorts never exposed to SP (2007, 2008 and part of 2009 cohort). The sample in Panel C includes all cohorts exposed to SP (the four colleges of the 2009 cohort and 2010-2017 cohorts).

Figure A7: RD Estimates by Cohort for Low-SES

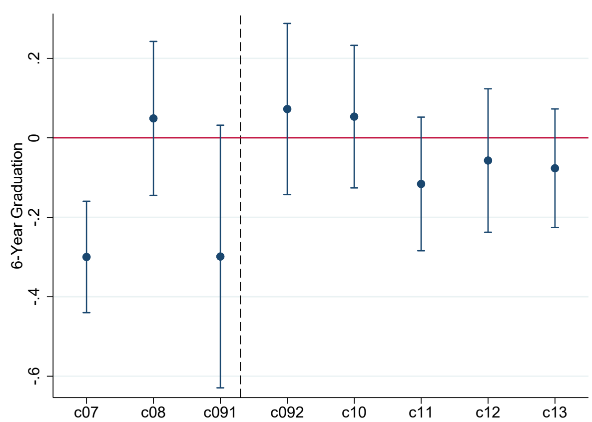
(a) Year 1 Dropout



(b) Standardized Q2+Q3 GPA



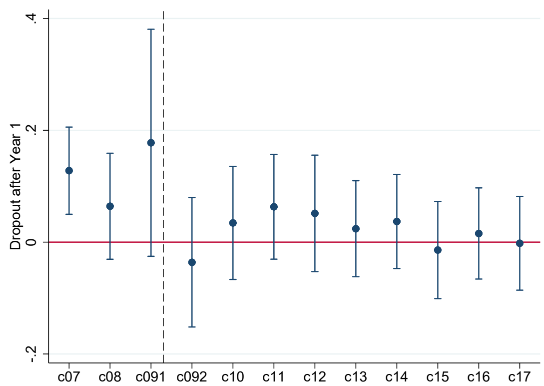
(c) 6-Year Graduation



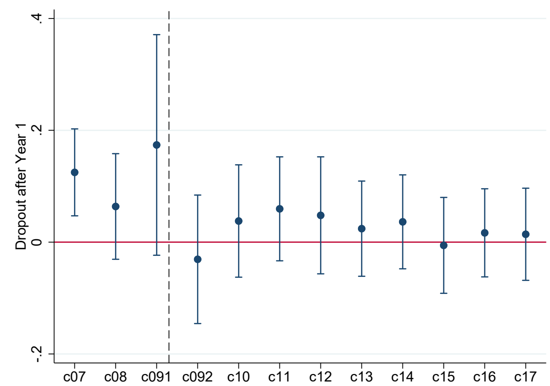
Notes: Figures include first-year students who are low-SES in entering fall cohorts 2007-2017. Each point estimate is derived from a separate RD regression for each cohort where treatment is defined as scoring below a 2.0 GPA. All point estimates to the left of the dotted line represent control cohorts (first-time freshman students enrolled in 2007, 2008 and two colleges in 2009, i.e. prior to the introduction of the SP). All point estimates to the right of the dotted line represent treated cohorts (first time freshman students enrolled in four colleges in 2009 and the 2010 to 2017 cohorts, i.e. after the introduction of SP). Bars represent upper and lower 95% confidence intervals for each point estimate.

Figure A8: Year 1 Dropout, RD Estimates by Cohort Alternative Specifications

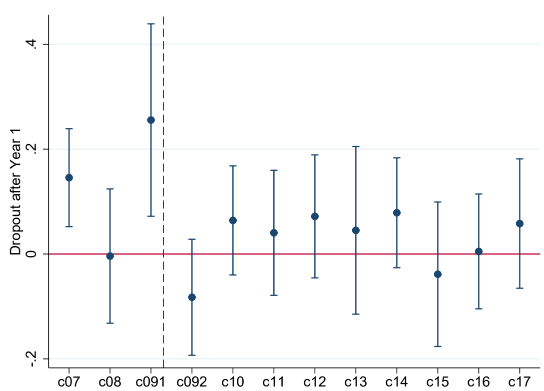
(a) Local Linear, Bandwidth=1



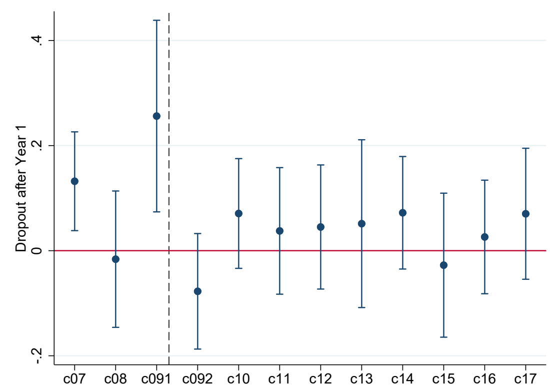
(b) Local Linear, Bandwidth=1, + Controls)



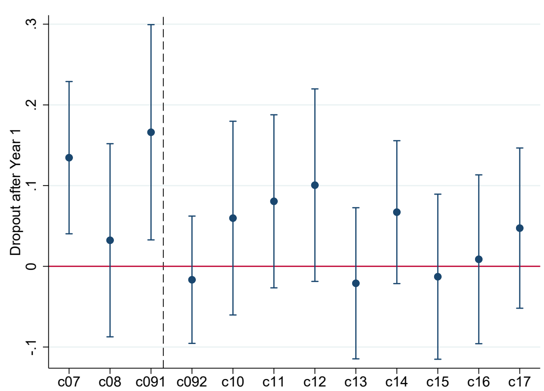
(c) Local Linear, CCT Bandwidth Selector



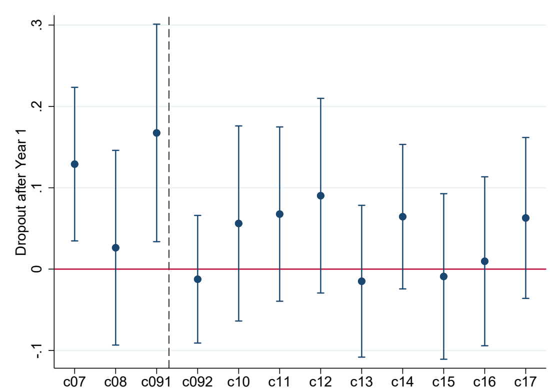
(d) Local Linear, CCT Bandwidth Selector, + Controls



(e) Quadratic, Bandwidth=1



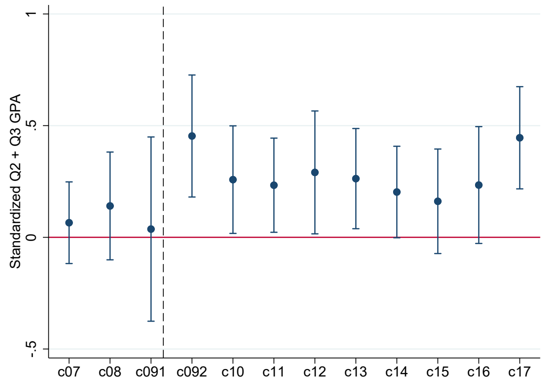
(f) Quadratic, Bandwidth=1, + Controls



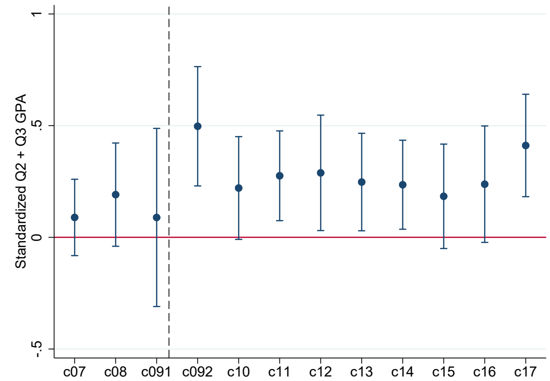
Notes: Figures include all first-year students enrolled at the university in entering fall cohorts 2007-2017. Each point estimate is derived from a separate RD regression for each cohort where treatment is defined as scoring below a 2.0 GPA. All point estimates to the left of the dotted line represent control cohorts (first-time freshman students enrolled in 2007, 2008 and two colleges in 2009, i.e. prior to the introduction of the SP). All point estimates to the right of the dotted line represent treated cohorts (first time freshman students enrolled in four colleges in 2009 and the 2010 to 2017 cohorts, i.e. after the introduction of SP). Bars represent upper and lower 95% confidence intervals for each point estimate.

Figure A9: Standardized Q2 + Q3 GPA, RD Estimates by Cohort Alternative Specifications

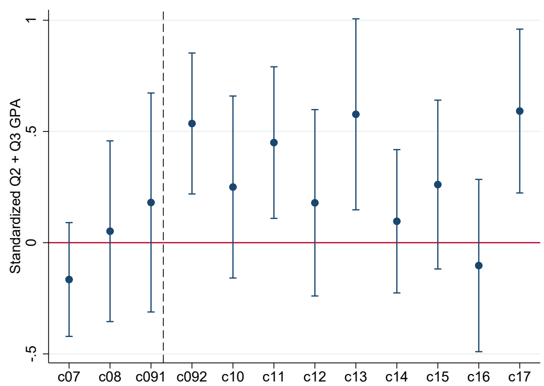
(a) Local Linear, Bandwidth=1



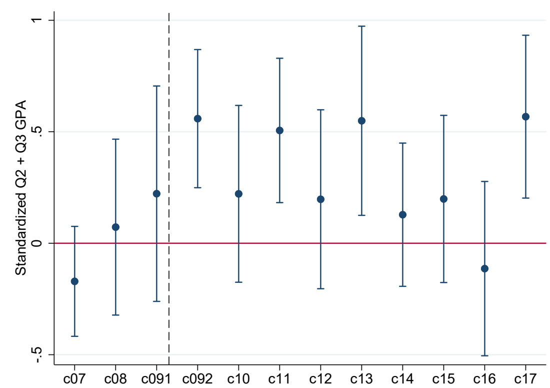
(b) Local Linear, Bandwidth=1, + Controls)



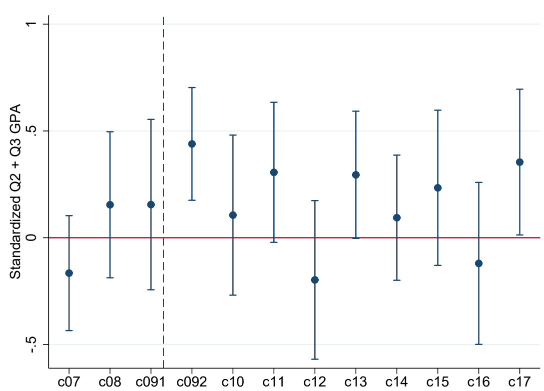
(c) Local Linear, CCT Bandwidth Selector



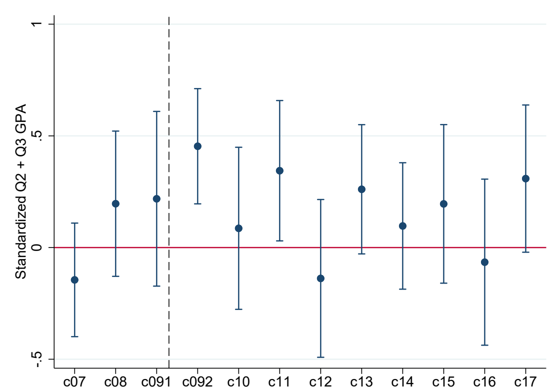
(d) Local Linear, CCT Bandwidth Selector, + Controls



(e) Quadratic, Bandwidth=1



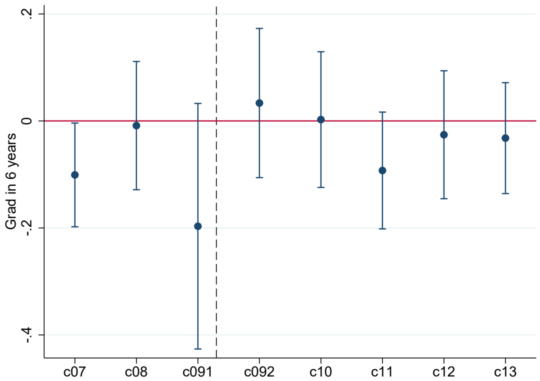
(f) Quadratic, Bandwidth=1, + Controls



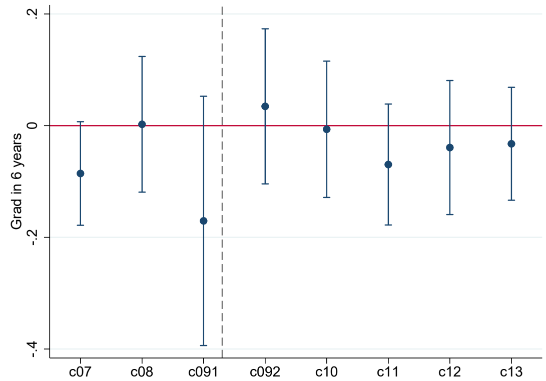
Notes: Figures include all first-year students enrolled at the university in entering fall cohorts 2007-2017. Each point estimate is derived from a separate RD regression for each cohort where treatment is defined as scoring below a 2.0 GPA. All point estimates to the left of the dotted line represent control cohorts (first-time freshman students enrolled in 2007, 2008 and two colleges in 2009, i.e. prior to the introduction of the SP). All point estimates to the right of the dotted line represent treated cohorts (first time freshman students enrolled in four colleges in 2009 and the 2010 to 2017 cohorts, i.e. after the introduction of SP). Bars represent upper and lower 95% confidence intervals for each point estimate.

Figure A10: 6 Year Graduation, RD Estimates by Cohort Alternative Specifications

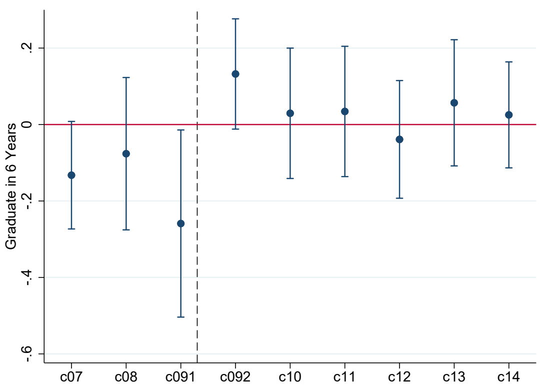
(a) Local Linear, Bandwidth=1



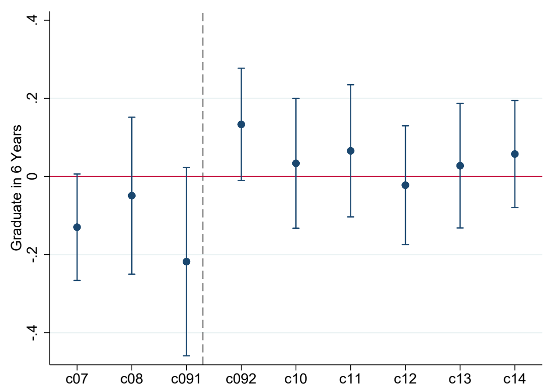
(b) Local Linear, Bandwidth=1, + Controls)



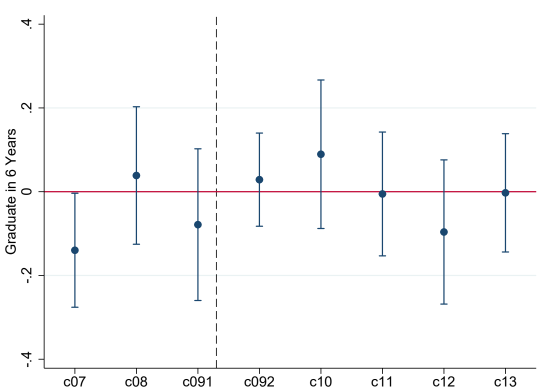
(c) Local Linear, CCT Bandwidth Selector



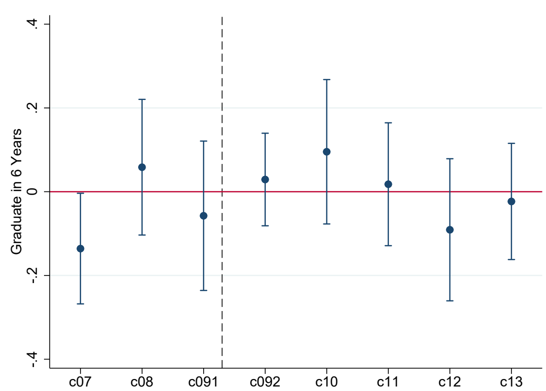
(d) Local Linear, CCT Bandwidth Selector, + Controls



(e) Quadratic, Bandwidth=1



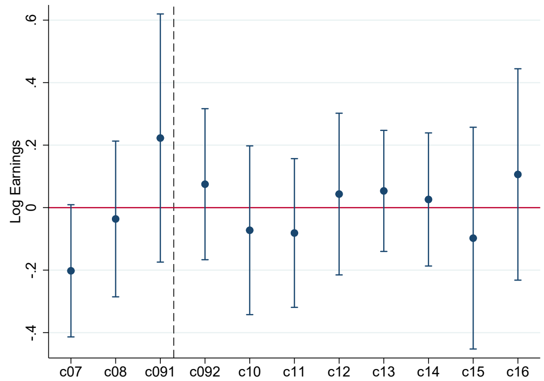
(f) Quadratic, Bandwidth=1, + Controls



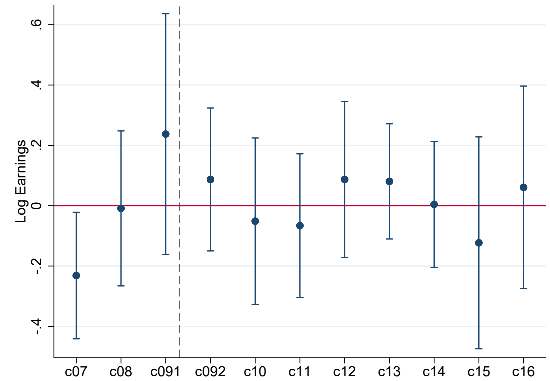
Notes: Figures include all first-year students enrolled at the university in entering fall cohorts 2007-2017. Each point estimate is derived from a separate RD regression for each cohort where treatment is defined as scoring below a 2.0 GPA. All point estimates to the left of the dotted line represent control cohorts (first-time freshman students enrolled in 2007, 2008 and two colleges in 2009, i.e. prior to the introduction of the SP). All point estimates to the right of the dotted line represent treated cohorts (first time freshman students enrolled in four colleges in 2009 and the 2010 to 2017 cohorts, i.e. after the introduction of SP). Bars represent upper and lower 95% confidence intervals for each point estimate.

Figure A11: Log Earnings, RD Estimates by Cohort Alternative Specifications

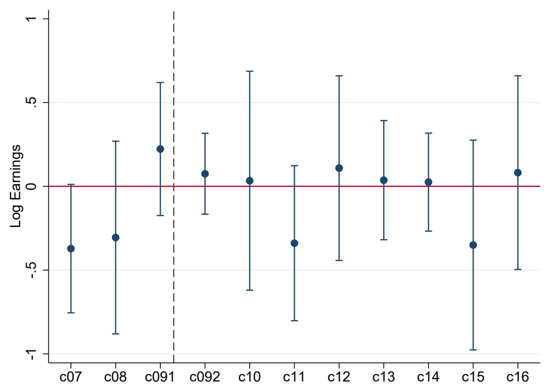
(a) Local Linear, Bandwidth=1



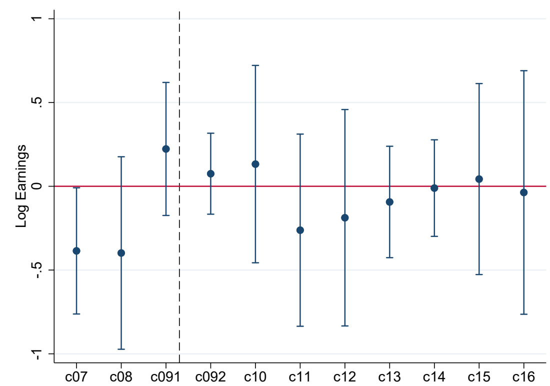
(b) Local Linear, Bandwidth=1, + Controls)



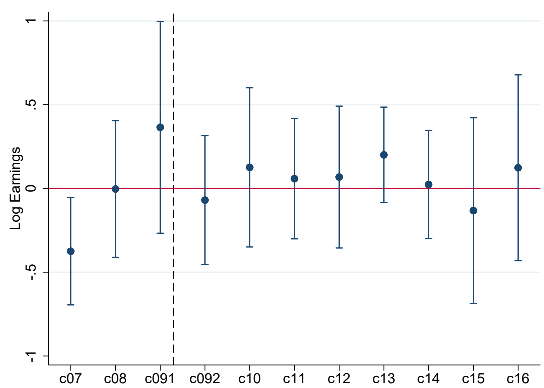
(c) Local Linear, CCT Bandwidth Selector



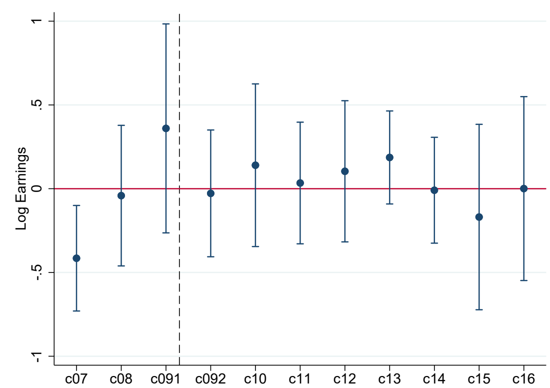
(d) Local Linear, CCT Bandwidth Selector, + Controls



(e) Quadratic, Bandwidth=1



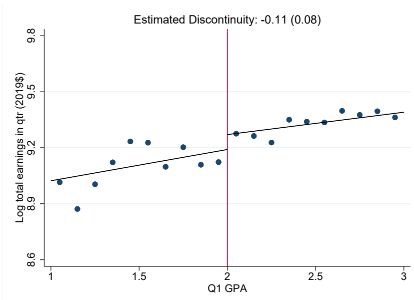
(f) Quadratic, Bandwidth=1, + Controls



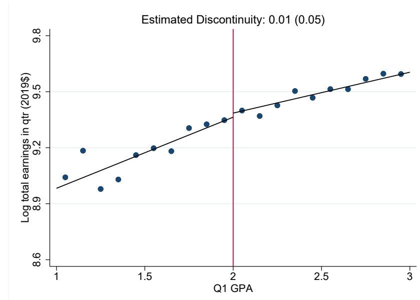
Notes: The sample is limited to quarters 7-9 years since the student enrolled at the university where the unit of observation is the student-quarter. The last calendar quarter included in the sample is 2022, Quarter 4. Dollar amounts have been adjusted to 2019 dollars using the Consumer Price Index for All Urban Consumers. Control variables include high school GPA, gender, non-white, Math and English remedial status, Pell Grant eligibility status, whether EFC is less than $30,000, indicators for whether parents attended college, and cohort and college fixed effects. All regressions use a bandwidth of 0.75 or 1 grade point on either side of the cutoff. Standard errors are cluster-robust at the student level, and regressions are weighted by one over the number of quarters in which a given student is present in the regression sample. *** p<0.01, ** p<0.05, * p<0.1.

Figure A12: RD Figures for Log Earnings by Control and Treated Cohorts for Subgroups

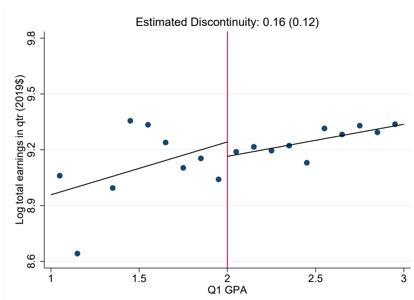
(a) All (Control Cohorts)



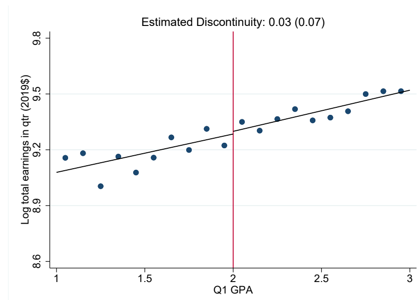
(b) All (Treatment Cohorts)



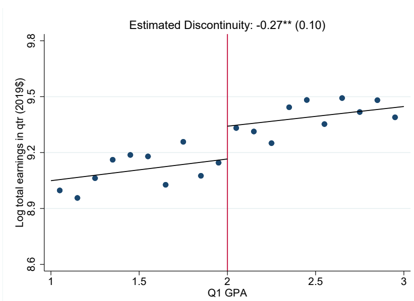
(c) Females (Control Cohorts)



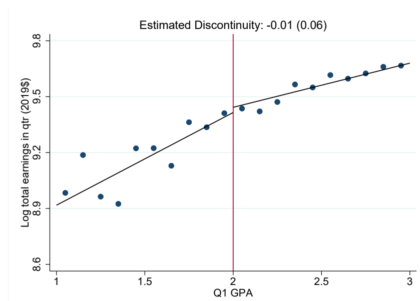
(d) Females (Treatment Cohorts)



(e) Males (Control Cohorts)



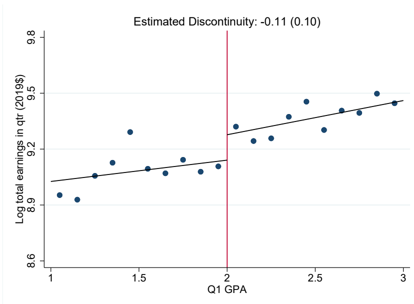
(f) Males (Treatment Cohorts)



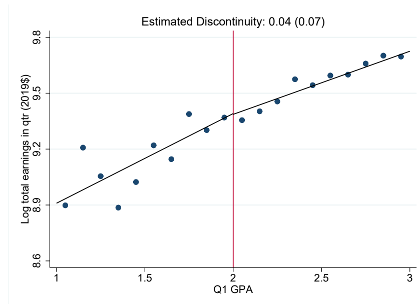
Notes: Control cohort figures include the sample of students enrolled at the university in 2007, 2008 and a subset of colleges in 2009. Treated cohort figures include the sample of students enrolled at the university in select colleges in 2009 and all cohorts from 2010 to 2017 (except for graduation outcomes which include up to the 2014 cohort). The STEM subsample includes students in the college of engineering, architecture and sciences. The non-STEM subsample includes students who are in the college of agriculture, business and liberal arts. Circles represent local averages over a 0.1 GPA range. Figures are drawn using a linear fit on either side of the cutoff. Estimates and standard errors (in parentheses) are reported above each figure.

Figure A13: RD Figures for Log Earnings by Control and Treated Cohorts for Subgroups

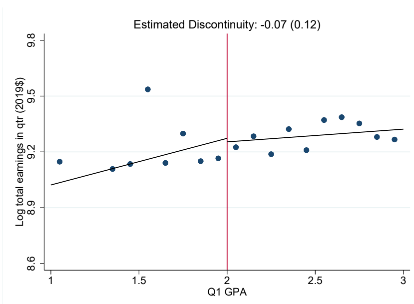
(a) STEM (Control Cohorts)



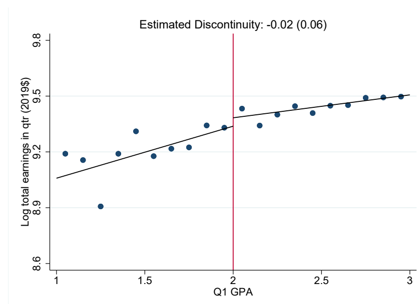
(b) STEM (Treatment Cohorts)



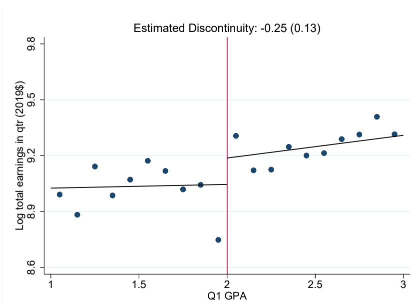
(c) Non-STEM (Control Cohorts)



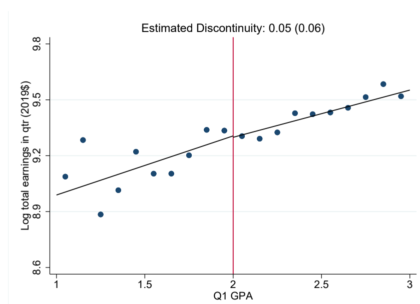
(d) Non-STEM (Treatment Cohorts)



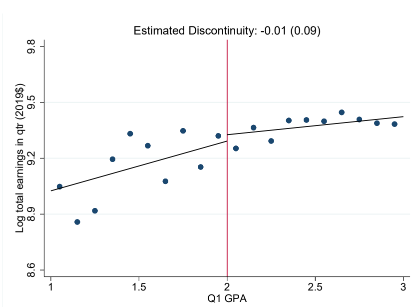
(e) Lower SES (Control Cohorts)



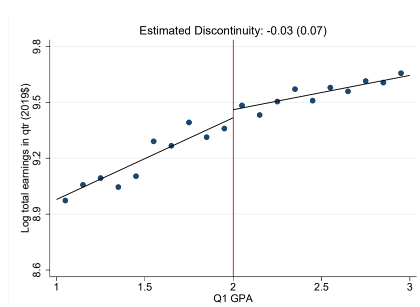
(f) Lower SES (Treatment Cohorts)



(g) Higher SES (Control Cohorts)



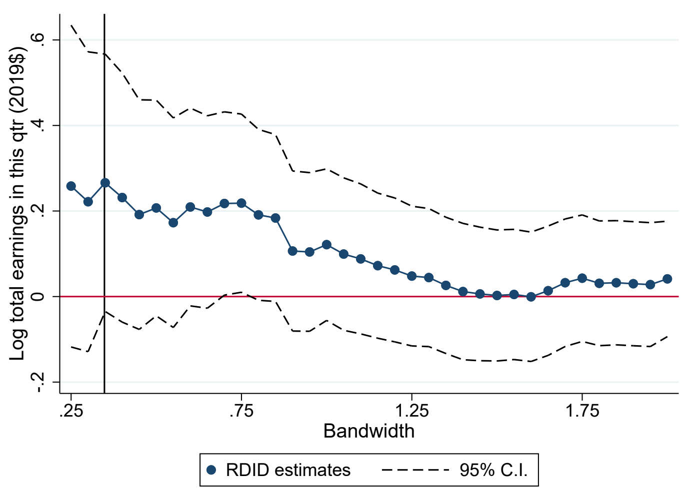
(h) Higher SES (Treatment Cohorts)



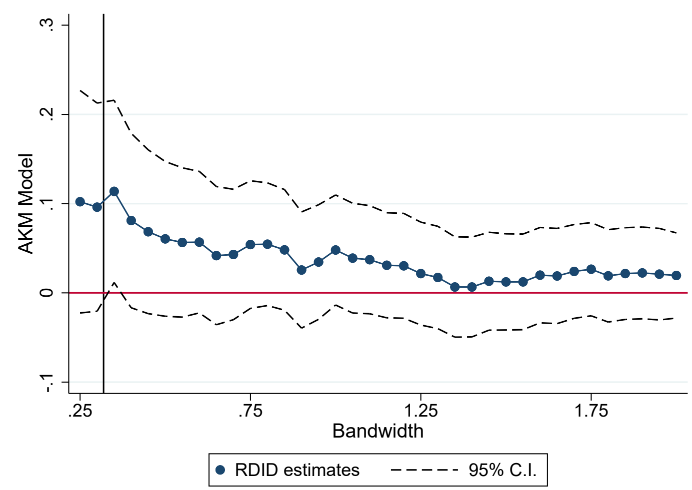
Notes: Control cohort figures include the sample of students enrolled at the university in 2007, 2008 and a subset of colleges in 2009. Treated cohort figures include the sample of students enrolled at the university in select colleges in 2009 and all cohorts from 2010 to 2017 (except for graduation outcomes which include up to the 2014 cohort). Lower SES is defined as students with an expected family contribution (EFC) less than \$30,000. Higher SES are those with an EFC greater than or equal to \$30,000, and includes those who have a missing EFC. Circles represent local averages over a 0.1 GPA range. Figures are drawn using a linear fit on either side of the cutoff. Estimates and standard errors (in parentheses) are reported above each figure.

Figure A14: DiRD Estimates for Earnings Outcomes by Bandwidth

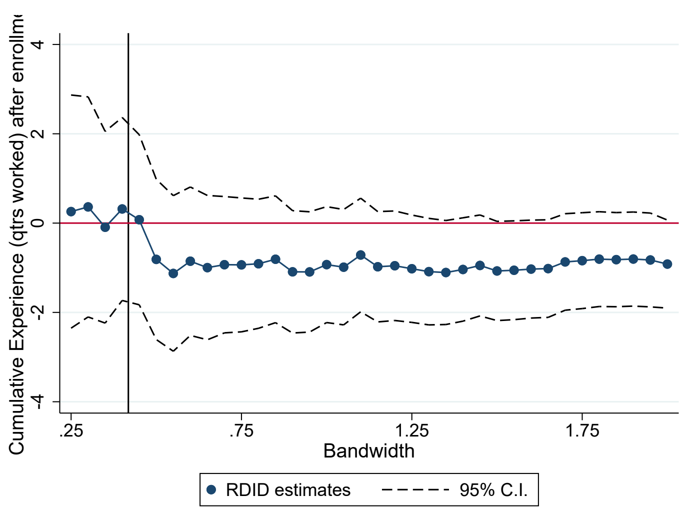
(a) Log Earnings



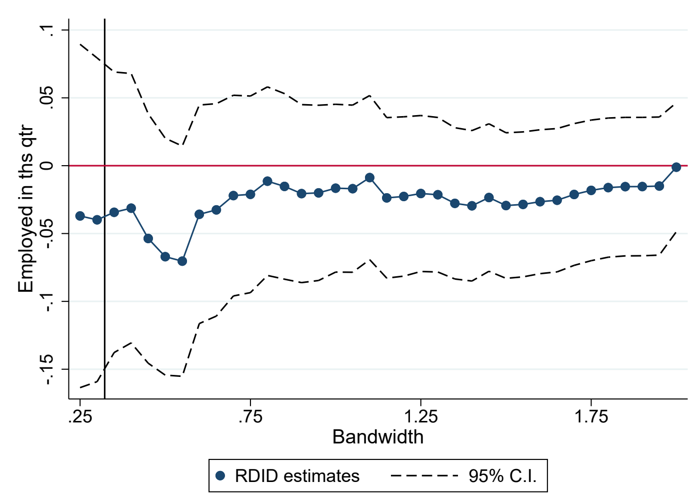
(b) AKM Estimates



(c) Cum.Experience (Qts Worked)



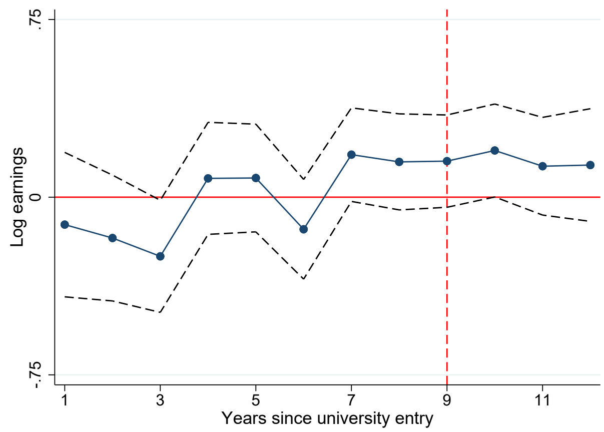
(d) Employed



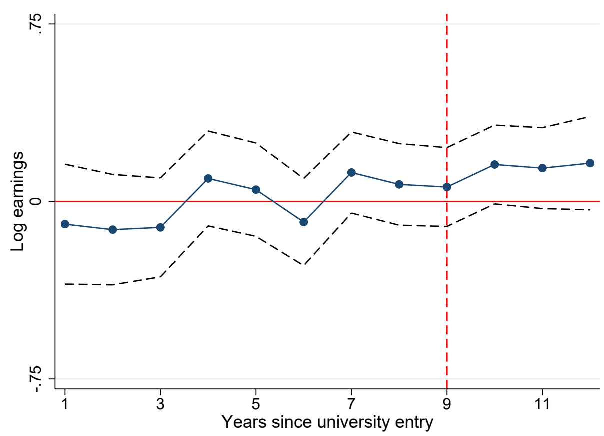
Notes: The sample is limited to quarters 7-9 years since the student enrolled at the university where the unit of observation is the student-quarter. The last calendar quarter included in the sample is 2022, Quarter 4. Dollar amounts have been adjusted to 2019 dollars using the Consumer Price Index for All Urban Consumers. Control variables include high school GPA, gender, non-white, Math and English remedial status, Pell Grant eligibility status, whether EFC is less than $30,000, indicators for whether parents attended college, and cohort and college fixed effects. Dashed lines represent 95% confidence intervals. DiRD estimates are equivalent to differencing two local linear RD regressions. The vertical black lines indicate the optimal bandwidth calculated using the CCT procedure as described in Calonico et al. (2014).

Figure A15: DiRD Wage Estimates Relative to University Enrollment Year for All Students

(a) Bandwidth=0.75 GPA points



(b) Bandwidth=1 GPA point



Notes: The sample is limited to quarters 7-9 years since the student enrolled at the university and to students who enrolled in or before 2013–where the unit of observation is the student-quarter. The last calendar quarter included in the sample is 2022, Quarter 4. The red-dashed line reported at 9 years represents the last year under which the sample is balanced. All longer run wage estimates from years 10 through 12 are from unbalanced regressions. Dollar amounts have been adjusted to 2019 dollars using the Consumer Price Index for All Urban Consumers. Control variables include high school GPA, gender, non-white, Math and English remedial status, Pell Grant eligibility status, whether EFC is less than $30,000, indicators for whether parents attended college, and cohort and college fixed effects. All regressions use a bandwidth of 0.75 or 1 grade point on either side of the cutoff. Standard errors are cluster-robust at the student level, and regressions are weighted by one over the number of quarters in which a given student is present in the regression sample. *** p<0.01, ** p<0.05, * p<0.1.

B Appendix Tables

Table B1: Covariates by Subgroup for Pre-Program and Program Years

	Female		Male		STEM		Non-STEM		Low-SES		High-SES	
	Pre	Prog.	Pre	Prog.	Pre	Prog.	Pre	Prog.	Pre	Prog.	Pre	Prog.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
HS GPA	3.69	3.79	3.59	3.72	3.69	3.83	3.56	3.68	3.64	3.77	3.63	3.74
	[0.41]	[0.41]	[0.42]	[0.46]	[0.41]	[0.48]	[0.41]	[0.40]	[0.41]	[0.43]	[0.42]	[0.45]
Female	1.00	1.00	0.00	0.00	0.32	0.35	0.56	0.52	0.43	0.46	0.42	0.43
	[0.00]	[0.00]	[0.00]	[0.00]	[0.47]	[0.48]	[0.50]	[0.50]]	[0.50]	[0.50]	[0.49]	[0.49]
Non-White	0.35	0.45	0.38	0.43	0.41	0.49	0.32	0.39	0.51	0.55	0.29	0.35
	[0.48]	[0.50]	[0.49]	[0.49]	[0.49]	[0.50]	[0.47]	[0.49]	[0.50]	[0.50]	[0.46]	[0.48]
Remedial Math	0.08	0.03	0.04	0.02	0.03	0.01	0.09	0.04	0.06	0.03	0.05	0.02
	[0.27]	[0.17]	[0.19]	[0.14]	[0.16]	[0.10]	[0.29]	[0.19]	[0.24]	[0.17]	[0.22]	[0.14]
Remedial English	0.12	0.04	0.14	0.05	0.13	0.04	0.14	0.06	0.20	0.06	0.10	0.04
	[0.33]	[0.20]	[0.35]	[0.22]	[0.33]	[0.19]	[0.35]	[0.23]	[0.40]	[0.24]	[0.30]	[0.19]
Pell Grant Eligible	0.15	0.21	0.14	0.19	0.17	0.22	0.12	0.18	0.41	0.45	0.00	0.00
	[0.36]	[0.41]	[0.35]	[0.39]	[0.37]	[0.42]	[0.32]	[0.39]	[0.49]	[0.50]	[0.00]	[0.02]
EFC $< \$30,000$	0.36	0.47	0.35	0.43	0.39	0.47	0.31	0.42	1.00	1.00	0.00	0.00
	[0.48]	[0.50]	[0.48]	[0.50]	[0.49]	[0.50]	[0.46]	[0.49]	[0.00]	[0.00]	[0.00]	[0.00]
Father College +	0.77	0.76	0.80	0.77	0.77	0.75	0.81	0.78	0.62	0.62	0.88	0.88
	[0.42]	[0.43]	[0.40]	[0.42]	[0.42]	[0.44]	[0.39]	[0.41]	[0.49]	[0.48]	[0.33]	[0.32]
Mother College +	0.80	0.79	0.80	0.80	0.78	0.78	0.83	0.81	0.66	0.68	0.88	0.89
	[0.40]	[0.40]	[0.40]	[0.40]	[0.42]	[0.41]	[0.37]	[0.39]	[0.47]	[0.47]	[0.33]	[0.31]
Obs.	1,969	7,829	2,659	9,768	2,592	8,143	2,036	9,454	1,639	7,852	2,989	9,745

Notes: The sample includes all first-time freshmen enrolled at the university in the entering fall cohorts 2007-2017 within 1 bandwidth of the 2.0 GPA threshold. "Pre" indicates pre-program years and "Prog." indicates program years. Standard deviations are in brackets.

Table B2: DiRD Estimates for Dropout in Yrs 1-3

	(1) Dropout Yr 1	(2) Dropout Yr 2	(3) Dropout Yr 3
Panel A: Bandwidth= 0.75			
	-0.075**	0.014	0.000
	(0.035)	(0.030)	(0.019)
Panel B: Bandwidth= 1			
	-0.074**	0.032	0.011
	(0.033)	(0.027)	(0.018)
Observations (BW=0.75)	12,978	12,978	12,978
Observations (BW=1)	20,009	20,009	20,009
Control Mean	0.22	0.12	0.05

Notes: The sample includes all first-time freshmen enrolled at the university in the entering fall cohorts 2007-2016. Dropout Yr t takes the value of 1 if a student has non-zero credits attempted in all years from entry including t, attempts no credits in year $t+1$, and does not eventually graduate, and zero otherwise. Control variables include high school GPA, gender, non-white, Math and English remedial status, Pell Grant eligibility status, whether EFC is less than $\$30,000$, indicators for whether parents attended college, and cohort and college fixed effects. Each point estimate is from a separate regression. DiRD estimates are equivalent to differencing two local linear RD regressions. The reported control mean is the average outcome of the pre-period cohorts with Q1 GPA $\in [1.75, 2.0)$. Robust standard errors reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B3: DiRD Estimates for Academic Outcomes by Gender and Field of Study (Bandwidth=1)

	(1) Dropout Q1	(2) Dropout Q2	(3) Dropout Q3	(4) GPA (Q2 + Q3)	(5) Earned Credits (Q2 + Q3)	(6) Grad. 4 yr.	(7) Grad. 6 yr.
Female	-0.020	0.002	-0.077	-0.152	0.837	-0.000	0.044
	(0.025)	(0.031)	(0.052)	(0.124)	(1.033)	(0.059)	(0.065)
Male	-0.001	0.001	-0.084**	0.303***	1.316*	-0.012	0.030
	(0.016)	(0.022)	(0.041)	(0.091)	(0.715)	(0.031)	(0.051)
STEM	-0.003	0.010	-0.125***	0.236**	0.962	-0.018	0.037
	(0.018)	(0.024)	(0.043)	(0.097)	(0.756)	(0.033)	(0.054)
Non-STEM	-0.018	-0.016	-0.022	-0.010	1.450	0.010	0.034
	(0.022)	(0.027)	(0.045)	(0.113)	(0.952)	(0.052)	(0.060)
Obs. (Female)	9,798	9,798	9,798	9,465	9,798	7,926	7,319
Control Mean	0.06	0.07	0.20	-0.54	23.21	0.21	0.63
Obs. (Male)	12,427	12,427	12,427	11,956	12,427	10,355	9,210
Control Mean	0.03	0.04	0.25	-0.98	22.54	0.08	0.51
Obs. (STEM)	10,735	10,735	10,735	10,339	10,735	8,971	8,009
Control Mean	0.04	0.05	0.27	-0.95	22.42	0.09	0.49
Obs. (Non-STEM)	11,490	11,490	11,490	11,082	11,490	9,310	8,520
Control Mean	0.04	0.05	0.16	-0.66	23.32	0.18	0.64

Notes: The sample includes all first-time freshmen enrolled at the university in the entering fall cohorts 2007-2017. Control variables include high school GPA, gender, non-white, math and English remedial status, Pell Grant eligibility status, whether EFC is less than $30,000, indicators for whether parents attended college, and cohort and college fixed effects. All regressions use a bandwidth of 1 grade point on either side of the cutoff. The STEM subsample includes students in the college of engineering, architecture and sciences. The non-STEM subsample includes students who are in the college of agriculture, business and liberal arts. Lower SES is defined as students with an expected family contribution (EFC) less than $30,000. Higher SES are those with an EFC greater than or equal to $30,000, and includes those who have a missing EFC. Each point estimate is from a separate regression. DiRD estimates are equivalent to differencing two local linear RD regressions. Table B8 shows that these results are robust to dropping the 2009 cohort, the year of the pilot program. The reported control mean is the average outcome of the pre-period cohorts with Q1 GPA $\in [1.75, 2.0)$. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B4: DiRD Estimates for Academic Outcomes by Gender and Field of Study (Bandwidth=0.75)

	(1) Dropout Q1	(2) Dropout Q2	(3) Dropout Q3	(4) GPA (Q2 + Q3)	(5) Earned Credits (Q2 + Q3)	(6) Grad. 4 yr.	(7) Grad. 6 yr.
Female	-0.012	0.011	-0.077	-0.118	0.805	0.013	0.102
	(0.031)	(0.037)	(0.061)	(0.150)	(1.234)	(0.071)	(0.078)
Male	-0.012	-0.008	-0.062	0.297***	1.561*	0.002	0.055
	(0.018)	(0.025)	(0.046)	(0.105)	(0.824)	(0.037)	(0.059)
STEM	-0.004	0.005	-0.120**	0.290**	1.414	-0.009	0.112*
	(0.021)	(0.028)	(0.051)	(0.113)	(0.888)	(0.038)	(0.064)
Non-STEM	-0.028	-0.010	-0.006	-0.080	0.969	0.022	0.015
	(0.024)	(0.031)	(0.050)	(0.130)	(1.083)	(0.061)	(0.070)
Obs. (Female)	6,004	6,004	6,004	5,780	6,004	4,893	4,468
Control Mean	0.06	0.07	0.20	-0.54	23.21	0.21	0.63
Obs. (Male)	8,403	8,403	8,403	8,041	8,403	7,002	6,179
Control Mean	0.03	0.04	0.25	-0.98	22.54	0.08	0.51
Obs. (STEM	7,032	7,032	7,032	6,743	7,032	5,898	5,221
Control Mean	0.04	0.05	0.27	-0.95	22.42	0.09	0.49
Obs. (Non-STEM)	7,375	7,375	7,375	7,078	7,375	5,997	5,426
Control Mean	0.04	0.05	0.16	-0.66	23.32	0.18	0.64

Notes: The sample includes all first-time freshmen enrolled at the university in the entering fall cohorts 2007-2017. Control variables include high school GPA, gender, non-white, Math and English remedial status, Pell Grant eligibility status, whether EFC scores are missing, indicators for whether parents attended college, and cohort and college fixed effects. All regressions use a bandwidth of 0.75 grade points on either side of the cutoff. The STEM subsample includes students in the college of engineering, architecture and sciences. The non-STEM subsample includes students who are in the college of agriculture, business and liberal arts. Lower SES is defined as students with an expected family contribution (EFC) less than \$30,000. Higher SES are those with an EFC greater than or equal to \$30,000, and includes those who have a missing EFC. Each point estimate is from a separate regression. DiRD estimates are equivalent to differencing two local linear RD regressions. The reported control mean is the average outcome of the pre-period cohorts with Q1 GPA $\in [1.75, 2.0)$. Robust standard errors reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B5: 'Donut' DiRD Estimates

	(1) Dropout Q1	(2) Dropout Q2	(3) Dropout Q3	(4) GPA (Q2 + Q3)	(5) Earned Credits (Q2 + Q3)	(6) Grad. 4 yr.	(7) Grad. 6 yr.
Bandwidth= 0.5	-0.017	-0.013	-0.094**	0.192*	1.435*	0.034	0.100
	(0.018)	(0.024)	(0.046)	(0.114)	(0.850)	(0.048)	(0.062)
With Controls	-0.016	-0.011	-0.094**	0.212*	1.403*	0.035	0.087
	(0.018)	(0.024)	(0.045)	(0.110)	(0.849)	(0.045)	(0.060)
Observations	8,369	8,369	8,369	7,989	8,369	6,930	6,131
Bandwidth= 0.75	-0.012	-0.005	-0.080**	0.174*	1.452**	0.041	0.079
	(0.016)	(0.021)	(0.038)	(0.093)	(0.706)	(0.039)	(0.050)
With Controls	-0.012	-0.005	-0.075**	0.149*	1.391**	0.026	0.065
	(0.016)	(0.021)	(0.037)	(0.089)	(0.703)	(0.036)	(0.048)
Observations	13,803	13,803	13,803	13,238	13,803	11,358	10,165
Bandwidth= 1	-0.009	-0.002	-0.090***	0.165**	1.304**	0.023	0.047
	(0.014)	(0.018)	(0.032)	(0.078)	(0.599)	(0.032)	(0.042)
With Controls	-0.009	-0.002	-0.087***	0.141*	1.238**	0.002	0.040
	(0.014)	(0.018)	(0.032)	(0.075)	(0.595)	(0.030)	(0.041)
Observations	21,621	21,621	21,621	20,838	21,621	17,744	16,047

Notes: The sample includes all first-time freshmen enrolled at the university in the entering fall cohorts 2007-2017. Control variables include high school GPA, gender, non-white, Math and English remedial status, Pell Grant eligibility status, whether EFC is less than $\$30,000$, indicators for whether parents attended college, and cohort and college fixed effects. Each point estimate is from a separate regression. 'Donut' DiRD estimates are equivalent to differencing two local linear RD regressions after excluding the heaping point at GPA = 2.0. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B6: Predicted Outcomes Based on Baseline Characteristics (Test of RD Assumption)

	(1) Predicted Dropout	(2) Predicted Q2 + Q3 GPA
Bandwidth= 0.5	0.004	-0.050
	(0.004)	(0.042)
Bandwidth= 0.75	-0.002	0.004
	(0.003)	(0.035)
Bandwidth= 1	-0.001	0.009
	(0.003)	(0.030)
Observations (BW=0.5)	8,973	8,973
Observations (BW=0.75)	14,407	14,407
Observations (BW=1)	22,225	22,225

Notes: The sample includes all first-year students entering the university in the fall cohorts 2007-2017. All outcomes are predicted based on the following control variables: high school GPA, whether a student is non-white, gender, Math and English remedial status, Pell eligibility status, whether EFC is less than $\$30,000$, indicators for whether parents attended college, and cohort and college fixed effects. Each point estimate is from a separate regression. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B7: DiRD Estimates Placebo Test: Running Variable=Yr 2, Q1 GPA

	(1) Dropout Yr 2
Panel A: Bandwidth=0.75	
All Students	-0.014
	(0.033)
Low SES	0.032
	(0.053)
High SES	-0.039
	(0.042)
Obs. (All)	12,910
Obs. (Low SES)	5,453
Obs. (High SES)	7,457
Panel B: Bandwidth=1	
All Students	-0.016
	(0.029)
Low SES	0.017
	(0.047)
High SES	-0.032
	(0.037)
Obs. (All)	18,452
Obs. (Low SES)	7,676
Obs. (High SES)	10,776

Notes: The sample includes all first-time freshmen enrolled at the university in the entering fall cohorts 2007-2015. Control variables include high school GPA, gender, non-white, Math and English remedial status, Pell Grant eligibility status, whether EFC is less than $\$30,000$, indicators for whether parents attended college, and cohort and college fixed effects. Each point estimate is from a separate regression. DiRD estimates are equivalent to differencing two local linear RD regressions. There is no SP for students who are placed on probation in their second year. As such, the DiRD estimates reported in this table act as a placebo test. Note that the sample is smaller than in Table 2 because some students dropped out during the first year and we cannot include the 2018 cohort when studying year 2 outcomes. *** p<0.01, **

Table B8: DiRD Estimates for Academic Outcomes by Subgroup Excluding 2009 Cohort

	(1) Dropout Q1	(2) Dropout Q2	(3) Dropout Q3	(4) GPA (Q2 + Q3)	(5) Earned Credits (Q2 + Q3)	(6) Grad. 4 yr.	(7) Grad. 6 yr.
All Students	-0.008	0.000	-0.081**	0.156**	1.054*	-0.019	0.040
	(0.014)	(0.019)	(0.033)	(0.077)	(0.611)	(0.030)	(0.042)
Female	-0.022	0.013	-0.054	-0.108	0.451	-0.024	0.062
	(0.027)	(0.035)	(0.055)	(0.136)	(1.146)	(0.063)	(0.071)
Male	-0.001	-0.004	-0.086**	0.283***	1.314*	-0.010	0.017
	(0.017)	(0.022)	(0.042)	(0.094)	(0.729)	(0.033)	(0.053)
STEM	0.002	0.013	-0.115**	0.202**	0.637	-0.032	0.025
	(0.018)	(0.024)	(0.045)	(0.101)	(0.785)	(0.034)	(0.057)
Non-STEM	-0.026	-0.027	-0.029	0.039	1.586	0.020	0.041
	(0.024)	(0.030)	(0.048)	(0.121)	(1.003)	(0.054)	(0.064)
Lower SES	-0.002	-0.004	-0.119**	0.222*	1.542*	-0.081*	0.145**
	(0.020)	(0.027)	(0.052)	(0.120)	(0.923)	(0.044)	(0.064)
Higher SES	-0.013	0.003	-0.056	0.109	0.599	0.029	-0.028
	(0.020)	(0.026)	(0.044)	(0.100)	(0.824)	(0.042)	(0.055)
Obs. (All)	20,291	20,291	20,291	19,575	20,291	16,347	14,595
Obs. (Female)	8,953	8,953	8,953	8,646	8,953	7,081	6,474
Obs. (Male)	11,338	11,338	11,338	10,929	11,338	9,266	8,121
Obs. (STEM)	9,770	9,770	9,770	9,432	9,770	8,006	7,044
Obs. (Non-STEM)	10,521	10,521	10,521	10,143	10,521	8,341	7,551
Obs. (Lower SES)	8,715	8,715	8,715	8,424	8,715	6,996	6,178
Obs. (Higher SES)	11,576	11,576	11,576	11,151	11,576	9,351	8,417

Notes: The sample includes all first-time freshmen enrolled at the university in the entering fall cohorts 2007-2017. Control variables include high school GPA, gender, non-white, math and English remedial status, Pell Grant eligibility status, whether EFC is less than $30,000, indicators for whether parents attended college, and cohort and college fixed effects. All regressions use a bandwidth of 1 grade point on either side of the cutoff. The STEM subsample includes students in the college of engineering, architecture and sciences. The non-STEM subsample includes students who are in the college of agriculture, business and liberal arts. Lower SES is defined as students with an expected family contribution (EFC) less than $30,000. Higher SES are those with an EFC greater than or equal to $30,000, and includes those who have a missing EFC. Each point estimate is from a separate regression. DiRD estimates are equivalent to differencing two local linear RD regressions. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B9: DiRD Estimates for Labor Market Outcomes: Pre-covid (through 2019 Q4)

	(1)	(2)	(3)	(4)
	Log Earnings	Firm Pay Premia	Employed	Cumulative Experience (Qtrs)
All Students	0.0951	0.0428	–0.0203	–1.28*
	(0.0961)	(0.0335)	(0.0342)	(0.699)
Lower SES	0.311**	0.0576	0.0195	0.0602
	(0.154)	(0.0503)	(0.0516)	(1.08)
Higher SES	–0.0740	0.0293	–0.0489	–2.22**
	(0.121)	(0.0449)	(0.0458)	(0.914)
Obs. (All)	92,200	92,069	127,626	127,626
Control Mean	13,030	0.574	0.729	16.7
Obs. (Lower SES)	37,292	37,256	51,566	51,566
Control Mean	12,213	0.522	0.686	16.3
Obs. (Higher SES)	54,908	54,813	76,060	76,060
Control Mean	13,553	0.608	0.760	17.1

Notes: The sample is limited to quarters 7-9 years since the student enrolled at the university where the unit of observation is the student-quarter. The last calendar quarter included in the sample is 2019, Quarter 4. Dollar amounts have been adjusted to 2019 dollars using the Consumer Price Index for All Urban Consumers. Control means are calculated among students in the pre-coaching group, with GPAs just below the probation cutoff (Q1 GPA $\in [1.75,2)$). Control means for log earnings outcomes are in dollars. All regressions use a bandwidth of 1 grade points on either side of the cutoff. Standard errors are cluster-robust at the student level, and regressions are weighted by one over the number of quarters in which a given student is present in the regression sample. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B10: DiRD Estimates for Labor Market Outcomes: Cohorts $\in$ [2007, 2013]

	(1)	(2)	(3)	(4)
	Log Earnings	Firm Pay Premia	Employed	Cumulative Experience (Qtrs)
All Students	0.137	0.0495	–0.0192	–1.23*
	(0.0947)	(0.0332)	(0.0338)	(0.714)
Lower SES	0.325**	0.0539	0.0299	0.0855
	(0.152)	(0.0499)	(0.0509)	(1.10)
Higher SES	–0.00146	0.0471	–0.0554	–2.13**
	(0.119)	(0.0445)	(0.0452)	(0.938)
Female Students	–0.133	–0.0772	–0.113**	–2.33**
	(0.147)	(0.0536)	(0.0546)	(1.10)
Male Students	0.291**	0.124***	0.0263	–0.723
	(0.122)	(0.0420)	(0.0430)	(0.922)
Non-STEM Students	0.0149	–0.0122	–0.0111	–1.60
	(0.142)	(0.0504)	(0.0506)	(1.06)
STEM Students	0.222*	0.0948**	–0.0262	–0.998
	(0.131)	(0.0449)	(0.0468)	(0.979)
Obs. (All)	118,546	118,266	166,296	166,296
Control Mean)	13,030	0.574	0.729	16.7
Obs. (Lower SES)	50,050	49,953	69,984	69,984
Control Mean	12,213	0.522	0.686	16.3
Obs. (Higher SES)	68,496	68,313	96,312	96,312
Control Mean	13,553	0.608	0.760	17.1
Obs. (Female Students)	51,074	50,950	70,752	70,752
Control Mean	12,120	0.576	0.779	18.6
Obs. (Male Students)	67,472	67,316	95,544	95,544
Control Mean	13,503	0.574	0.704	15.8
Obs. (Non-STEM Students)	60,021	59,830	81,852	81,852
Control Mean	14,208	0.568	0.739	18.1
Obs. (STEM Students)	58,525	58,436	84,444	84,444
Ctrl Mean	12,310	0.578	0.723	15.9

Notes: The sample is limited to students who enrolled in or before 2013 and to quarters 7-9 years since the student enrolled at the university where the unit of observation is the student-quarter. The last calendar quarter included in the sample is 2022, Quarter 4. Dollar amounts have been adjusted to 2019 dollars using the Consumer Price Index for All Urban Consumers. Control means are calculated among students in the pre-coaching group, with GPAs just below the probation cutoff (Q1 GPA $\in$ [1.75,2)). Control means for log earnings outcomes are in dollars. All regressions use a bandwidth of 1 grade points on either side of the cutoff. Standard errors are cluster-robust at the student level, and regressions are weighted by one over the number of quarters in which a given student is present in the regression sample. *** p<0.01, ** p<0.05, * p<0.1.

Table B11: Summary Statistics for Labor Market Outcomes

	Full Sample mean/sd	Bandwidth=1 Q1 GPA$\in[1-3]$ mean/sd	Pre-program yrs. Q1 GPA$\in[1-2)$ Probation mean/sd	Pre-program yrs. Q1 GPA$\in[2-3]$ No Probation mean/sd	Program yrs. Q1 GPA$\in[1-2)$ Probation + FYSP mean/sd	Program yrs. Q1 GPA$\in[2-3]$ Neither mean/sd
Total earnings this qtr	12504.17	11659.89	9235.91	10962.19	9458.12	12411.99
	(14835.13)	(13016.54)	(11563.87)	(11526.23)	(11243.50)	(13655.63)
Log total earnings this qtr	9.51	9.43	9.12	9.34	9.22	9.52
	(0.92)	(0.93)	(1.04)	(0.93)	(1.04)	(0.89)
Employed in this qtr	0.70	0.71	0.70	0.73	0.67	0.70
	(0.46)	(0.46)	(0.46)	(0.44)	(0.47)	(0.46)
Cumulative qtrs worked	15.90	16.13	16.34	16.77	15.53	15.98
	(8.05)	(8.18)	(8.70)	(8.20)	(8.63)	(8.06)
Firm pay premium	0.74	0.71	0.60	0.67	0.64	0.74
	(0.37)	(0.37)	(0.38)	(0.35)	(0.38)	(0.36)
Age	24.84	24.86	24.93	24.92	24.84	24.83
	(0.91)	(0.92)	(0.92)	(0.92)	(0.92)	(0.92)
Student-Qtrs	407,068	204,436	11,016	40,392	21,288	116,254
Students	40,545	19,882	918	3,366	2,125	11,951

Notes: The sample is limited to years 7–9 after first enrollment at the university. Cohorts include 2007-2017 The table reports means and standard deviations in brackets. The last calendar quarter included in the sample is 2022 Q4.

Table B12: DiRD Estimates for Labor Market Outcomes by Gender and Field of Study

	(1)	(2)	(3)	(4)
	Log Earnings	Firm Pay Premia	Employed	Cumulative Experience (Qtrs)
Female Students	–0.128	–0.0502	–0.0986*	–2.05**
	(0.138)	(0.0509)	(0.0513)	(1.02)
Male Students	0.259**	0.104***	0.0212	–0.431
	(0.117)	(0.0399)	(0.0400)	(0.851)
Non-STEM Students	0.0432	0.00872	0.00830	–1.21
	(0.136)	(0.0483)	(0.0482)	(0.999)
STEM Students	0.162	0.0686	–0.0384	–0.779
	(0.126)	(0.0422)	(0.0429)	(0.895)
Obs. (Female Students)	63,036	62,745	88,400	88,400
Control Mean	12,120	0.576	0.779	18.6
Obs. (Male Students)	81,261	80,869	116,036	116,036
Control Mean	13,503	0.574	0.704	15.8
Obs. (Non-STEM Students)	75,104	74,631	103,418	103,418
Control Mean	14,208	0.568	0.739	18.1
Obs. (STEM Students)	69,193	68,983	101,018	101,018
Control Mean	12,310	0.578	0.723	15.9

Notes: The sample is limited to quarters 7-9 years since the student enrolled at the university where the unit of observation is the student-quarter. The last calendar quarter included in the sample is 2022, Quarter 4. Dollar amounts have been adjusted to 2019 dollars using the Consumer Price Index for All Urban Consumers. Control means are calculated among students in the pre-coaching group, with GPAs just below the probation cutoff (Q1 GPA $\in [1.75,2)$). Control means for log earnings outcomes are in dollars. All regressions use a bandwidth of 1 grade points on either side of the cutoff. Standard errors are cluster-robust at the student level, and regressions are weighted by one over the number of quarters in which a given student is present in the regression sample. *** p<0.01, ** p<0.05, * p<0.1.

Table B13: Windsorized DiRD Estimates for Log Earnings

	(1) Windsorized Log Earnings
All Students	0.115
	(0.085)
Low SES	0.261**
	(0.135)
High SES	0.008
	(0.042)
Female Student	-0.132
	(0.108)
Male Student	0.251**
	(0.109)
STEM student	0.166
	(0.116)
Non-STEM	0.028
	(0.128)
Obs. (All)	144,297
Obs. (Low SES)	61,715
Obs. (High SES)	82,582
Obs. (Female)	63,036
Obs. (Male)	81,261
Obs. (STEM)	69,193
Obs. (Non-STEM)	75,104

Notes: Notes: The sample is limited to quarters 7-9 years since the student enrolled at the university where the unit of observation is the student-quarter. The last calendar quarter included in the sample is 2022, Quarter 4. Dollar amounts have been adjusted to 2019 dollars using the Consumer Price Index for All Urban Consumers. All regressions use a bandwidth of 1 grade points on either side of the cutoff. Standard errors are cluster-robust at the student level, and regressions are weighted by one over the number of quarters in which a given student is present in the regression sample. *** p<0.01, ** p<0.05, * p<0.1.

Table B14: DiRD Estimates: Additional Academic Outcomes

	(1) Re-Enrollment	(2) Switch Major
All Students	-0.001	0.008
	(0.007)	(0.033)
Low-SES	0.007	0.002
	(0.006)	(0.050)
High-SES	-0.008	0.014
	(0.011)	(0.045)
Obs. (All)	20,009	20,009
Control Mean	0.003	0.22
Obs. (Low SES)	8,493	8,493
Control Mean	0.002	0.23
Obs. (High SES)	11,516	11,516
Control Mean	0.01	0.21

Notes: The sample for Column 1 includes all first-time freshmen enrolled at the university in the entering fall cohorts 2007-2016. The outcome "Re-enrollment" takes the value of 1 if a student earns no credits in year 2 and then earns credits in each quarter of year 3. Control variables include high school GPA, gender, non-white, Math and English remedial status, Pell Grant eligibility status, whether EFC is less than $\$30,000$, indicators for whether parents attended college, and cohort and college fixed effects. Each point estimate is from a separate regression. DiRD estimates are equivalent to differencing two local linear RD regressions. The reported control mean is the average outcome of the pre-period cohorts with Q1 GPA $\in [1.75, 2.0)$. Robust standard errors reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. While we do not find evidence that students adjust by switching major (column 2) in this context, it is possible that in another setting this could be a margin by which they adjust, as only 22% of students at this university switch majors compared to a national average of about 50%.

Table B15: Individual FE Estimates for Each Outcome in the PCA Indices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Effort (2013)								
	Attend Class	Post-Exam Review	Study Hours/Week	Office Hours/Week	Complete Readings	Edit Writing	Study in Advance for Exams	Prioritize Studying
Post	0.042	4.465***	1.098***	0.098***	0.121***	0.177***	0.140***	0.330***
	(0.034)	(0.411)	(0.306)	(0.027)	(0.026)	(0.033)	(0.026)	(0.034)
Pre-Program Mean	0.73	13.55	4.26	0.84	0.84	0.73	0.82	0.52
Panel B: Self-Confidence (2013)								
	Confident Taking Exams	Confident in Writing Ability						
Post	0.112***	0.065***						
	(0.034)	(0.024)						
Pre-Program Mean	0.76	0.86						
Panel C: Motivation (2013)								
	Motivated to Focus on School	Motivated to Earn Good Grades When Uninterested in Course	Study Despite Other Distractions					
Post	0.009	0.209***	0.065**					
	(0.017)	(0.034)	(0.025)					
Pre-Program Mean	0.95	0.67	0.85					
Panel D: Resources/Community (2013)								
	Ask Fam/Friends For Help	Connected to Comm. at Uni	Know About Student Services					
Post	0.167***	0.000	0.749***					
	(0.034)	(0.026)	(0.030)					
Pre-Program Mean	0.75	0.87	0.19					
Panel E: Leisure (2013)								
	Av. Hours Sleep/Night	Av. Hours Socializing Per Week	Av. Hours TV/Gaming/Soc. Media Per Week					
Post	-0.186	-0.279	-0.586*					
	(0.147)	(0.247)	(0.299)					
Pre-Program Mean	8.40	7.59	5.27					

Notes: The data come from pre- and post-SP program survey responses. An observation is a student-survey, where there are two survey waves (pre and post SP). All regressions include a post program indicator "Post" and individual fixed effects. The sample is a balanced panel of 430 observations for those participating in SP in 2013. The response rate is about 91%, as 316 students from this cohort completed the SP and 288 responded to at least one of the surveys. The survey administered in 2103 includes a different set of questions from that administered in the following years (2015-2018), as such, the responses cannot be aggregated across these two periods. Standard errors are reported in parentheses and are clustered at the individual level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table B16: Individual FE Estimates for Each Outcome in the 2013 PCA Indices (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel F: Study Habits (2013)							
	Class Notes	Retain Assign. Info	Comm. Thoughts	Recall Lecture Material	Ask Questions	Summarize Lecture Material	Adjust Learning to Teaching Style
Post	0.056**	0.074***	0.033	0.116***	0.214***	0.079***	0.242***
	(0.025)	(0.024)	(0.021)	(0.029)	(0.036)	(0.023)	(0.031)
Pre-Program Mean	0.90	0.87	0.92	0.81	0.65	0.88	0.69
Panel G: Time Management (2013)							
	Manage Time Well						
Post	0.260***						
	(0.033)						
Pre-Program Mean	0.64						
Panel H: Self Confidence (2015-18)							
	Conf. in Time Management Skills	Conf. in Decision to Attend This Uni.	Conf. Will Improve Grades	Conf. Will Grad. From This Uni.			
Post	0.180***	0.052***	-0.016**	-0.008			
	(0.021)	(0.010)	(0.007)	(0.006)			
Pre-Program Mean	0.55	0.91	0.98	0.98			
Panel I: Resources/Community (2015-18)							
	Know Not Only One on Aca. Probation	Can ID a Staff/ Faculty Who Cares	Feel Connected to This Uni.	Know of a Resource to Get Back on Track			
Post	0.228***	0.425***	-0.302***	-0.042**			
	(0.015)	(0.017)	(0.020)	(0.017)			
Pre-Program Mean	0.75	0.48	0.79	0.81			
Panel J: Motivation (2015-18)							
	Motivated to Focus on Academics						
Post	-0.148***						
	(0.013)						
Pre-Program Mean	0.97						

Notes: The data come from pre- and post-SP program survey responses. An observation is a student-survey, where there are two survey waves (pre and post SP). All regressions include a post program indicator "Post" and individual fixed effects. The sample for Panels F, G and H is a balanced panel of 430 observations for those participating in SP in 2013. The response rate is about 91%, as 316 students from this cohort completed the SP and 288 responded to at least one of the surveys. The sample for Panels I and J is a balanced panel of 1,842 observations from those participating in SP in 2015-2018. The survey administered in 2103 includes a different set of questions from that administered in the following years (2015-2018), as such, the responses cannot be aggregated across these two periods. Standard errors are reported in parentheses and are clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1

C SP Workshop Materials

Campus Resource Guide

Campus Resources (Scan QR codes to learn more!)

1-3: Study Skills: Find resources and videos on popular topics such as:

- Study strategies, text anxiety tips and study guides
- Flashcards, video tutorials, interactive exercises
- Lecture note taking
- Memorization
- Learning style (Vark Questionnaire)

Study Skills

4-5: Tutoring on Campus (Location Varies)

- Free 1-1 or group tutoring for a variety of classes
- Supplemental Workshops: 1 unit workshop to go alongside certain Science and Math classes. Recap information from class, get help with study skills, test prep, and group studying.
- Study sessions: Weekly sessions made of 8-15 students for multiple subjects. Submit a request through the portal

Tutoring

6-7: Office hours can sometimes be intimidating and confusing on what you should ask. Here are some helpful tips:

- If office hours conflict with your schedule, contact professors for an alternate time to meet. They are more than happy to help out!
- Show problems on homework or tests that you were confused about, ask the professor to walk you through each steps.
- Bring in class notes that you would like more explanations on
- Explain your study strategies and ask about additional tips or tricks

Professors are very knowledgeable in their field and know of many outside resources and sometimes even internship or research opportunities, get to know them!

8-9: Associated Students, Inc. (ASI):

- Student Government
- Clubs and Organization
- Craft Center Classes
- The Recreation Center
- Activities and events

ASI

Dean of Students

- Club Sports
- Center For Service in Action
- Center for Leadership
- Fraternity & Sorority Life

Dean of Students

10: Cross Cultural Centers(Location varies)

- **Gender Equity Center-** Educating and empowering feminist, womxnist, mujerista moments though an intersectional lens and striving for social justice.
- **Men & Masculinity-** Creates spaces to express and evaluate masculinity and intersections with other identities through programs, dialogs and trainings.
- **Multicultural Center-** Provides space and events for people across all races, ethnicities, gender, sexual orientation, disability, economic class, religion, citizenship and their intersections
- **Pride Center-** Provides brave spaces and events to all sexualities, gender identities and expressions. Check out their peer mentoring program!

CCCs

Student Academic Services (Location Varies)

- **Dream Center-** Offers an inclusive space and a multitude of events for all undocumented students, those in mixed-status families and their allies. Stop by for a space to study, or to hang out with friends!
- **Black Academic Excellence Center (BAEC)-** Offers a supportive and enriching environment to promote excellence among Black students on campus. Stop by their center to say hello or attend one of their events!

SAS

11-14: Campus Health and Wellbeing (Bldg. 27):
Health Services:

- Mostly free services
- Walk in or make an appointment for medical attention or advice
- Educational programs about drugs, alcohol, sexuality and other topics
- On site lab testing, X-rays, and Shots (e.g Flu, TB tests)
- Discounted Pharmacy
- After-hours nurse advice
- PULSE peer mentoring health education program, covering topics such as, drugs and alcohol use and/or recovery, physical and sexual health and mental well being.

Health & Wellbeing

Counseling Services:

- Individual, couples, group therapy sessions
- Emotional Well Being Workshops
- End of Quarter Survival Kit Workshops
- Clinicians specializing in: Anxiety, Eating Disorders, Multicultural issues, Trauma, Alcohol and Drug Abuse, Suicide Prevention and many more.

Counseling Services

15-18: Basic Needs & Crisis Services (Location Varies)

Food Insecurity:

- CalFresh- Provides monthly payments to eligible students that can be used where food is sold like grocery stores, and farmers markets.
- Meal Vouchers- Students experiencing short-term financial need, can dine at 805 Kitchen during the school year and The Avenue for the summer.
- Food Pantry (Bldg 27, Lower Level)- Students can access free, packed and canned foods, frozen meals and personal hygiene products.
- Food Bank Distribution- Once a week on Mott Lawn, bags of fresh produce and food for free.

Financial Hardships:

- Cares Grant- One-time grant for unexpected emergencies like, paying for tuition, academic supplies, medical expenses, emergency housing and other temporary hardships.
- Professional Clothing Closet- Free, high-quality work clothes for interviews and future internships and jobs
- Financial Aid Office: Offers daily drop in hours where students can meet with a counselor to discuss ways to cover the cost of college.

Basic Needs

15-18 (Cont.): Crisis Services

- Safer (Bldg 65, Rm 217)- Provides confidential crisis counseling, advocacy and education and support resources by state-certified advocates. Learn about your options, rights, and other resources about sexual assault or misconduct, dating or domestic violence and stalking.

Safer

Reporting Hate Crimes

- Bias Incident Report- If you believe you have witnessed an act of discrimination or harassment on or off campus, you may file a report online though the Dean of Students.

Bias Incident Report

19-20: Career Services (Bldg. 124): Drop in or make an appointment with the Freshman Focus Team or any other Career Counselor to talk about:

- Career exploration
- Major Exploration
- Interviewing skills
- Resume and Cover letter

Career Services

21-22: Advising Centers: Have a question and don't know where to start? Visit an advisor!

College Advising Center: CAED, CAFES, CENG, CLA, CSM, OCOB,

- Course planning
- Navigating your curriculum
- Major and support related classes
- Tracking progress to degree
- Concentration

Specialty Advising

- Pre-Health Career Advising
- International/Study Abroad

Success Center (Bldg 52-D37):

- Referring you to academic and/or on-campus resources
- Understanding university and college-specific policies and procedures
- Navigating tools such as PASS and Student Center
- Change of Major process
- GE Classes
- Minors
- Transfer courses

Advising

23-24: Conflict resolution Ombuds (Library 35-113):

- If you feel that you got an unfair grade in a class
- If you feel that you got treated unfairly by someone in the University Community
- If you want to discuss a sensitive question or issue

Ombuds

25: Disability Resource Center (Bldg. 124): Provides services to those with long-term or short-term disabilities:

- How to request services
- Eligibility
- Information on testing for Learning Disabilities
- Possible accommodations - in and outside the classroom
- Peer Mentor Program

DRC

xx

My Success Plan

I decided to attend University X because:

____

____

A positive experience I have had at University X:

____

____

My favorite part of University X:

____

____

During my time at University X, I am most looking forward to:

____

____

Creating S.M.A.R.T. Goals

S	M	A	R	T
Specific Make your goal detailed and specific to know what you are working towards	**Measurable** Set parameters so that you can identify tangible evidence towards achieving your goal	**Attainable** Draft realistic goals that challenge you but you are confident to achieve	**Relevant** Make sure each goal is consistent with other goals you have established and fits with your immediate and long-term plans	**Time-Bound** Set a time that you would like to achieve your goal by

<u>Original Goal:</u> **<u>SMART Goal:</u>**

I want to read more	**S**: 12 books a year **M**: 1 book a month **A**: 1 hour at lunch, 1 hour before bed **R**: More than currently reading **T**: 1 book a month, 12 a year	I am going to read 1 book a month by reading for an hour during lunch and an hour before bed for a total of 12 books a year.

xx

WINTER QUARTER SMART GOAL:

Action steps:

I will complete the following action steps:

Personal	Academic	Social
(ex: *I will create a calendar/schedule to keep me on track with attending classes & completing assignments)*	(ex: *I will use time in between classes to study, read, and review notes)*	(ex: *I will refrain from social outings, TV, parties, social media, video games, etc. until all my homework is complete for that day)*

Resources:

I will utilize the following resources to help me achieve my goal:

(Example: I will visit the Success Center, , by Week 3 to discuss Change of Major)

1. I will visit ______________________________located in_________________________by__________

   to discuss____
2. I will visit ______________________________located in_________________________by__________

   to discuss____

Challenges: (What could stop/de-motivate me along the way?)

1. ____
2. ____

Ways to overcome my challenges:

1. ____
2. ____

Who's got your back? (The person(s) in my life that I will share my action plan with and ask to help keep me accountable for accomplishing my goal & action steps)

____

I am fully committed to completing this success plan this quarter.

SIGN: ____DATE:__________

You will receive an email from your Coach in Week 5 to follow up on your goals and action steps. Once you have communicated with your coach and completed a post-survey, your requirements with the First Year Success Program will be fulfilled.

ACADEMIC SKILLS CENTER

Weekly Schedule

	Sun	Mon	Tues	Wed	Thurs	Fri	Sat
6:00 AM							
7:00 AM							
8:00 AM							
9:00 AM							
10:00 AM							
11:00 AM							
12:00 PM							
1:00 PM							
2:00 PM							
3:00 PM							
4:00 PM							
5:00 PM							
6:00 PM							
7:00 PM							
8:00 PM							
9:00 PM							
10:00 PM							
11:00 PM							
12:00 PM							
1:00 AM							
2:00 AM							

Quarter Schedule

	Sun	Mon	Tues	Wed	Thurs	Fri	Sat
Week 1							
Week 2							
Week 3							
Week 4							
Week 5							
Week 6							
Week 7							
Week 8							
Week 9							
Week 10							
Week 11							
Finals							

D SP Pre- and Post-Surveys

All surveys were administered via the online platform SurveyMonkey.

Pre-Survey for Students Qualifying for SP in Fall 2013

Q1: What is your first name?

Q2: What is your last name?

Q3: What is your university email address?

Q4: Which college are you from?

Q5: In the Freshman Success Workshop, you will participate in a small group that will be led by an academic coach. An academic coach will lead a discussion and help group members develop an action plan to achieve success after being put on academic probation. Please check all the ways you wish to work with an academic coach.

- ☐ Identify resources to improve my study skills
- ☐ Generally improve my academic performance
- ☐ Identify ways to achieve my goal GPA
- ☐ Identify why my grades do not reflect my effort
- ☐ Stay motivated and on track to achieve my academic goals
- ☐ Learn about relevant policies
- ☐ Reduce anxiety and stress about my academic performance
- ☐ Complete the Freshman Success Workshop
- ☐ Other, please explain

Q6: On average, how many hours per week do you study?

Q7: On average, how many hours do you sleep each night?

Q8: How many times did you attend faculty office hours last quarter?

Q9: How many hours each day do you spend socializing or doing extracurricular activities?

Q10: How many hours each day do you spend watching TV, going on Facebook, gaming, etc.?

Q11: Please read the below prompts and respond to each with one of the following options: "Always, Sometimes or Rarely".

- ☐ I feel motivated to focus on school.
- ☐ I complete the assigned reading for all my classes.
- ☐ My class notes help me prepare adequately for a test.
- ☐ I retain the information I read for homework assignments.
- ☐ I feel confident about my writing ability.
- ☐ I take the time to revise my writing to make it clear, correct, and consistent.
- ☐ I easily and effectively communicate my thoughts.

☐ When I do not understand my professor, I ask the right questions to clarify.

☐ I easily remember things I learn in class.

☐ At the end of a lecture, I can summarize what was presented.

☐ I feel confident when taking an exam.

☐ When I think I did poorly on a test I just finished, I go back to my notes and review all the information I had forgotten.

☐ I prepare in advance for a test rather than "cramming" the night before.

☐ I manage my time well.

☐ I change my other priorities to have enough time for studying and completing course assignments.

☐ I can successfully balance many aspects of my life (such as friends, family, school, work, extracurricular, etc.).

☐ I study even when less important things distract me.

☐ When I have to take a course that doesn't interest me, I find a way to motivate myself to earn a good grade.

☐ I attend my classes regularly.

☐ I ask for help from family members, friends, or other appropriate individuals when needed.

☐ I know about the student services offered by Cal Poly and know how to use them.

☐ I easily adjust my learning style to my instructors' teaching styles.

☐ I feel connected to a community at Cal Poly.

Post-Survey for Students Qualifying for SP in Fall 2013

Q1: What is your first name?

Q2: What is your last name?

Q3: What is your university email address?

Q4: Which college are you from?

Q5: What day and time did you attend a workshop?

Q6: Please rate your academic coach in the following areas by selecting "Excellent, Average, Below Average, or Not Applicable" for each of the following:

- ☐ Approachability
- ☐ Knowledge
- ☐ Preparation

Q7: Which part of the Freshman Success Program was most effective? (select one)

- ☐ The presentation at the beginning
- ☐ The breakout session
- ☐ Both were equally effective

Q8: What did you find most beneficial from the big session? (please select one)

- ☐ Identified resources to improve my study skills
- ☐ Identified ways to achieve my goal GPA
- ☐ Identified why my grades do not reflect my effort
- ☐ Learned about relevant policies
- ☐ Learned how to improve my academic performance
- ☐ More motivated and on track to achieve my academic goals
- ☐ Reduced anxiety and stress about my academic performance
- ☐ I didn't find anything beneficial from this session
- ☐ Other (please explain)

Q9: What did you find most beneficial from the small group breakout session? (please select one)

- ☐ Discussion with other students
- ☐ Learning about resources
- ☐ SMART goals/goal setting
- ☐ The Self-Evaluation
- ☐ I didn't find anything beneficial from this session
- ☐ Other (please explain)

Q10: As a result of attending the Freshman Success Program, I am more likely to...(check all that apply)

- ☐ Attend class
- ☐ Do the assigned reading
- ☐ Manage my time more efficiently
- ☐ Seek out resources I need
- ☐ None of the above
- ☐ Other (please explain)

Q11: Next year, if we were to incorporate a student panel (video) segment into the big presentation of previous students on academic probation, would you be interested in participating?

Q12: In what area do you think your behavior has changed the most this quarter? (please select one)

- ☐ Increased the number of hours of sleep per night
- ☐ Increased the number of hours spent studying per day
- ☐ Increased the number of visits to office hours
- ☐ Managing my time better
- ☐ Utilizing campus resources

Q13: So far this quarter, how many hours per week do you study?

Q14: So far this quarter, on average, how many hours do you sleep each night?

Q15: So far this quarter, how many times have you been to faculty office hours?

Q16: So far this quarter, how many hours each week do you spend socializing or doing extracurricular activities?

Q17: So far this quarter, how many hours each week do you spend watching TV, going on Facebook, gaming, etc.?

Q18: Please read the below prompts and respond to each with one of the following options: Always, Sometimes or Rarely.

- ☐ I feel motivated to focus on school.
- ☐ I complete the assigned reading for all of my classes.
- ☐ My class notes help me prepare adequately for a test.
- ☐ I retain the information I read for homework assignments.
- ☐ I feel confident about my writing ability.
- ☐ I take the time to revise my writing to make it clear, correct, and consistent.
- ☐ I easily and effectively communicate my thoughts.
- ☐ When I don't understand my professor, I ask the right questions to clarify.
- ☐ I easily remember things I learn in class.
- ☐ At the end of a lecture, I am able to summarize what was presented.
- ☐ I feel confident when taking an exam.

☐ When I think I did poorly on a test I just finished, I go back to my notes and locate all the information I had forgotten.

☐ I prepare in advance for a test rather than "cramming" the night before.

☐ I manage my time well.

☐ I change my other priorities to have enough time for studying and completing course assignments.

☐ I can successfully balance many aspects of my life (such as friends, family, school, work, extracurricular, etc.).

☐ I study even when less important things distract me.

☐ When I have to take a course that doesn't interest me, I can find a way to motivate myself to earn a good grade.

☐ I attend my classes regularly.

☐ I ask for help from family members, friends, or other appropriate individuals when needed.

☐ I know about the student services offered by Cal Poly and know how to use them.

☐ I easily adjust my learning style to my instructors' teaching styles.

☐ Although I exert great effort, my grades are lower than I expect them to be.

☐ I feel connected to a community at Cal Poly.

Q19: Do you have any additional comments you would like to add?

Pre-Survey for Students Qualifying for SP in Fall or Winter 2015-2018

Q1: What is your first name?

Q2: What is your last name?

Q3: What is your university email address?

Q4: Which college are you from?

Q5: What is your major?

Q6: After taking the StrengthsFinder assessment, what are your top five strengths?

Q7: Which statement best applies to you?

- ☐ I know I am not the only one on academic probation at Cal Poly.
- ☐ I feel as though I am the only one on academic probation at Cal Poly.

Q8: Which statement best applies to you?

- ☐ I can identify a staff or faculty member at Cal Poly who cares about my success.
- ☐ I am looking for a staff or faculty member at Cal Poly who cares about my success.

Q9: Looking back at Fall Quarter, were there internal factors affecting your academic performance? Mark up to three that apply to you regarding your Fall Quarter academic difficulties.

- ☐ I could not find motivation to focus on academics.
- ☐ I felt like I did not have the appropriate study skills to succeed.
- ☐ I managed my time poorly.
- ☐ I did not attend all my classes.
- ☐ I recognized that I was having difficulty, but I was not comfortable seeking campus resources.
- ☐ I focused on extracurricular activities more than I should have.
- ☐ None of the above (no internal factors affected my academic performance).
- ☐ Other (please explain).

Q10: Looking back at Fall Quarter, were there external factors affecting your academic performance? Mark up to three that apply to you regarding your Fall Quarter academic difficulties.

- ☐ I had roommate issues that kept me from studying.
- ☐ I do not like my major and, therefore, did not do well in my classes.
- ☐ I had mostly General Education classes and was not interested in my classes.
- ☐ I got sick and missed too many classes.
- ☐ I had a personal crisis and had to focus my energy in other areas besides school.
- ☐ I had a bad professor(s) during Fall Quarter, which led to me being on Academic Probation.
- ☐ I did not have a choice in my block enrolled schedule, so I didn't like the times I had classes.

☐ None of the above (no external factors affected my academic performance).

☐ Other (please explain).

Q11: Which statement best applies to you?

☐ I know of at least one campus resource that will help me get back on track.

☐ I do not know of at least one campus resource that will help me get back on track.

Q12: List your involvement in campus clubs, organizations, or activities.

Q13: List your interests outside of your academic life.

Q14: Which statement best applies to you?

☐ I am motivated to focus on my academics at Cal Poly.

☐ I am not motivated to focus on my academics at Cal Poly.

Q15: Which statement best applies to you?

☐ I feel connected to Cal Poly.

☐ I do not feel connected to Cal Poly.

Q16: Which statement best applies to you?

☐ I am confident in my time management skills.

☐ I am not confident in my time management skills.

Q17: How confident are you in your decision to attend Cal Poly?

Q18: How confident are you that you will be able to get your grades up enough to be taken off academic probation by the end of Winter Quarter?

Q19: How confident are you that you will graduate from Cal Poly?

Post-Survey for Students Qualifying for SP in Fall or Winter 2015-2018

Q1: What is your first name?

Q2: What is your last name?

Q3: What is your university email address?

Q4: Which college are you from?

Q5: After the First Year Success Program, which statement best applies to you?

- ☐ I know I am not the only one on academic probation at Cal Poly.
- ☐ I feel as though I am the only one on academic probation at Cal Poly.

Q6: After the First Year Success Program, which statement best applies to you?

- ☐ I can identify a staff or faculty member at Cal Poly who cares about my success.
- ☐ I still have not yet found a staff or faculty member at Cal Poly who cares about my success.

Q7: Looking back at Fall Quarter, were there internal factors affecting your academic performance? Mark up to three that apply to you regarding your Fall Quarter academic difficulties.

- ☐ I could not find motivation to focus on academics.
- ☐ I felt like I did not have the appropriate study skills to succeed.
- ☐ I managed my time poorly.
- ☐ I did not attend all my classes.
- ☐ I recognized that I was having difficulty, but I was not comfortable seeking campus resources.
- ☐ I focused on extracurricular activities more than I should have.
- ☐ None of the above (no internal factors affected my academic performance).
- ☐ Other (please explain).

Q8: Looking back at Fall Quarter, were there external factors affecting your academic performance? Mark up to three that apply to you regarding your Fall Quarter academic difficulties.

- ☐ I had roommate issues that kept me from studying.
- ☐ I do not like my major and, therefore, did not do well in my classes.
- ☐ I had mostly General Education classes and was not interested in my classes.
- ☐ I got sick and missed too many classes.
- ☐ I had a personal crisis and had to focus my energy in other areas besides school.
- ☐ I had a bad professor(s) during Fall Quarter, which led to me being on Academic Probation.
- ☐ I did not have a choice in my block enrolled schedule, so I didn't like the times I had classes.
- ☐ None of the above (no external factors affected my academic performance).

☐ Other (please explain).

Q9: Do you feel that incorporating your top five strengths helped you come up with a relevant and productive Winter Quarter goal?

☐ Yes
☐ No
☐ Other (please explain)

Q10: After the First Year Success Program, which statement best applies to you?

☐ I know of at least one campus resource that will help me get back on track.
☐ I do not know of at least one campus resource that will help me get back on track.

Q11: Which statement best applies to you?

☐ After identifying a resource (academic advising, Career Services, professor's office hours, etc.) in the First Year Success Program, I have not utilized this resource by the time of completing this survey.
☐ After identifying a resource (academic advising, Career Services, professor's office hours, etc.) in the First Year Success Program, I have utilized this resource by the time of completing this survey.

Q12: After the First Year Success Program, which statement best applies to you?

☐ I am more motivated to focus on my academics at Cal Poly.
☐ I am equally as motivated to focus on my academics at Cal Poly as before the program.

Q13: After the First Year Success Program, which statement best applies to you?

☐ I feel more connected to Cal Poly.
☐ I feel equally as connected to Cal Poly as before the program.

Q14: After the First Year Success Program, which statement best applies to you?

☐ I am more confident in my time management skills.
☐ I am equally as confident in my time management skills as before the program.

Q15: After the First Year Success Program, how confident are you in your decision to attend Cal Poly?

Q16: After the First Year Success Program, how confident are you that you will be able to get your grades up enough to be taken off academic probation by the end of Winter Quarter?

☐ Very confident
☐ Confident
☐ Somewhat confident
☐ Not confident

Q17: After the First Year Success Program, how confident are you that you will graduate from Cal Poly?

E Estimating Firm-Specific Pay Premia

Since we use the same data as Flamang and Kancherla (2024) to estimate firm-specific pay premia, we closely follow their approach. Below is an overview of the statistical model, data, sample restrictions, and estimation that we use. Additional detail can be found Appendix B of Flamang and Kancherla (2024). Additional background on similar Abowd-Kramarz-Margolis (AKM) regressions more generally can be found in Abowd et al. (1999); Card et al. (2013); Song et al. (2019).

Model. We run the following regression:

$$y_{it} = \alpha_i + \psi_{\mathbf{J}(i,t)} + \mathbf{X}'_{it}\beta + \epsilon_{it} \tag{3}$$

where y_{it} is the log of worker i's earnings in year t, $\mathbf{X}$ contains calendar year dummies and controls for experience, and $\phi_{\mathbf{J}(i,t)}$ are the firm effects that we are interested in. We measure experience as the number of years since the worker first worked in CA.

Data and Sample Restrictions. Starting with the full the CA quarterly earnings data for the years 1995-2021, we first move to the worker-year level by assigning each worker to their highest paying employer in that year. We next impose three sample restrictions:

1. Drop if earnings $<$ \$4k

2. Drop if worker is not employed by firm for $>$ 1 consecutive year

3. Drop the first two years of the data (1995-1996)

The first two restrictions are meant to remove workers that are minimially attached to their firms. The third is meant to improve the accuracy of our experience measure—since some workers who first appear in our data in 1995 or 1996 will actually have pre-1995 CA work experience that we do not observe. For the same reason, we also allow the experience effects for workers who first appear in the data in 1995-1996 to differ from the experience effects of workers who enter in 1997+ (i.e., the experience effect in $\mathbf{X}_it$ is interacted with a dummy for entrance in 1995-1996).

Estimation and Specification Checks. As emphasized in prior work using similar approaches (e.g., Abowd et al., 1999; Card et al., 2013), our AKM regression is only identified within a set of firms that are connected by workers moving between firms. Following this earlier work we limit our estimation sample to the largest connected set, which includes over 99% of worker-years in the earnings data. We then estimate the model using the zig-zag method.

The main identification assumption is the unobserved component of earnings (ϵ_{it}) is unrelated to the sequence of firms that workers move between (the so-called "exogenous mobility" assumption). Flamang and Kancherla (2024) provides a standard specificaiton check in support of this assumption, which shows that workers moving from lower to higher paying firms (as measured by the wages of their coworkers) have relatively flat earnings profiles leading up to the move, and that the same is true for workers moving in the other direction (from high to low paying firms).