Fall Forward or Spring Back?

Evaluating Student Outcomes of UC Berkeley's Fall Program for Freshmen

Patrick A. Lapid*

this version: March 13, 2018

DRAFT: Please do not cite or circulate.

Abstract

UC Berkeley Extension's Fall Program for Freshmen (FPF) is a first-year learning community for undergraduates admitted for the spring semester to the College of Letters and Science at UC Berkeley. Participants choose from a subset of introductory courses and receive advising at facilities near UC Berkeley, while living and participating in activities with other Berkeley students during the fall semester. FPF participants matriculate to the main campus in the spring semester. I assess the treatment effects of FPF on college outcomes by using regression and propensity score methods to control for students' backgrounds at admission and adjusting for differences between FPF participants and fall-semester matriculants at Berkeley. FPF participants are similar to fall-semester matriculants in their admission characteristics and predicted graduation rates. I find that FPF participants have a 3–4 percentage-point increase in their four- and six-year graduation rates compared to fall-semester matriculants, with a small decrease in college GPA among graduates (0.03 points). FPF's effects are concentrated among Berkeley students with below-median high school GPAs and SAT scores. Bias-adjusted estimates using reasonable unobservable selection assumptions are similar in magnitude and direction. These findings are robust across a variety of model specifications.

[JEL: I21, I23, I28, H75, J24]

*Office of Research, Consumer Financial Protection Bureau, patrick.lapid@cfpb.gov. This paper is the result of the author's independent research and does not necessarily represent the views of the Consumer Financial Protection Bureau or the United States. [additional acknowlegements here]

Contents

1 Introduction **1**

2 Program background **4**

3 Research design **7**
3.1 Treatment effects under selection-on-observables . . . 8
3.2 Selection in the admissions process . . . 10

4 Data and descriptive statistics **14**
4.1 Data sources . . . 14
4.2 Summary statistics . . . 15
4.3 Predicted outcomes . . . 18

5 Empirical models and findings **19**
5.1 Regression estimates of treatment effects . . . 19
5.2 Fixed-effects estimates . . . 23
5.3 Heterogeneous effect estimates . . . 25
5.4 Coefficient stability of treatment effects . . . 27
5.5 Discussion . . . 30

6 Conclusion **31**

A Alternative regression specifications **35**
A.1 Logit and probit model estimates . . . 35
A.2 Separate regressions for treatment and control groups . . . 37
A.3 Inverse-probability weighting with propensity scores . . . 39
A.4 Combining regression with inverse-probability weighting . . . 40

B FPF and other Spring enrollees **41**

C Additional figures **44**

1 Introduction

College completion has become more salient in the higher education research and policy arenas in recent years. Previously overlooked in discussions focusing on college preparation and access, which have both improved over time, there is concern that rates of degree completion have stalled in recent birth cohorts, with a growing gap in college completion rates between students from poorer and richer families (Bailey and Dynarski 2011) and longer time-to-degree for lower-income students (Bound et al. 2012). Much of this focus has been on improving graduation rates at broad

access (also known as open-access or less selective) postsecondary institutions in the United States. Both in 2007 and in 2015, six-year graduation rates were under 50 percent at broad access four-year colleges that accept 80 percent or more of applicants (Crisp et al. 2017). Yet even students at selective colleges struggle to finish their degrees: at public flagship universities, one out of five students from middle- and upper-income families with college-educated parents did not earn a bachelor's degree from any institution within six years of their initial enrollment; non-completion rates of lower-income students were even higher (Bowen et al. 2009).

In response to this "crisis in completion" (Slaughter 2017), a variety of policies and interventions have been proposed and implemented to boost completion rates. Merit aid, combined with course completion and grade requirements, can encourage students to graduate on time (Scott-Clayton 2012). Need-based aid for cost-constrained students can promote continued college attendance and help them to focus on studying versus working to pay tuition and living expenses (Castleman and Long 2016; Goldrick-Rab 2016). Other interventions aim to boost individual students' academic and psychological abilities to succeed in college (Lavecchia et al. 2016). Some institutions, such as Georgia State University, are turning to predictive analytics, combined with targeted advising and peer tutoring, to improve graduation rates, time to degree, and other student outcomes (Kamenetz 2016; Quinton 2013).

But what else can we learn from changing the structure of the undergraduate experience? How would students perform under different environments, even at the same institution? At many selective private universities and liberal arts colleges, students have smaller classes, greater access to staff and faculty advisors for assistance, and more choice constraints in course selection and major programs. Public universities have created their own undergraduate honors colleges and learning communities to replicate some of these characteristics from smaller private institutions. Structure may matter: in a comparison of Ivy League universities and liberal arts colleges, six-year graduation rate gaps between lower- and upper-income students are small or non-existent, versus significant differences in graduation rates by family income at public flagship universities. However, honors college students at these same public flagships also had little difference in completion rates by income status, in contrast to their non-honors college peers (Bowen et al. 2009).

In this paper, I compare the academic outcomes of students exposed to different first-year, fall-semester academic structures and experiences at a flagship public university. Since 1983, the

University of California, Berkeley (UC Berkeley) has offered the Fall Program for Freshmen (FPF) to first-year students who have been offered admission to UC Berkeley's College of Letters and Science in the spring semester. FPF participants choose from a limited menu of first-year courses, attend classes exclusively with other FPF students, and receive advising at FPF's offices near the main Berkeley campus during the fall semester, while living and participating in extra-curricular activities with other Berkeley students. FPF participants then matriculate to UC Berkeley in the spring semester. In contrast to honors colleges and many learning communities in other institutions, FPF participation is not completely self-selected; only applicants offered deferred admission to Berkeley in the spring semester can participate in FPF for the preceding fall term.

Because of this unique application process, I argue that the admission process to UC Berkeley is unlikely to introduce positive bias in treatment effect estimates of FPF participation on college outcomes, and that treatment effect estimates of FPF participation can be identified under a "selection-on-observables" assumption. I then provide descriptive statistics from UC Berkeley admissions and registration data, showing how fall-semester admits and matriculants have greater mean high school grade point averages and admission test scores compared to their spring-semester peers. I also show that predicted graduation rates of FPF participants, based on their admission characteristics, are similar to fall-semester matriculants.

Using regression and propensity score methods, I assess the effect of FPF participation on graduation rates and other outcomes, controlling for students' admission and demographic characteristics that may differ between FPF participants and fall-semester matriculants. My estimates show that FPF participants are more likely to graduate in four years, by about 4 percentage points compared to fall-semester matriculants, and are more likely to graduate in six years, by about 3 percentage points. FPF students' college grade point averages (GPAs) at graduation were 0.03 grade points lower, a small decrease that is only a tenth of a standard deviation of the college GPA distribution. FPF's effect on graduation rates are concentrated among students with weaker academic backgrounds, as measured by below-median high school GPAs and SAT test scores. The estimated treatment effects are of similar magnitude and direction across models, including logit, probit, the inclusion of high-school fixed effects, and a variety of propensity score weighting and matching methods.

This paper contributes to an expanding literature in the economics of higher education that

analyzes structural components of the college experience and their impacts on grades, major choice, retention, and college completion.[1] Within the higher education research community, scholars have argued for greater use of causal inference research methods to evaluate programs and institutional effectiveness.[2] Quasi-experimental techniques for analyzing observational data are useful for higher education program evaluations, either because we would like to assess the past performance of an existing program, or because an experimental design is not feasible or ethical (Reynolds and DesJardins 2009). Lastly, this paper draws upon the selection-on-observables research methods in applied econometrics, particularly evaluation methods used previously to assess the effectiveness of job training and other labor market programs, such as propensity score weighting and matching.[3]

In the next section, I discuss the institutional context of the program, followed by a discussion in section 3 of the selection on observables research design and identifying causal impacts of the program. Section 4 describes the data sources, summary statistics, and predicted outcomes of FPF participants. Section 5 goes into detail with the various empirical models and estimates I use, and section 6 concludes.

2 Program background

UC Berkeley Extension's Fall Program for Freshmen (FPF) is a semester-length learning community program offered to first-year students admitted to UC Berkeley's College of Letters and Science in the Spring semester. Both spring admittance and the Fall Program for Freshmen have taken place at UC Berkeley since 1983.[4] Students applying to UC Berkeley as prospective first-year (freshman) admits all submit applications by November 30 of their final (senior) year in high school. The

[1] Recent work includes the effects of non-tenure-track faculty versus tenure-track professors on student outcomes such as major choice and later course performance (Bettinger and Long 2010; Figlio et al. 2015); modifications to student placement in and structure of developmental/remedial education courses (Bailey et al. 2015; Bettinger et al. 2013); redesigned course "pathways" to degree completion (Bailey et al. 2015; Jenkins and Rodríguez 2013); and implementing "intrusive" advising and other changes in student support (Bailey et al. 2015).

[2] Examples include Pike et al. (2011); Reynolds and DesJardins (2009); and Titus (2007).

[3] Imbens (2004) provides a detailed survey of treatment effect inference and estimation methods relying on selection-on-observables assumptions of exogeneity. Reynolds and DesJardins (2009) discuss applications of these methods for higher education research.

[4] At UC Berkeley and other higher education institutions, spring-term admittance is primarily used to maintain enrollment capacity between academic terms. Graduations and other departures from the university after the fall term result in lower tuition and financial aid revenue, unused or underused instructional space, and spare housing capacity, all from reduced enrollment.

majority of first-year students are admitted to the College of Letters and Science, which houses the bulk of academic departments in the liberal arts at UC Berkeley.[5]

The admissions process has two selection stages that separate students into fall-semester matriculant and FPF participant cohorts. The first stage is when UC Berkeley's Office of Undergraduate Admissions reviews applications and makes one of the following choices for each applicant: admission for the fall semester, admission for the spring semester, or rejection. Students are notified of this decision by the end of March. The second stage is slightly different between admission groups. Fall-semester admits choose to attend UC Berkeley or not; spring-semester admits choose to participate in FPF in the fall semester and matriculate to UC Berkeley in the following spring semester, to matriculate to UC Berkeley in the spring semester only,[6] or to not attend UC Berkeley.

In the second stage, fall-semester admitted students decide to attend UC Berkeley by submitting a Statement of Intent to Register (SIR) before May 1 and officially matriculating (registering) at UC Berkeley. Spring-semester admits also submit a SIR before May 1 to matriculate, but they also are informed about the Fall Program for Freshmen within a few days of receiving their admission offer. FPF positions are first-come, first-served; potential participants can sign up before May 1. This results in some spring admits signing up for FPF but eventually not matriculating to UC Berkeley. Since there are approximately 600–700 FPF slots but over 2,000 spring-semester admits each year, the FPF staff maintain a waiting list of all interested students. Even so, there are some spring-semester matriculants who would like to join FPF but are unable to get in, as well as other spring-semester matriculants who do not express interest in FPF.[7]

In FPF, participants can choose from a subset of introductory Berkeley courses for first-year students. FPF instruction and advising occurs at a facility four blocks south of the Berkeley

[5] Other colleges at UC Berkeley open to undergraduates (such as the Colleges of Chemistry, Engineering, Environmental Design, and Natural Resources) only admit first-year students during the fall semester.

[6] These students are notified that they can travel, work, and/or take classes at a community college or other institution during the fall semester.

[7] The following section discusses applicant admission characteristics in more detail.

Table 1 shows first-year admission pool data for UC Berkeley's College of Letters and Science. The first-year admission pool between the 2006-2007 and 2008-2009 academic years increased from 30,000 to 34,000 applicants per year to the College of Letters and Science, with approximately 30 percent in each cohort offered admission to UC Berkeley. Three-quarters of these offers were for regular admission during the fall semester; the remainder of admission offers were deferred to the spring semester. Student yield (the total number of students who register to attend UC Berkeley, divided by the total number of admitted first-year students) for the College of Letters and Science is slightly over 40 percent of all admitted students. Yield rates are similar across academic years, for the fall and spring semesters.

campus. FPF students can apply for student housing and are able to participate in social and cultural activities with regular UC Berkeley students. The courses offered through FPF are all approved by their academic departments at UC Berkeley. FPF courses must use the same textbook as the regular Berkeley course and follow similar course content and topics. FPF instructors hired by UC Berkeley Extension must also be approved by the relevant UC Berkeley academic department and the Academic Senate. These instructors primarily consist of graduate students from UC Berkeley with prior teaching background, as well as experienced teaching faculty from other colleges in the San Francisco Bay Area. Courses in FPF are smaller in size compared to regular UC Berkeley courses; for example, introductory lecture courses in Mathematics, Psychology, and Statistics range from 60 to 100 students in FPF, versus 300 to 400 students on the main campus.

FPF also has its own advising staff that work with students before, during, and after their participation in FPF. This is in addition to the standard advising, orientation, and other student services open to all Berkeley students. One key FPF advising policy is that students who have under a 2.0 grade-point-average (GPA) in their coursework by the eighth week of instruction (roughly mid-semester) must meet with their advisor for guidance and assistance to improve their academic performance. In addition, informal discussion between program instructors and advisors occurs regularly during the program to discuss student performance.

Tuition costs and housing policies are slightly different between regular Berkeley students and FPF participants. Since UC Berkeley Extension programs are self-supporting, no state aid directly supports the instructional and advising staff for FPF. In addition, FPF and Extension must negotiate access to campus student services that regular UC Berkeley undergraduates receive. Both factors result in slightly higher tuition prices for students participating in FPF.[8] Before 2003, FPF students and their families paid out-of-pocket for program costs and housing, which may have constrained lower-income students from participating in FPF. However from 2003 onward, FPF participants can use federal and state financial aid to pay for program tuition and other expenses, similar to fall-semester matriculants to UC Berkeley. The other main difference between FPF participants and other first-year Fall students is UC Berkeley's housing priority for these two groups. Before 2011, FPF participants' priority for UC Berkeley's campus housing was lower than for fall-semester

[8] The FPF difference is approximately $500 more for the fall semester during the 2006 to 2008 period studied in this paper.

matriculants; since then, both groups are treated similarly in the on-campus housing application process. Still, approximately 60 percent of FPF participants lived in campus housing during the pre-2011 period, with the remainder in the local community.

Internal analyses from UC Berkeley point out that the six-year graduation rates of FPF participants are slightly greater than for fall-semester matriculants, and much greater than for non-FPF spring-semester maticulants, in 2007 and 2008. These analyses also report higher GPAs among Berkeley alumni who participated in FPF versus other Berkeley students. It is possible that participation in FPF has a positive impact on graduation rates and other academic outcomes, assuming that FPF participants and fall-semester matriculants are comparable. In the following sections, I will describe the similarities between these two groups, the differences that come up during the admission and selection process, and the controls and adjustments I make in my empirical estimates and models that adjust for these differences when possible.

3 Research design

In this paper I seek to estimate the average treatment effects of student participation in UC Berkeley Extension's Fall Program for Freshmen on graduation and other academic outcomes. I define the treatment group as first-year, spring-semester matriculants to UC Berkeley who participate in FPF during the preceding fall semester. I will refer to students in the treatment group as "FPF participants". The primary control group in my main analysis are first-year students admitted and enrolled in UC Berkeley during the fall semester, who I will call "fall-semester matriculants". I focus on comparing outcomes between FPF participants and fall-semester matriculants because both groups of students are exposed to many similar college experiences during the fall semester, including living near UC Berkeley and participating in extracurricular activities. Academic majors and course sequences may also be designed for students who begin in the fall semester; non-FPF spring-semester matriculants to UC Berkeley may be unable to enter a course sequence if prerequisites are not taken or offered at their local community college or other institution attended in the preceding fall term.[9]

[9] Although non-FPF spring-semester matriculants make up a potential secondary control group, I believe that the differences in timing and experiences between these students and the other two groups will complicate any comparison of outcomes.

3.1 Treatment effects under selection-on-observables

In a basic treatment effect model, for any outcome of interest Y_i, every college student i has two potential outcomes that may be observed: Y_i^0, the outcome that would be observed if the student attended UC Berkeley in the fall semester, and Y_i^1, the outcome that would be observed if the student participated in FPF in the fall semester and matriculated to UC Berkeley in the spring semester. Participation in FPF is indicated by the binary treatment selection variable T_i. We then define the average treatment effect (ATE) as the mean difference in these outcomes from program participation:

$$ATE = \mathbb{E}\left[Y_i^1 - Y_i^0\right]$$

However we only observe Y_i^0 for control group members and Y_i^1 for treatment group members in the data. To estimate the ATE we have to first predict two sets of counterfactual outcomes, Y_i^1 for individuals in the control group and Y_i^0 for individuals in the treatment group.[10]

I attempt to identify the ATE under a *selection-on-observables* assumption, requiring both *conditional independence* and *common support*. The conditional independence assumption states that for individual i, binary treatment T_i, potential outcomes Y_i^1 and Y_i^0, and other observable covariates X_i:

$$\left(Y_i^0, Y_i^1\right) \perp\!\!\!\perp T_i \mid X_i.$$

This means that the potential outcomes of i are statistically independent of the treatment T_i, conditional on observable covariates X_i, such that the selection into treatment is plausibly exogenous.

The second assumption of common support (or overlap) means that no set of observable

[10] Besides estimating the ATE for FPF participation, we could also be interested in estimating the treatment-on-the-treated (TOT) effect, the benefit of FPF participation for treatment group participants only:

$$TOT = \mathbb{E}\left[Y_i^1 - Y_i^0 \mid T_i = 1\right]$$

Estimating the TOT only requires predicted outcomes of Y_i^0 for individuals in the treatment group.

I focus on estimating the ATE versus TOT, because of the many similarities I find between fall-semester matriculants and FPF participants. The TOT estimate would be relevant if major differences between the treatment and control groups of students were apparent, such that there was less overlap in academic or demographic covariates or in the propensity scores for FPF selection. In my empirical models, I find that each model's ATE and TOT estimates are similar in value.

characteristics will perfectly predict assignment to either the treatment or control group:

$$0 < \hat{P}(X) < 1, \forall X,$$

where $\hat{P}(X)$ is the predicted propensity score for assignment to the treatment group. This assumption is necessary for estimation using propensity score and matching methods, but not required for regression.[11]

Under the selection on observables assumption, treatment assignment is assumed to be exogenous conditional on the observed covariates; adjusting for these covariates (via regression, propensity score matching, and/or propensity score re-weighting) will allow us to argue that our estimates of the effect of Fall Program for Freshmen participation are causal. However, there is no guarantee that the researcher has controlled for all factors affecting both selecting into treatment and academic outcomes. The omission of unobserved variables to researchers, but possibly observed by staff in the admissions selection process, may result in bias that overestimates or underestimates the effect of FPF. This is the case if these unobservable variables are both related to being selected into FPF (T_i) as well as to outcomes (Y_i). Our estimates of FPF's effect on an academic outcome (e.g., the probability of graduating in four years) would be positively biased if an unobserved variable was correlated positively (or negatively) both with FPF selection and the outcome of interest. However, the estimated effect is negatively biased if the unobserved variable was positively correlated with the treatment but negatively correlated with the outcome (or negatively correlated with the treatment but positively correlated with the outcome).[12]

[11] With linear regression and other parametric estimation methods, if there is no common support for an observable variable X, or for particular values of X, for the treatment or control group, we could extrapolate those values from the other group where we have observable data. However, we may question the plausibility of these estimates.

[12] Specifically, let $Y_i = \alpha^L + \beta^L T_i + \gamma W_i + \varepsilon_i^L$, the long regression of the outcome Y_i on T_i, treatment into FPF, and another variable W_i. The treatment effect of interest is the coefficient β^L. However, if variable W_i is unobserved, we can only estimate the short regression $Y_i = \alpha^S + \beta^S T_i + \varepsilon_i^S$, with β^S being the coefficient of interest. The omitted variable bias formula relates β^S and β^L as follows:

$$\begin{aligned}\beta^S &= \beta^L + \pi \cdot \gamma \\ &= \beta^L + \frac{Cov(W_i, T_i)}{Var(T_i)} \cdot \frac{Cov(Y_i, W_i)}{Var(W_i)}\end{aligned}$$

The final two terms describe how unobserved W_i covaries both with the treatment T_i and outcome Y_i; if the covariance of W_i with the treatment or outcome is zero, then there is no bias from omitting W_i in the short regression, so $\beta^S = \beta^L$.

However, in the case that W_i covaries in the same direction (either positively or negatively) with both the treatment T_i and outcome Y_i, there will be a positive bias in the short regression estimate of the treatment effect, so $\beta^S > \beta^L$,

3.2 Selection in the admissions process

Keeping this in mind, we must account for selection in the college admissions process. The admission and acceptance process to UC Berkeley will complicate any analysis of the impact of the Fall Program for Freshmen on participants and comparisons to fall-semester matriculants. As mentioned before, three key stages of selection occur during the admissions process: (1) when UC Berkeley's Office of Undergraduate Admissions accepts applicants for the fall or spring semester, or rejects applicants for admission; (2) when admitted students decide to attend UC Berkeley; and for spring-semester admits who choose to attend UC Berkeley, (3) when they choose to participate in FPF and come to Berkeley in the fall semester, or pursue alternative plans and start their Berkeley classes in the spring semester. If the control condition is being offered fall-semester admission and attending UC Berkeley, the treatment is a combination of being offered spring-semester admission and attending UC Berkeley and participating in FPF during the fall semester.

At the initial selection stage, staff members in UC Berkeley's Office for Undergraduate Admissions review applicants' academic and personal backgrounds. Researchers are able to observe some applicant information, including high school GPA, SAT test scores, and demographic characteristics, but other applicant data that may be relevant in predicting college outcomes are unobserved by researchers, such as extracurricular activities, participation in college-preparatory programs, work experience, and others. Students with stronger academic backgrounds, based on observable variables, are more likely to be offered any admission to UC Berkeley (versus rejection), with higher academically performing students again more likely to receive a fall (versus spring) admission offer.

Admissions counts and academic profiles for fall and spring admits are shown in Table 1. We see that high school GPAs and SAT scores are slightly higher among fall-admitted students versus

and our short regression estimates of FPF's treatment effect will be larger than the true effect. On the other hand, if there are unobserved factors W_i to the researcher but are observed to admissions officers and admissions knows that $Cov(Y_i, W_i) > 0$, that individuals with W_i are more likely to graduate or have other positive outcomes in college, it is unlikely in this case that $Cov(T_i, W_i) > 0$, that the admissions office would also offer these individuals participation in FPF via spring admission—these students would more likely be admitted for the Fall term.

Of course, there may be an interesting case of some unobserved variable W_i not known to admissions that both positively predicts FPF participation (conversely negatively predicts Fall admission) but also has a positive relationship with graduation, GPA at graduation, and other outcomes—further work to reduce this possibility is needed.

In the other case with W_i covarying with treatment T_i in one direction, but with college outcomes Y_i in the other direction, such as with observable variables like high school GPA and SAT scores, the omitted variable bias will be negative in sign, resulting in $\beta^S < \beta^L$, an underestimate of the treatment effect of FPF toward zero. This kind of bias then works against finding an effect of FPF; a statistically significant estimate of the effect of FPF on college outcomes would be a lower bound, with the true effect possibly greater in magnitude.

spring admits, by 0.10 to 0.12 grade points and 35 to 47 SAT points. The average high school GPAs and SAT scores for both fall and spring admits are both considerably larger than for the initial application pool. If unobservable variables that are correlated with graduation are also correlated with observable variables more likely to be found (or greater in value) among fall-admit students (and less likely to be found among spring-admit students), it is possible that graduation rates and other academic outcomes are higher among fall-semester versus spring-semester admits, before taking into account FPF participation. In this example of omitted variable bias, we would be less likely to estimate a positive effect of FPF participation on academic outcomes, due to negative selection of spring-semester versus fall-semester admits by admissions staff.

In the second selection stage, students admitted to UC Berkeley for the fall and spring semesters can choose whether or not to attend UC Berkeley. In both the fall and spring admit pools, students with stronger academic and extra-curricular records are more likely to apply and receive admission to other selective universities, both within the University of California system and to other institutions. Therefore, yield rates are lower for groups of students with higher GPAs and test scores.[13] We may believe that the strongest spring admit students are particularly less likely to choose UC Berkeley, unless their desire to attend Berkeley outweighs other considerations.

As shown in Table 1, yield rates are around 40 percent of both fall and spring admits. Within the fall admit group, the average high school GPA and SAT scores for matriculants (students who accept admission to UC Berkeley and submit a Statement of Intent to Register) declines slightly for all the admission cohorts studied. This is also the case for admitted versus matriculated spring-semester students, although the decrease in average high school GPA and SAT scores is smaller in magnitude. The gap in the mean high school GPA between fall and spring matriculants and across admission cohorts is 0.04 to 0.06 grade points. For average SAT scores, the gap between fall and spring matriculants is 9 to 21 SAT points across cohorts. The observed academic profiles of spring matriculants are considerably closer to fall matriculants, compared to spring versus fall admits. Even so, the mean high school GPAs and SAT scores for spring matriculants are slightly lower than the means for fall matriculants in the same admission cohort. It appears that the negative selection bias for spring- versus fall-semester matriculants is lessened, compared to spring versus fall admits, at least in terms of academic profiles.

[13] This is based on observations from UC Berkeley admissions and institutional research staff.

Table 1: Admission and academic characteristics of UC Berkeley applicants

Admission year	2006-2007	2007-2008	2008-2009
Applicant count	30,289	31,635	34,343
Admitted			
Fall semester	7,238	7,430	7,403
Spring semester	2,112	2,272	2,198
Matriculated (SIR submitted)			
Fall semester	3,016	3,157	2,975
Spring semester	888	1,002	886
Admit rate			
Fall semester	23.9%	23.5%	21.6%
Spring semester	7.0%	7.2%	6.4%
Yield rate			
Fall semester	41.7%	42.5%	40.2%
Spring semester	42.0%	44.1%	40.3%
Weighted high school GPA (means)			
Applicants	3.92	3.93	3.93
Admitted			
Fall semester	4.32	4.32	4.34
Spring semester	4.21	4.20	4.24
Difference: Fall vs. Spring	0.11	0.12	0.10
Matriculated (SIR submitted)			
Fall semester	4.23	4.23	4.23
Spring semester	4.19	4.17	4.19
Difference: Fall vs. Spring	0.04	0.06	0.04
SAT math and verbal score (means)			
Applicants	1313	1318	1313
Admitted			
Fall semester	1408	1412	1411
Spring semester	1373	1365	1374
Difference: Fall vs. Spring	35	47	37
Matriculated (SIR submitted)			
Fall semester	1368	1376	1368
Spring semester	1353	1355	1359
Difference: Fall vs. Spring	15	21	9

Note: All counts, mean GPAs and SAT scores are for applicants, admits, and matriculated students to the College of Letters and Science. Admit rates are percentages of applicants that are admitted for the academic year. Yield rates are percentages of admitted students that matriculate to UC Berkeley for the indicated semester. Weighted high school GPA includes additional grade points for AP and honors courses taken in high school. Source: Office of Undergraduate Admissions, UC Berkeley.

In addition to choosing to attend UC Berkeley, spring matriculants decide either to attend the Fall Program for Freshmen and come to Berkeley for FPF in the fall semester, or decline (or are unable to attend) FPF and come to Berkeley a semester later during the spring. It appears that FPF participants have academic and socio-economic backgrounds that are positively correlated with graduation and academic success in college, compared to non-FPF spring matriculants. Out of the spring admit pool who choose to attend UC Berkeley, FPF participants have greater high school GPAs and SAT scores, are more likely to come from high schools with greater academic preparation and resources, and are less likely to be low-income and first-generation college students.[14] However, the majority of spring matriculants (about two-thirds) do participate in FPF, and their average characteristics are generally similar to fall matriculants, as I will discuss below in the data description and summary statistics.

Based on what we know about the admissions process and the available data, I believe the control group of fall-semester matriculants to UC Berkeley and the treatment group of FPF participants are generally similar to each other, and that the differences would not result in biased estimates of FPF having a positive impact on academic outcomes. I argue that FPF participants are negatively selected at the admissions stage, as spring-semester admits, based on academic background. This will make estimating a positive effect of FPF on academic outcomes more difficult. Also, since FPF participants mainly live in Berkeley and can participate in the same extra-curricular activities as fall-semester matriculants, differences in the college experience outside of the classroom are minimized across both groups. Even though we see a narrower gap in academic backgrounds between fall- and spring-semester students at the matriculation stage, and a positive selection effect among the Spring registrants who participate in FPF, it does not appear that FPF participants would have academically outperformed their fall-matriculant counterparts if FPF participants were instead admitted during the fall semester.

In the following section, I will provide summary statistics and predicted outcomes from the data, arguing that FPF participants' graduation rates and college GPAs would be similar to fall-semester matriculants in the absence of the program. Both groups of students also exhibit a high degree of overlap, in terms of individual characteristics as well as in estimates of the propensity of being

[14] Appendix A compares the differences across spring-semester matriculants, both FPF participants and non-participants, in academic and demographic backgrounds.

an FPF participant versus a fall-semester admit. In my empirical estimates of FPF's effects on graduation and academic outcomes, I control for observable differences in the data; we will see that these estimates change only slightly when more controls are included.

4 Data and descriptive statistics

4.1 Data sources

I analyze UC Berkeley data provided by the Office of Undergraduate Admissions and the Office of the Registrar for enrolled students starting between Fall 2006 and Spring 2009. This includes three admission-year waves, which can each be divided into fall-semester matriculants; Fall Program for Freshmen participants, who participate in FPF in Berkeley during the fall semester and enroll in regular UC Berkeley classes in the following spring semester; and spring-semester matriculants who do not participate in FPF. I will refer to these groups as fall-semester matriculants, FPF participants, and spring-only matriculants. with my primary analysis comparing outcomes between fall-semester matriculants and FPF participants.[15] Furthermore, my primary analysis sample only includes fall-semester matriculants and FPF participants who are admitted to the College of Letters and Science (L&S), the largest undergraduate college at UC Berkeley.[16] As a result, spring-semester matriculants are initially admitted to the College of Letters and Science, including all FPF participants. The main comparisons that follow will treat fall-semester matriculants in L&S as the control group, and FPF participants (also in L&S) as the treatment group, with the treatment defined as spring-semester admission and matriculation to UC Berkeley plus participation in FPF in the preceding fall semester.

Data from the Office of Undergraduate Admissions includes student academic and demographic data before enrolling at UC Berkeley, including high school grade point average (GPA), SAT I math and verbal scores, first-generation college student status, California high school academic performance index, gender, ethnicity, and California residency status. The data from the Office of the Registrar includes final college outcomes, including graduation and time to graduation, as well

[15] Appendix A provides more detail on spring-only matriculants and compares them to FPF participants.

[16] A small number of Spring admits and FPF enrollees during this period were admitted to the College of Engineering during this period. These students are not in the analysis sample; the addition of these students has little impact on the estimated effects of FPF.

as intermediate college outcomes, such as major choice, first-year probation status, and retention at UC Berkeley from one to six years after initial enrollment.

I focus on four-year and six-year graduation rates, as well as college GPA at graduation, because they are discussed in FPF's own program materials and because four-year and six-year graduation rates are key measures of on-time graduation in higher education data, such as the Integrated Postsecondary Education Data System (IPEDS) and the College Scorecard.

4.2 Summary statistics

Table 2 compares summary statistics of key outcome and background variables between Fall Program for Freshmen participants and fall-semester matriculants in the College of Letters and Science. Overall, graduation rates are high for UC Berkeley students in L&S: over 90 percent of first-year enrollees graduate from UC Berkeley in six years, and above 70 percent graduate within four years. We see that four- and six-year graduation rates are 3 to 4 percentage points higher among FPF participants compared to fall-semester matriculants.[17] FPF participants do graduate with slightly lower college grade-point averages versus fall-semester matriculants. I will show that these mean differences in outcomes will be similar to the estimated treatment effects that control for observable characteristics, as discussed in Section 5.

The overall differences in mean admission characteristics between fall-semester matriculants and FPF participants are small in magnitude. Individual academic preparation is slightly better among fall-semester matriculants than FPF participants; mean high school grade point averages and SAT scores are higher by a few points. The distribution of high school GPAs and test scores of FPF participants are less dispersed: kernel densities of high school GPAs (Figure 1) and of SAT scores (Figure 2) show that very high and very low high school GPAs and SAT scores are more likely to be found among fall-semester matriculants. These distributions suggests that FPF participants are academically similar to Fall-term students.

On the other hand, FPF participants are more likely to come from the highest academic performance decile of California public high schools, and less likely to have attended a public high school with a below-median ranking in academic performance. There are also a lower percentage of first-generation college students, defined as students with parents that have either

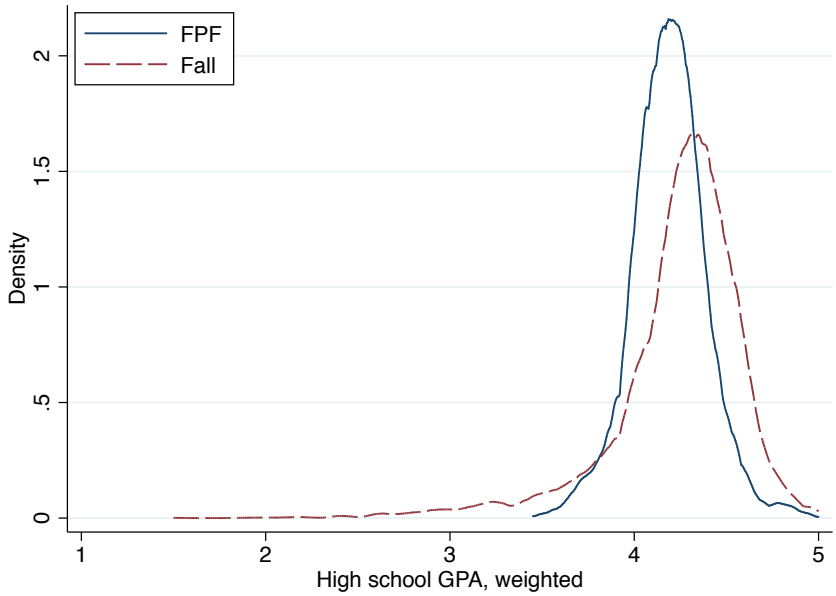
[17] FPF students also have higher graduation rates at five years, as well as at any time from UC Berkeley.

Table 2: Summary statistics

	Fall	FPF	Diff.	p-value
Outcome variables				
Graduate in 4 years	0.71	0.76	-0.04	0.00
Graduate in 6 years	0.91	0.94	-0.03	0.00
UC Berkeley GPA at graduation	3.41	3.36	0.04	0.00
Admission variables				
High school GPA, unweighted	3.81	3.77	0.03	0.00
High school GPA, weighted	4.23	4.19	0.05	0.00
SAT Math score	665	662	2.92	0.21
SAT Verbal score	640	634	5.70	0.02
API 1st-5th decile	0.21	0.15	0.05	0.00
API 6th-9th decile	0.33	0.34	-0.01	0.40
API 10th decile	0.23	0.29	-0.06	0.00
CA HS, no API	0.13	0.15	-0.01	0.09
HS outside CA	0.10	0.07	0.03	0.00
Parent education				
No 4-yr degree	0.28	0.24	0.04	0.00
HS diploma or less	0.17	0.13	0.05	0.00
Female	0.60	0.60	-0.00	0.93
CA resident	0.88	0.92	-0.04	0.00
African American/Black	0.04	0.02	0.02	0.00
Latino/Hispanic	0.12	0.09	0.03	0.00
Native American	0.00	0.00	0.00	0.05
Pacific Islander	0.00	0.00	0.00	0.33
Asian American	0.43	0.47	-0.05	0.00
White	0.30	0.32	-0.02	0.13
Other	0.07	0.09	-0.02	0.00
International	0.03	0.00	0.03	0.00
Observations	8,763	1,866		

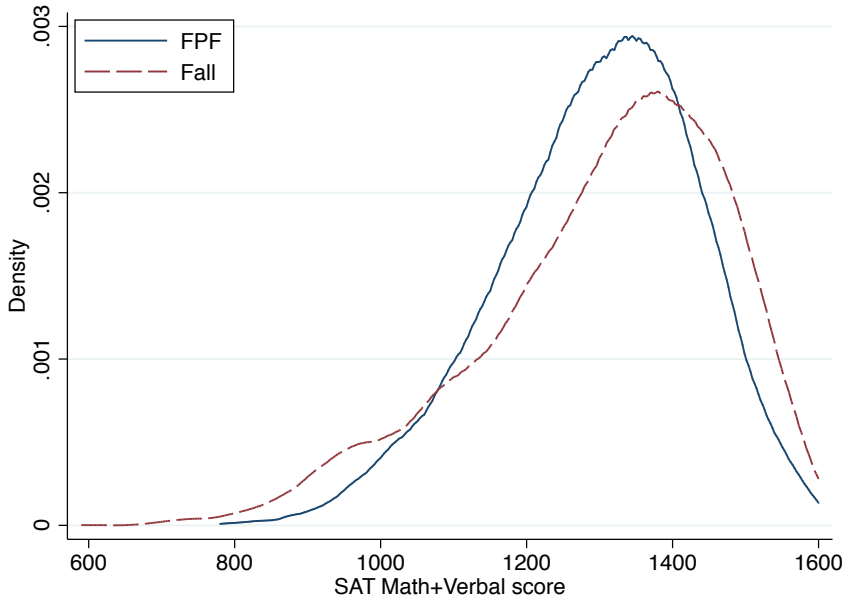
Note: Fall students and FPF participants in College of Letters and Science only. API stands for California Academic Performance Index for public schools only. Source: UC Berkeley.

Figure 1: Kernel densities of weighted high school GPA



Note: Fall students and FPF participants in College of Letters and Science only. Source: UC Berkeley.

Figure 2: Kernel densities of combined math and verbal SAT scores



Note: Fall students and FPF participants in College of Letters and Science only. Source: UC Berkeley.

not completed a bachelor's degree, or have an educational attainment equivalent to a high school diploma or less. Across demographic variables, FPF students are more likely to identify as White or Asian American, and less likely to identify as African American or Latino.

4.3 Predicted outcomes

To further compare the similarities and differences between the treatment group of FPF participants and the control group of fall-semester matriculants, I predict graduation and college GPA outcomes of FPF participants by first estimating regression models of these outcomes on fall-semester matriculants, then by using these coefficients to predict these outcomes for FPF participants. Specifically, I estimate the following:

$$Y_i = \alpha + \gamma X_i + \varepsilon_i.$$

Students are indexed by i. Y_i are the college outcomes I seek to predict, such as the probability of graduation in four or six years and the college GPA at graduation. X_i is a vector of control variables from students' admission applications, including the following: high school GPA; SAT math and verbal scores; first-generation college student status (neither parent having a college degree); high school academic performance index (API) decile dummy variables for California public school graduates (including dummy variables for California private school and out-of-state school attendance); and other student demographic variables, such as gender, California residency, and ethnicity. Continuous outcomes, such as time-to degree in years and UC Berkeley GPA for graduates, are estimated and predicted by ordinary-least-squares (OLS) linear regression. For binary outcomes, such as graduation, I estimate and predict with logistic regression; predictions from OLS or probit regression are similar in magnitude.

Table 3 compares predicted outcomes for FPF participants and fall-semester matriculants. Predicted outcomes of FPF participants are very similar to fall-semester matriculants; FPF participants' predicted graduation rates in four and six years from UC Berkeley are 1 percentage point greater than for fall-semester matriculants. Mean predictions of college GPA and graduating with a STEM degree are slightly lower for FPF participants. Similar to the high school GPA and admission test score distributions discussed earlier, the distributions of predicted graduation probabilities at four

Table 3: Predicted student outcomes, using Fall student coefficients

	FPF	Fall	Diff.	p-value
Graduate in 4 years (Logit)	0.72	0.71	0.0095	0.02
Graduate in 6 years (Logit)	0.92	0.91	0.012	0.00
UC Berkeley GPA at graduation (OLS)	3.39	3.39	-0.0064	0.21
Observations	8,763	1,866		

Note: Fall students and FPF participants in College of Letters and Science only. Regression model in parentheses. Predictions based on regressing outcome on admission variables for Fall students, then predicting outcomes for Fall and FPF students. Source: UC Berkeley.

years (Figure A2) and six years (Figure A3) as well as college GPAs at graduation (Figure A4) also show less variance among FPF participants; larger shares of fall-semester matriculants at UC Berkeley have very high or very low predicted graduation probabilities and college GPAs, versus FPF participants.

These slightly higher predicted graduation rates (by one percentage point) for the FPF participants suggest that, as a group, predictions of lower graduation rates from lower high school grades and test scores are offset by other high school and family background variables that predict higher graduation rates. However, the predicted graduation rates for FPF participants are a few percentage points lower than the actual graduation outcomes shown in Table 2, suggesting that FPF may increase students' likelihood of graduation.

5 Empirical models and findings

In this section I present my analyses of treatment effect estimates from FPF participation. I will present a brief description of each estimation model and the resulting estimates for the following outcomes: four-year and six-year graduation rates, as well as college GPA at graduation.

5.1 Regression estimates of treatment effects

I first estimate FPF's effect on college outcomes with basic ordinary least squares (OLS) regression models. These are similar to the predicted outcome models discussed earlier, with the addition of

participation in FPF, T_i, as a binary treatment variable:

$$Y_i = \alpha + \beta T_i + \gamma X_i + \varepsilon_i.$$

Similar to before, i indexes students; Y_i indicates college outcomes affected by FPF participation; X_i is a vector of academic preparation, school background, and demographic control variables (as described in the previous section on predicted outcomes); and ε_i are unobserved factors affecting Y_i.

The coefficient of interest, β, describes the mean change in outcome Y_i among FPF participants, controlling for any covariates X_i that may influence outcome Y_i as well as selection into FPF, T_i. We can interpet β as the causal impact of FPF participation only if the conditional independence assumption holds, that any unobserved variables captured in ε_i are orthogonal to FPF participation. Because T_i is a single indicator variable, these models assume that the effects of other covariates on college outcomes are identical for FPF participants and fall-semester matriculants. All standard errors are clustered at the high-school level when possible.

Table 4 shows estimated effects of FPF's on the probability of graduation in four years, with each column representing the inclusion of additional control variables. The mean four-year graduation rate for fall-semester matriculants at UC Berkeley is 71 percent. The effect of high school grade point averages (HS GPA) on college outcomes are scaled down to one-third point increments, similar to a student moving from a B-plus to an A-minus grade average. SAT math and verbal score effects are scaled up to 100-point increments.

Column (1) only controls for admission-year fixed effects; we see that FPF participants are more likely to graduate in four years by 4.4 percentage points. Column (2) include student academic characteristics, such as the student's high school GPA and SAT scores at admission; the estimated effect of FPF on four-year graduation increases to 5.8 percentage points. Since high school GPAs and admission test scores are positively correlated with graduation, but slightly negatively correlated with spring admission and selection into FPF, the exclusion of high school GPA and SAT scores lowers the estimated coefficient from FPF participation in column (1) compared to the same coefficient in column (2) when academic controls are included. The addition of high school academic performance decile fixed effects and an indicator for first-generation college student status

Table 4: Graduation in 4 years, linear model estimates

	(1)	(2)	(3)	(4)	(5)
FPF (Treatment)	0.044***	0.058***	0.049***	0.046***	0.043***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
High school GPA, weighted		0.076***	0.081***	0.072***	0.073***
(× 1/3rd point)		(0.0051)	(0.0054)	(0.0053)	(0.0053)
Math SAT		0.045***	0.032***	0.036***	0.045***
(× 100 points)		(0.0063)	(0.0068)	(0.0071)	(0.0074)
Verbal SAT		0.046***	0.033***	0.029***	0.024***
(× 100 points)		(0.0060)	(0.0062)	(0.0062)	(0.0063)
Academic background	No	No	Yes	Yes	Yes
Demographic background	No	No	No	Yes	Yes
Intended major FEs	No	No	No	No	Yes
R-squared	0.0016	0.082	0.090	0.109	0.113

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors (in parentheses) clustered at high-school level. Treatment variable is FPF participation and spring admission to UC Berkeley. High school GPA and SAT coefficients scaled as indicated. All models include year FEs. Outcome $\overline{Y} = 0.712$ for Fall students. $N = 10629$.

in column (3) reduces the magnitude of FPF participation on predicted four-year graduation to 4.9 percentage points; this falls to 4.6 percentage points when indicator variables for ethnic and gender self-identification, as well as California residency, are included in column (4). Lastly, column (5) includes controls for students' intended fields of study, bringing the FPF coefficient down to 4.3 percentage points. With the inclusion of all control variables, the predicted increase from attending FPF on the probability of graduation in four years is larger than a 100-point increase in a student's verbal SAT score, about the same as a 100-point increase in a students math SAT score, and two-thirds of the effect from increasing a student's high school GPA by a third of a grade point.

Even with the inclusion of control variables, the estimated effects of FPF on predicted four-year graduation rates across models are similar in magnitude to the unadjusted mean difference in graduation rates between FPF participants and Fall enrollees shown earlier in Table 2. Across models, the estimated coefficients for FPF participation are statistically significant ($p < .001$) and the standard errors for the FPF estimates are similar in magnitude. Since the mean graduation rate in four years for regular Fall students is 71 percent, the FPF cohort has a 6 percent greater share of students graduating in four years during the 2006–2009 study period.

Table 5: Graduation in 6 years, linear model estimates

	(1)	(2)	(3)	(4)	(5)
FPF (Treatment)	0.034***	0.040***	0.035***	0.033***	0.033***
	(0.0066)	(0.0065)	(0.0065)	(0.0066)	(0.0066)
High school GPA, weighted		0.040***	0.043***	0.039***	0.039***
(× 1/3rd point)		(0.0044)	(0.0044)	(0.0043)	(0.0043)
Math SAT		0.018***	0.015**	0.014**	0.014**
(× 100 points)		(0.0043)	(0.0047)	(0.0050)	(0.0052)
Verbal SAT		0.011**	0.0065	0.0038	0.0035
(× 100 points)		(0.0038)	(0.0041)	(0.0041)	(0.0041)
Academic background	No	No	Yes	Yes	Yes
Demographic background	No	No	No	Yes	Yes
Intended major FEs	No	No	No	No	Yes
R-squared	0.0026	0.043	0.052	0.067	0.068

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors (in parentheses) clustered at high-school level. Treatment variable is FPF participation and spring admission to UC Berkeley. High school GPA and SAT coefficients scaled as indicated. All models include year FEs. Outcome $\overline{Y} = 0.911$ for Fall students. $N = 10629$.

Table 5 displays estimates of FPF effect on six-year graduation rates; FPF participation is associated with a 3 to 4 percentage-point increase in this outcome. Each column has the same regression models as in Table 4. Again, these estimates are similar to the mean FPF-Fall student difference in the six-year graduation rate. 91 percent of fall-semester matriculants graduate from Berkeley in six years, so the increase in the graduation rate from attending FPF is small but still statistically significant (with $p < .001$ across models). The estimated increase to the probability of six-year graduation from attending FPF is similar in scale to a student having a one-third point increase in their high school GPA, and larger than an 100-point increase in a student's math or verbal SAT score.

However, the cumulative UC Berkeley GPA at graduation for FPF participants does not appear to be larger than for fall-semester matriculants admitted to the College of Letters and Science (Table 6). In the basic model in column (1), FPF students graduate with a lower GPA by 0.04 grade points. Controlling for students' high school GPAs and SAT scores, as well as other controls, reduces the GPA gap at graduation to 0.03 points. This estimated decrease in GPA from FPF participation is statistically significant ($p < .01$), but small in magnitude, compared to the mean GPA at graduation

Table 6: UCB GPA at graduation, linear model estimates

	(1)	(2)	(3)	(4)	(5)
FPF (Treatment)	-0.044***	-0.016	-0.025**	-0.023**	-0.028**
	(0.0095)	(0.0087)	(0.0087)	(0.0085)	(0.0085)
High school GPA, weighted		0.074***	0.082***	0.083***	0.087***
(× 1/3rd point)		(0.0049)	(0.0053)	(0.0051)	(0.0051)
Math SAT		0.022***	0.0010	0.017**	0.032***
(× 100 points)		(0.0058)	(0.0063)	(0.0065)	(0.0065)
Verbal SAT		0.13***	0.11***	0.10***	0.092***
(× 100 points)		(0.0049)	(0.0052)	(0.0049)	(0.0050)
Academic background	No	No	Yes	Yes	Yes
Demographic background	No	No	No	Yes	Yes
Intended major FEs	No	No	No	No	Yes
R-squared	0.0022	0.20	0.22	0.24	0.26

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors (in parentheses) clustered at high-school level. Treatment variable is FPF participation and spring admission to UC Berkeley. High school GPA and SAT coefficients scaled as indicated. All models include year FEs. Outcome $\overline{Y} = 3.406$ for Fall students. $N = 9901$.

of 3.41 among fall-semester matriculants. By comparison, a one-third grade point increase in high school GPA is associated with a predicted 0.09 increase in college GPA at graduation, while a 100-point verbal SAT increase is associated with a 0.09 increase in college GPA. The change in college GPA associated with a Math SAT increase of 100 points is considerably smaller, only 0.03 points.

The following subsections will explore the addition of high school fixed effects to these linear models, heterogeneous impacts of FPF among particular groups, and adjusting estimated effects of FPF due to selection bias of unobservables. Robustness checks of these main estimates using alternative regression and matching models are included in Appendix A; these alternative models result in point estimates and standard errors that are similar in direction and magnitude to the basic models described above.

5.2 Fixed-effects estimates

In the previous regressions, I included dummy variables for a California high school's Academic Performance Index decile, as well as indicators if a student attended a California private or out-

Table 7: High-school fixed-effect estimates

	w/o fixed effects	w/ fixed effects
Graduate in 4 years	0.036**	0.035**
[$\overline{Y}$ = 0.74]	(0.012)	(0.012)
Graduate in 6 years	0.022**	0.022**
[$\overline{Y}$ = 0.93]	(0.0069)	(0.0072)
UCB GPA	-0.027**	-0.029**
[$\overline{Y}$ = 3.42]	(0.0089)	(0.0094)

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Each cell is a separate regression. Standard errors (in parentheses) clustered at high-school level. All models include year, academic, demographic, and admission controls; HS API is absorbed into HS FEs in column 2. Outcome means shown for Fall students. N = 8182 for graduation estimates; N = 7734 for GPA estimates.

of-state high school. An alternative method to control for students' academic preparation for college, as mediated through unobserved high school quality, would be to include high-school-level fixed effects in lieu of API dummy variables. Among students from the same high school, fixed effects control for unobserved high-school-specific variation that may affect students' future college outcomes.

However, in order to use fixed effects, I only include high schools that have sent students both as regular Fall enrollees and as FPF participants to UC Berkeley. After dropping high schools that only have enrolled Berkeley students in one category, the sample is reduced to 8,182 students (from 10,629 in the full sample). I estimate the regression model with full controls (column (5) in tables 4–6) on this subsample, and I compare these estimates to a regression with over 600 high-school fixed effects.[18] These students are more likely to have attended high school in California, and are more likely than in the full sample to be from schools in the highest API deciles.

I present these fixed-effects results in Table 7, comparing them to the non-fixed-effects regression with high school API controls, as well as the outcome means for this smaller sample. Compared to UC Berkeley students from the same high school, FPF participants are more likely to graduate in four years by 3.5 percentage points, more likely to graduate in six years by 2.2 percentage points, and have a GPA that is lower by almost 0.03 points. If we compare these estimates to the full sample, we see that the graduation rate and college GPA outcome means of Fall students are slightly higher

[18]The fixed-effects models do not include the API decile dummy variables, because they are highly collinear with high school fixed effects, except for some schools that may move from one decile to another from year to year.

than for the full sample, and that the estimated effects of FPF participation are smaller in magnitude than the full sample, but the treatment effects are in the same direction and still significant.

The inclusion of high-school fixed effects has little effect on the estimated FPF treatment effect compared to models without fixed effects. It appears that any treatment effect from FPF participation has little influence from students' high school of origin.[19]

5.3 Heterogeneous effect estimates

Education interventions often have heterogeneous treatment effects, with greater returns to educational attainment and other outcomes from program participation among groups such as underrepresented minority students, lower-income students, and students with below-average high school GPAs and test scores (Card 2001). In this subsection, I provide evidence that FPF participation has heterogeneous effects for students in particular academic and demographic categories. I first compare regression estimates of students having below-median high school GPAs and SAT scores, followed by regressions of FPF's effects among three categories of underrepresented students at UC Berkeley.

I first look at FPF's impact on students with weaker academic preparation for college, as expressed by their high school grades and SAT scores. I add indicator variables for students with below-median high school GPAs and combined math and verbal SAT scores, interacted with FPF participation.[20] In Table 8, we see increases in four-year graduation rates are concentrated among FPF participants with a below-median SAT score, with a smaller increase for low-GPA FPF students. The sign of the FPF coefficent switches from positive to negative, and is no longer statistically significant ($p > .05$). Combining coefficients, we see that students with a low SAT score who participate in FPF are 5.4 percentage points more likely to graduate in four years. Since the mean four-year graduation rates of low-SAT and low-GPA Fall Berkeley students are 8–9 percentage points lower than the overall mean, FPF participation appears to cut this gap by a half or more for students with below-median SAT scores.

[19] High-school fixed effects do provide more explanatory power, with an increase in R-squared by about .10 for each outcome. The estimated fixed effects themselves also have a wide range of positive and negative values, so school origin does matter for college outcomes, but separately from FPF participation.

[20] I use all students in the College of Letters and Science admitted in the Fall term or participating in FPF to determine the median GPA and SAT score; FPF students are more likely to have below-median high school GPA and SAT scores.

Table 8: Heterogeneous effects of FPF, students with below-median GPA and SAT scores

	Graduate in 4 years		Graduate in 6 years		UCB GPA	
FPF	0.043***	-0.036	0.033***	-0.022	-0.028**	-0.069***
	(0.011)	(0.021)	(0.0066)	(0.012)	(0.0085)	(0.017)
Low HS GPA		0.041		0.052***		0.013
× FPF		(0.022)		(0.013)		(0.017)
Low SAT score		0.090***		0.034**		0.056**
× FPF		(0.022)		(0.012)		(0.017)
Fall mean (all)	0.71		0.91		3.41	
Fall mean (low GPA)	0.63		0.87		3.30	
Fall mean (low SAT)	0.62		0.88		3.27	

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Each column is a separate regression. Standard errors (in parentheses) clustered at high-school level. All models include year, academic, demographic, and admission controls. Outcome means shown for Fall students. $N = 10629$ for graduation estimates; $N = 9901$ for GPA estimates.

Similarly, the relationship between FPF participation and graduation in six years is concentrated among students with a low GPA in high school and smaller effects for students with a low SAT score. Both are strongly significant ($p < .01$). Again, FPF students with higher grades and SAT scores do not appear to have higher graduation rates compared to regular Fall students. Lastly, the negative relationship between FPF and UC Berkeley GPA for graduates is concentrated among FPF participants with stronger academic backgrounds. FPF students with low SAT scores do not graduate with lower college GPAs compared to other Fall students.

FPF's impact among historically underrepresented groups of students at UC Berkeley is less clear. I include indicator variables for underrepresented minority (URM) students, first-generation college students, and students from California high schools with below-median API rankings, all interacted with FPF participation. These results are shown in Table 9. Students from these groups who attend FPF appear to have higher four- and six-year graduation rates versus regular Fall students, but the estimates are marginally significant for underrepresented minority students, and not statistically significant for first-generation college students and students from low-API high schools. FPF's effect on college GPA appear to be unchanged with the inclusion of these indicator variables.

Table 9: Heterogeneous effects of FPF, students from historically underrepresented groups

	Graduate in 4 years		Graduate in 6 years		UCB GPA	
FPF	0.043***	0.026*	0.033***	0.020**	-0.028**	-0.030**
	(0.011)	(0.013)	(0.0066)	(0.0074)	(0.0085)	(0.0098)
URM student × FPF		0.065		0.042		0.035
		(0.039)		(0.023)		(0.030)
1st-gen student × FPF		0.020		0.015		-0.00071
		(0.026)		(0.015)		(0.021)
Low-API HS × FPF		0.028		0.023		-0.011
		(0.032)		(0.021)		(0.025)
Fall mean (all)	0.71		0.91		3.41	
Fall mean (URM)	0.52		0.83		3.21	
Fall mean (1-G)	0.61		0.87		3.25	
Fall mean (lowAPI)	0.59		0.87		3.25	

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Each column is a separate regression. Standard errors (in parentheses) clustered at high-school level. All models include year, academic, demographic, and admission controls. Outcome means shown for Fall students. $N = 10629$ for graduation estimates; $N = 9901$ for GPA estimates.

5.4 Coefficient stability of treatment effects

As shown before, OLS estimates of the treatment effect of FPF are stable across specifications as we include control variables that affect both treatment selection and outcomes. While these explanatory controls do little to adjust FPF's effects, this does not rule out the possibility that unobservable variables may bias the treatment effect. Altonji et al. (2005) and Oster (2017) discuss methods to estimate a bias-adjusted treatment effect, or determine the relative degree of selection on unobservables (versus selection on observables) that would negate any treatment effects, by setting assumptions on the degree of selection by unobservable variables and well as on explaining the outcomes of interest In this section, I test the sensitivity of my results to possible bias from unobservable variables, to see if FPF's treatment effects may still be significant.

Oster (2017) derives the following approximation of β^*, the bias-adjusted treatment effect:

$$\beta^* = \beta_1 - \delta \cdot (\beta_0 - \beta_1) \cdot \frac{(R_{max} - R_1)}{(R_1 - R_0)}$$

δ is a proportion of the degree of selection on unobservables relative to the degree of selection on

observables. If we assume $\delta = 1$, we allow for equal selection from unobservables and observables. β_0 and R_0 are the estimated coefficient and R-squared values from a regression of the outcome variable on FPF participation, with no control variables. β_1 and R_1 are the coefficent and R-squared from a regression of the same outcome on FPF and all controls. R_{max} is the possible maximum R-squared from a regression of the outcome variable on the treatment variable and all observed and unobserved variables. R_{max} is assumed to equal 1 by Altonji et al. (2005), but Oster (2017) argues that it is reasonable to assume R_{max} is some value less than one due to measurement error or idiosyncratic variation in the outcome variable. By assuming specific values of R_{max} and δ, we can then calculate β^* from our previous regression results. We could also vary the value of R_{max} from R_1 to 1, or the degree of selection δ, and observe how the bias-adjusted treatment effect changes.

One option for setting R_{max} is setting it equal to 1.3 times the R-squared of the regression with full controls. Oster (2017) provides evidence from randomized studies that most treatment effect estimates from these studies would still be significantly different from zero with this value. This rule also assumes that the observable variables were chosen for their explanatory power on both treatment and outcome variables, and that likely unobservable variables that could be included would have less explanatory power.

In these tests, I set $\delta = 1$, meaning that the proportional degree of selection on unobservables into treatment is equal to the degree of selection on observables. With these assumptions on R_{max} and δ, we see in Table 10 that the bias-adjusted estimates for FPF on all outcomes are slightly smaller in magnitude and closer to zero, but still very similar to the OLS regressions with full controls. For example, β^* for graduating in 4 years is 0.0427, compared to 0.0430 for FPF with all controls. The β^* for graduating in 6 years is 0.0322, slightly smaller than the estimate with all controls, 0.0326. Lastly, the estimated β^* for college GPA is less negative, at -0.0227 versus -0.0278.

I also estimate the degree of selection of unobservables, the value of δ, that would result in β^* equal to zero. Assuming $R_{max} = 1.3R_1$, we see that the degree of selection on unobservables would have to be quite large for the bias-adjusted treatment effect of FPF to be zero: δ is approximately 60 for graduating in 4 years, about 45 for graduating in 6 years, and greater than 5 for UCB GPA at graduation.

An alternative assumption for R_{max} is that the inclusion of unobservable variables will have the

Table 10: Coefficient stability of FPF treatment effects

	Grad 4y	Grad 6y	UCB GPA
β_0 (no controls)	0.0440	0.0337	-0.0439
R_0 (R-squared)	(0.002)	(0.003)	(0.002)
β_1 (all controls)	0.0430	0.0326	-0.0278
R_1 (R-squared)	(0.113)	(0.068)	(0.260)
β^*(for $\delta = 1$)			
$R_{max} = 1.3R_1$	0.0427	0.0322	-0.0227
$R_{max} = 2R_1 - R_0$	0.0419	0.0313	-0.0106
$R_{max} = 1$	0.0322	0.0035	0.0253
δ(for $\beta^* = 0$)			
$R_{max} = 1.3R_1$	59.8	44.9	5.2
$R_{max} = 2R_1 - R_0$	18.6	14.8	1.5
$R_{max} = 1$	2.4	1.0	0.5
Observations	10,629	10,629	9,901

Note: β_0 and R_0 are the coefficient and R-squared values from a regression of the outcome variable on FPF participation only. β_1 and R_1 are the FPF coefficent and R-squared from a regression of the same outcome on FPF participation and all controls. β^* is the bias-adjusted treatment effect. R_{max} is the feasible R-squared from regressing the outcome variable on the FPF participation and all observed and unobserved variables. δ is the ratio of the degree of selection on unobservables to the degree of selection on observables. Source: UC Berkeley.

same amount of explanatory power as the addition of the full set of observable variables. If R_0 is the R-squared from the short regression of the outcome variable on the treatment variable, and R_1 is the R-squared from regressing the outcome variable on the treatment and all observable controls, then $R_{max} = R_1 + (R_1 - R_0) = 2R_1 - R_0$, where $R_1 - R_0$ is the added increase in R-squared from the inclusion of unobservables. Generally, we would have a larger maximum R-squared and larger bias calculation from unobservable variables than when $R_{max} = 1.3R_1$. We see that the bias-adjusted estimates β^* for each outcome are smaller in magnitude and closer to zero, but still in the same direction and similar in value to the estimated coefficients with all controls. The degree of selection of unobservables δ when $R_{max} = 2R_1 - R_0$ for β^* to be equal to zero is still large for the graduation outcomes (18.6 for graduation in 4 years, 14.8 for graduation in 6 years), but only 1.5 for college GPA.

5.5 Discussion

[TO EDIT FURTHER]

Based on the results above, FPF participants appear to graduate from UC Berkeley at slightly higher rates, but with slightly lower GPAs than their regular fall-term counterparts. We also see that the benefits of FPF are concentrated among students with below-median high school grades and SAT scores. One interpretation for FPF participants having higher graduation rates but lower GPAs at graduation, compared to regular Fall students, is that FPF helps retain academically-weaker students at Berkeley who would have struggled as regular Fall or as Spring non-FPF students. These FPF students' classroom performance may slightly lag their peers at Berkeley, but they are still more likely to persist until graduation.

[many possible stories; stick to a few here]

One possible mechanism is that FPF students gain institutional knowledge that helps them in navigating the university in the following semesters, such as in major preparation and course selection. This relationship FPF students have with their program advisors and instructors is not restricted to the time they have in the program, but may benefit them in later years. Also, FPF participants may view their FPF advisor as their primary point of contact to other advising services at UC Berkeley, simplifying the advising process compared to traditional Fall students. This mechanism can help explain why FPF students with lower high school GPAs or SAT test scores have larger increases in their graduation rates in four and six years; FPF may improve aspects of college completion outside of classroom performance.

Another mechanism is a possible cohort/peer effect of FPF by taking classes and socializing with other peers with the same official status as Spring admits to Berkeley. Prior work on spring admission to college discusses a negative shock to self-confidence by receiving a spring admission letter—a prospective student may wonder "why can't I attend in the fall? Am I good enough?" Instead of arriving in - cohort effect of coming to a new environment with others in a similar situation,

non-cognitive skills: psychology and economics literature looking at self-efficacy (Benabou and Tirole) and/or identity (Akerlof and Kranton)

- possible negative shock to self-confidence by receiving spring admit letter, even with admit to

Berkeley; not "good enough" for regular Fall admission; other work with Spring admission suggests this.

- FPF may improve students' self-confidence or identification with UC Berkeley, improving their experience as Berkeley college students even if their classroom performance is unaffected.

Could FPF's structure and programming have no true effect on outcomes, but that FPF participants have a "chip on their shoulder" by receiving the spring admission offer, motivating them to work harder than they would as a regular Fall student? Perhaps, but the impacts I analyze take place four years or more after school entry. This could be a possible reason for first-year probation and other early outcomes being better for FPF students, but less relevant for later outcomes. Surveying FPF participants and non-FPF Spring admits on how they reacted to Spring admission and their subsequent behaviors at Berkeley can help shed light on what mechanisms are at work.

-Why would the effects for non-FPF Spring admitted students be uniformly worse than for FPF participants then?

- bias-adjusted estimates to account for unobservables before college enrollment suggest the effects of FPF would be similar in magnitude;

- of course there are other factors after year one that would affect college outcomes, but we do not want to "control" for these in determining FPF's treatment effects.

6 Conclusion

In this paper I evaluated the effect of UC Berkeley Extension's Fall Program for Freshmen on graduation rates and other outcomes of first-year enrollees at UC Berkeley who were offered admission during the spring semester. I show that FPF participants are similar to fall-semester matriculants on many observable characteristics, but have slightly lower high school grade point averages and SAT scores. Estimated effects from a variety of models suggest that FPF participants have a small and statistically significant increase in their likelihood to graduate in four years and in six years, compared to fall-semester matriculants. When all control variables are taken into account, the probability of graduation within four years increased by 3.4 to 4.6 percentage points, and the probability of graduation within six years increased by 2.4 to 3.3 percentage points. However, college GPAs at graduation were 0.02 to 0.04 grade points lower for FPF participants. These

estimated treatment effects are robust across a variety of models. FPF participation had larger effects on graduation rates of students entering Berkeley with below-median high school GPAs and SAT scores, with suggestive evidence of effects among for underrepresented minority students, students with lowest-quartile high school GPAs, and students with other academic and demographic characteristics that generally predict a lower likelihood of graduation.

One extension for this project is analyzing the potential response of fall-semester versus spring-semester admitted students to a deferred admission offer. It is plausible that spring-semester matriculants, both FPF participants and non-FPF students, would prefer to attend UC Berkeley as fall-semester matriculants, if given the opportunity. However, would fall-semester matriculants still attend Berkeley if they were offered spring-semester admission? If so, would they choose to participate in FPF, or pursue other activities during the fall semester? I have shown that FPF participants are similar or slightly weaker academically compared to fall-semester matriculants, and that FPF participants are stronger than the spring-admitted students who do not participate in FPF; would this difference alone explain FPF's estimated effects on graduation rates? With additional admissions data that includes students who were admitted and did not attend Berkeley, we could be able to determine the relative strength of fall- and spring-semester admits who choose to attend Berkeley, including the FPF participants. This data would also help control for selection for attending Berkeley for both fall- and spring-admitted students, and assessing further how similar fall-semester matriculants are to FPF participants.

Another path for further research would be to test the different mechanisms that may be operating through spring admission and FPF participation. Is it the smaller menu of courses, the smaller class sizes, or the program advising that has the largest impact? Which of these components has the greatest impact on marginal students? Or perhaps all of these elements work together—not in isolation, such as in the City University of New York's Accelerated Study in Associate Programs (ASAP) model for academic courses and services. If these program elements are adaptable and scalable, they may be implemented at other universities, both selective and broad-access institutions, to improve graduation rates and other academic outcomes for first-year students.

References

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 113(1):151–184.

Bailey, M. J. and Dynarski, S. M. (2011). Gains and Gaps: Changing Inequality in U.S. College Entry and Completion. Working Paper 17633, National Bureau of Economic Research.

Bailey, T. R., Jaggars, S. S., and Jenkins, D. (2015). *Redesigning America's Community Colleges: A Clearer Path to Student Success*. Cambridge, MA: Harvard University Press.

Bettinger, E. P., Boatman, A., and Long, B. T. (2013). Student Supports: Developmental Education and Other Academic Programs. *The Future of Children*, 23(1):93–115.

Bettinger, E. P. and Long, B. T. (2010). Does Cheaper Mean Better? The Impact of Using Adjunct Instructors on Student Outcomes. *The Review of Economics and Statistics*, 92(3):598–613.

Bound, J., Lovenheim, M. F., and Turner, S. (2012). Increasing time to baccalaureate degree in the united states. *Education Finance and Policy*, 7(4):375–424.

Bowen, W. G., Chingos, M. M., and McPherson, M. S. (2009). *Crossing the Finish Line: Completing College at America's Public Universities*. Princeton, NJ: Princeton University Press.

Card, D. (2001). Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. *Econometrica*, 69(5):1127–1160.

Castleman, B. L. and Long, B. T. (2016). Looking Beyond Enrollment: The Causal Effect of Need-Based Grants on College Access, Persistence, and Graduation. *Journal of Labor Economics*, 34(4):1023–1073.

Crisp, G., Doran, E., and Salis Reyes, N. A. (2017). Predicting Graduation Rates at 4-year Broad Access Institutions Using a Bayesian Modeling Approach. *Research in Higher Education*, pages 1–23.

Figlio, D. N., Schapiro, M. O., and Soter, K. B. (2015). Are Tenure Track Professors Better Teachers? *The Review of Economics and Statistics*, 97(4):715–724.

Goldrick-Rab, S. (2016). *Paying the Price: College Costs, Financial Aid, and the Betrayal of the American Dream*. University of Chicago Press.

Hirano, K., Imbens, G. W., and Ridder, G. (2003). Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score. *Econometrica*, 71(4):1161–1189.

Imbens, G. W. (2004). Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review. *The Review of Economics and Statistics*, 86(1):4–29.

Jenkins, D. and Rodríguez, O. (2013). Access and Success with Less: Improving Productivity in Broad-Access Postsecondary Institutions. *The Future of Children*, 23(1):187–209.

Kamenetz, A. (2016). How One University Used Big Data To Boost Graduation Rates. NPR.org, October 30. Accessed July 15, 2017. http://www.npr.org/sections/ed/2016/10/30/499200614/how-one-university-used-big-data-to-boost-graduation-rates.

Lavecchia, A. M., Liu, H., and Oreopoulos, P. (2016). Behavioral Economics of Education: Progress and Possibilities. In Hanushek, E. A., Machin, S., and Woessmann, L., editors, *Handbook of the Economics of Education*, volume 5, chapter 1, pages 1–74. Elsevier.

Oster, E. (2017). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 0(0):1–18.

Pike, G. R., Hansen, M. J., and Lin, C.-H. (2011). Using Instrumental Variables to Account for Selection Effects in Research on First-Year Programs. *Research in Higher Education*, 52(2):194–214.

Quinton, S. (2013). Georgia State Improved Its Graduation Rate by 22 Points in 10 Years. *The Atlantic*. September 23. Accesssed July 15, 2017. https://www.theatlantic.com/education/archive/2013/09/georgia-state-improved-its-graduation-rate-by-22-points-in-10-years/279909/.

Reynolds, C. L. and DesJardins, S. L. (2009). The Use of Matching Methods in Higher Education Research: Answering Whether Attendance at a 2-Year Institution Results in Differences in Educational Attainment. In Smart, J. C., editor, *Higher Education: Handbook of Theory and Research*, pages 47–97. Springer.

Scott-Clayton, J. (2012). On Money and Motivation: A Quasi-Experimental Analysis of Financial Incentives for College Achievement. *Journal of Human Resources*, 46(3):614–646.

Slaughter, A.-M. (2017). The Broken Promise of Higher Education. *The Atlantic*. May 17. Accesssed July 15, 2017. https://www.theatlantic.com/education/archive/2017/05/the-broken-promise-of-higher-education/526995/.

Titus, M. A. (2007). Detecting Selection Bias, Using Propensity Score Matching, and Estimating Treatment Effects: An Application to the Private Returns to a Master's Degree. *Research in Higher Education*, 48(4):487–521.

Wooldridge, J. M. (2007). Inverse Probability Weighted Estimation for General Missing Data Problems. *Journal of Econometrics*, 141(2):1281–1301.

Appendix

A Alternative regression specifications

Here I provide alternative estimates as robustness checks of my main OLS results. The following estimates include: logit and probit models for binary outcomes; separate regressions for treatment and control groups; using propensity scores for inverse-probability weighting; and combining regression with inverse-probability weighting. These methods all produce estimates that are in the same direction as the basic OLS estimates, with similar or slightly smaller effect sizes.

A.1 Logit and probit model estimates

In the above linear regression models, we have binary outcomes for four-year and six-year graduation. One issue in linear probability models is that the predicted probability of the outcome can be either less than zero or greater than one for some observations. Another issue with a dummy variable specification for treatment in a linear regression is that the predicted marginal effect of FPF participation will not vary across students. Instead, I can estimate the treatment effect of FPF participation for a binary outcome Y_i using a latent-variable model,

$$Y_i^* = \alpha + \beta T_i + \gamma X_i + \varepsilon_i, \quad Y_i = \mathbf{1}\left[Y_i^* > 0\right]$$

and fit the model using logit or probit regression.

I compare linear OLS estimates of the mean effect of FPF participation to the equivalent average marginal effects (AME) calculated from logit and probit models. This is done by first calculating for each student the discrete change in the probability of graduation if they participated in FPF, and then averaging the sum of these changes. This is because the marginal effect of FPF participation will vary for each student in a logit or probit model—students with very low or very high probabilities of graduation do not gain a large effect from FPF participation, compared to students with estimated graduation probabilities in the middle of the distribution.

Table A1 compares these AMEs from logit and probit models predicting four-year graduation to the linear probability model coefficients shown before; all models use the same academic record, background, and demographic controls shown in Column (4) in Table 4. The direction and magnitude of the estimated effects of FPF participation are similar across models, as well as for other covariates. In general, the AMEs estimated from logit and probit models of FPF participation are slightly smaller in magnitude than the equivalent OLS estimate: a 3.6 percentage point increase (in both logit and probit models) in the likelihood of graduating in four years, versus 4.6 percentage

Table A1: Graduation in 4 years, average marginal effects comparison

	(1) OLS	(2) Logit (AME)	(3) Probit (AME)
FPF (Treatment)	0.046***	0.036**	0.036***
	(0.011)	(0.011)	(0.011)
High school GPA, weighted	0.072***	0.063***	0.064***
(× 1/3rd point)	(0.0050)	(0.0047)	(0.0047)
SAT Math	0.036***	0.033***	0.033***
(× 100 points)	(0.0071)	(0.0067)	(0.0068)
SAT Verbal	0.029***	0.027***	0.026***
(× 100 points)	(0.0061)	(0.0058)	(0.0058)
Mean outcome, Fall students	0.71	0.71	0.71
(Pseudo) R-squared	0.11	0.090	0.090
Log-likelihood	-5962.2	-5739.5	-5742.0
Observations (N)	10629	10629	10629

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors in parentheses. Treatment variable is FPF participation and spring admission to UC Berkeley. All models include academic and demographic admission variables. High school GPA and SAT coefficients scaled as indicated.

Table A2: Graduation in 6 years, average marginal effects comparison

	(1) OLS	(2) Logit (AME)	(3) Probit (AME)
FPF (Treatment)	0.033***	0.026**	0.024**
	(0.0063)	(0.0079)	(0.0074)
High school GPA, weighted	0.039***	0.023***	0.025***
(× 1/3rd point)	(0.0042)	(0.0026)	(0.0027)
SAT Math	0.014**	0.011*	0.011**
(× 100 points)	(0.0049)	(0.0043)	(0.0042)
SAT Verbal	0.0038	0.0040	0.0035
(× 100 points)	(0.0039)	(0.0037)	(0.0037)
Mean outcome, Fall students	0.91	0.91	0.91
(Pseudo) R-squared	0.067	0.100	0.099
Log-likelihood	-1056.9	-2749.7	-2751.4
Observations (N)	10629	10629	10629

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors in parentheses. Treatment variable is FPF participation and spring admission to UC Berkeley. All models include academic and demographic admission variables. High school GPA and SAT coefficients scaled as indicated.

points from the equivalent linear OLS model. In Table A2, we also see similar estimates for FPF participation on six-year graduation outcomes across models, with an estimated increase of about 2.5 percentage points from logit and probit models, versus 3.3 percentage points in the linear OLS model.

A.2 Separate regressions for treatment and control groups

In the previous regression models, I estimated the average treatment effect of FPF by including FPF participation as a binary indicator (or dummy) variable, with all other control variables entered linearly. The previous specifications do not allow us to flexibly estimate possible treatment effect heterogeneity of FPF participation for any particular group of students, nor do they adjust for the likelihood of selection into the treatment or control groups. In the following sections, I will repeatedly refer to the results in Table A3 of selected alternative models of average treatment effects for the three main outcomes discussed above: the probability of both four-year and six-year graduation from UC Berkeley, and college GPA at graduation. Overall, the estimated treatment effects across the models in Table A3 are similar in direction and magnitude, and are also similar to the estimated effects of FPF from above.

One alternate regression approach would fully interact the indicator variable for FPF participation with all control variables (and the constant term). By doing so, we allow the effects of FPF participation to vary across participants' academic and demographic characteristics. This is equivalent to estimating separate regression models for the treatment and control groups:

$$\mu_1(X_i) = \mathbb{E}\left[Y_i^1 | X_i\right]$$
$$\mu_0(X_i) = \mathbb{E}\left[Y_i^0 | X_i\right]$$

and using these regressions to predict the potential control-group outcomes for treatment-group members, and the potential treatment-group outcomes for control-group members. The estimated ATE is the difference in expected outcomes between the treatment and control groups, with the expectation taken over the distribution of X_i for the combined sample:[21]

$$ATE = \mathbb{E}\left[\mu_1(X_i) - \mu_0(X_i)\right]$$

Since the effect of FPF participation is interacted with all possible covariates, each individual's marginal effect from FPF participation will be different depending on their background characteristics. These marginal effects are then averaged to determine the ATE.

[21] Estimating the TOT would use the distribution of X_i for the treated FPF participants only.

Table A3: Comparison of alternative FPF average treatment effect estimates

(a) Graduation in 4 years

Estimation method	Outcome model	FPF propensity score model	FPF ATE	Robust SE	p-value
RA	linear	N/A	0.040	0.013	0.002
RA	logit	N/A	0.035	0.013	0.005
RA	probit	N/A	0.037	0.013	0.003
IPW	weighted mean	logit	0.034	0.013	0.008
IPW	weighted mean	probit	0.034	0.013	0.007
IPWRA	logit	logit	0.034	0.013	0.007
IPWRA	probit	probit	0.035	0.013	0.006

(b) Graduation in 6 years

Estimation method	Outcome model	FPF propensity score model	FPF ATE	Robust SE	p-value
RA	linear	N/A	0.032	0.007	0.000
RA	logit	N/A	0.030	0.007	0.000
RA	probit	N/A	0.031	0.006	0.000
IPW	weighted mean	logit	0.030	0.007	0.000
IPW	weighted mean	probit	0.030	0.007	0.000
IPWRA	logit	logit	0.029	0.007	0.000
IPWRA	probit	probit	0.029	0.007	0.000

(c) UC Berkeley GPA at graduation

Estimation method	Outcome model	FPF propensity score model	FPF ATE	Robust SE	p-value
RA	linear	N/A	-0.031	0.009	0.001
IPW	weighted mean	logit	-0.041	0.009	0.000
IPW	weighted mean	probit	-0.039	0.010	0.000
IPWRA	linear	logit	-0.030	0.009	0.001
IPWRA	linear	probit	-0.030	0.009	0.001

Note: Estimation methods include regression adjustment (RA) with separate regressions by FPF treatment level, using the indicated outcome model; inverse-probability weighting (IPW) (see Hirano et al. 2003); and inverse-probability-weighted regression adjustment (IPWRA), a "doubly-robust" estimator combining RA and IPW methods (see Wooldridge 2007). "FPF propensity score model" indicates the model used to predict FPF treatment selection. All outcome and FPF propensity score models include full set of academic and demographic admission variables.

In Table A3, the first column for each sub-table indicates these models with the acronym "RA" (regression adjustment). Since the first two outcomes, four-year and six-year graduation, are binary, I estimate linear, logit, and probit regression models, as indicated in the second column of the table. All control variables are included, similar to the estimates in column (4) of Tables 4, 5, and 6.[22] We see that these RA models allowing for interactions between FPF participation and all control variables result in similar estimates to the regression models presented before. FPF participation is associated with a 3.5 to 4 percentage-point increase in the four-year graduation rate, a 3 to 3.2 percentage-point increase in the six-year graduation rate, and a -0.03 decrease in the UC Berkeley GPA at graduation.

A.3 Inverse-probability weighting with propensity scores

The next empirical model in Table A3 is labeled "IPW" for inverse-probability weighting. The previous estimates were parametric in nature; we estimated regression coefficients for FPF participation as well as for control variables (and interactions of these variables). A less-parametric method using the selection-on-observables assumptions is to first estimate propensity scores for selection into FPF using logit or probit regression. These predicted propensity scores are used as weights to match as closely as possible the covariate distributions of the treatment and control groups. The average treatment effect is the difference in weighted mean outcomes between the two groups (see Hirano et al. 2003; Imbens 2004; Reynolds and DesJardins 2009).

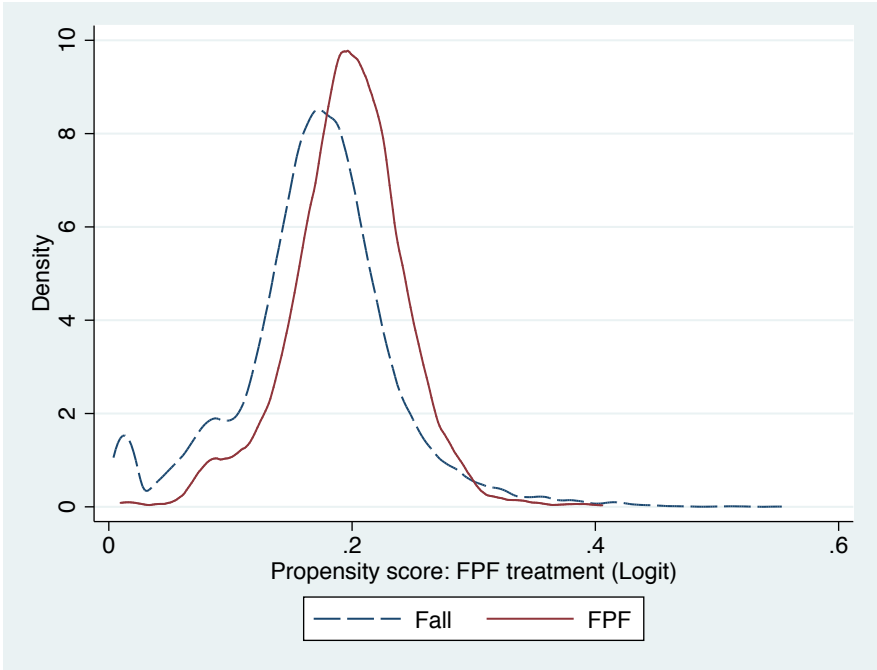
We can first assess the suitability of inverse-probability weighted estimates by first comparing propensity score distributions of FPF selection across the treatment and control groups. Using logistic regression, I first estimate propensity scores predicting selection into FPF, using both fall-semester matriculants and FPF participants in the data:

$$T_i = \alpha + \theta X_i + \varepsilon_i.$$

X_i are the same academic and demographic control variables as mentioned before. T_i is the binary treatment variable, with $T_i = 1$ for FPF participants in the treatment group and $T_i = 0$ for fall-semester matriculants in the control group. From this model I can estimate a predicted propensity score into treatment, $\widehat{T_i}$, for each student. We see in Figure A1 that there is a high degree of overlap of predicted propensity scores between fall-semester matriculants and FPF participants. In contrast to other program evaluation contexts where we are concerned about a lack of overlap between the treatment and control groups, we see that the majority of individuals appear to be similarly likely

[22] The third column listing is "N/A", since propensity score models for FPF participant selection are not used here. Subsequent estimates will use propensity-score weighting to control for differences between the treatment and control groups.

Figure A1: Predicted selection into Fall Program for Freshmen



Note: Fall students and FPF participants in College of Letters and Science only. Source: UC Berkeley.

to be selected into FPF (or fall-semester admittance).

The ATE estimates with inverse-probability weights are then calculated as follows:

$$ATE = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{T_i Y_i}{\widehat{T_i}} - \frac{(1-T_i)Y_i}{(1-\widehat{T_i})}\right]$$

Larger weights are given to observations predicted to be in one group (control or treatment), but are actually members of the other group. Referring back to the IPW estimates in Table A3, we see that the ATE estimates for FPF participation are similar to the previous results: a 3.4 percentage-point increase in the four-year graduation rate, a 3 percentage-point increase in the six-year graduation rate, and a -0.04 decrease in the UC Berkeley GPA at graduation. I include logit and probit estimates for the propensity score treatment model for FPF selection, using all control variables; the ATE estimates are similar with both methods.

A.4 Combining regression with inverse-probability weighting

The last type of estimates shown in Table A3 are listed "IPWRA", for inverse-probability weighted regression adjustment. In these models, we combine the above methods of fully-interacted regression estimates of FPF participation with propensity-score weighting. Estimators combining

regression with propensity-score weighting are considered "doubly-robust" because the estimates are robust to model misspecification in either the outcome model or the treatment model, but not in both (Wooldridge 2007).

In Table A3, I provide estimates that assume either a logit or probit model for selection into FPF treatment. For the binary graduation outcomes at four and six years, both the outcome and treatment models are the same (logit or probit). The outcome and treatment models use the same set of academic and demographic control variables from before. Again, we see that the IPWRA average treatment effect estimates for FPF participation are similar to the other model specifications: a 3.4-3.5 percentage-point increase in the four-year graduation rate, a 2.9 percentage-point increase in the six-year graduation rate, and a -0.03 decrease in college GPA.

B FPF and other Spring enrollees

In this section I compare FPF participants to spring-semester matriculants who did not participate in FPF, which are referred to as "Spring-only" students in this section tables.

Table A4 describes summary statistics of Spring-only students' outcome and background variables, with comparisons to FPF participants. Compared to FPF participants, the Spring-only students have slightly lower high school grades and SAT scores on average, have a higher proportion of Hispanic/Latinx and African American students, and have a larger share of first-generation college students whose parents did not have a bachelor's degree. Spring-only student outcomes are worse than for FPF participants and fall-semester matriculants: their six-year graduation rate was 8 percentage points lower than for FPF participants and 5 percentage points lower than for fall matriculants, and the differences in the five- and four-year graduation rates are considerably larger. In addition, college GPAs at graduation are slightly lower and time-to-degree is longer for Spring-only students.

I also use the prediction model for fall-semester matriculants from section 4.3 to compare predicted outcomes for Spring-only students to FPF participants, as shown in Table A5. The predicted graduation rates and other outcomes for Spring-only students are not as low as their actual outcomes. Based on their admissions data alone, the average predicted outcomes for Spring-only students are almost the same as the predicted outcomes for FPF participants.

Table A4: Summary statistics for FPF participants and Spring-only students

	FPF	Spring-only	Difference	p-value
Outcome variables				
Graduate in 4 years	0.76	0.46	0.29	0.00
Graduate in 6 years	0.94	0.86	0.08	0.00
UC Berkeley GPA at graduation	3.36	3.30	0.06	0.00
Admission variables				
High school GPA, unweighted	3.77	3.76	0.01	0.21
High school GPA, weighted	4.19	4.16	0.03	0.00
SAT Math score	662	643	19.1	0.00
SAT Verbal score	634	616	17.6	0.00
API 1st-5th decile	0.15	0.28	-0.13	0.00
API 6th-9th decile	0.34	0.31	0.03	0.23
API 10th decile	0.29	0.21	0.08	0.00
CA HS, no API	0.15	0.15	-0.00	0.82
HS outside CA	0.07	0.05	0.03	0.02
Parent education				
No 4-yr degree	0.24	0.37	-0.13	0.00
HS diploma or less	0.13	0.21	-0.08	0.00
Female	0.60	0.59	0.01	0.57
CA resident	0.92	0.93	-0.01	0.36
African American/Black	0.02	0.04	-0.02	0.01
Latino/Hispanic	0.09	0.17	-0.08	0.00
Native American	0.00	0.00	-0.00	0.21
Pacific Islander	0.00	0.00	-0.00	0.10
Asian American	0.47	0.44	0.03	0.16
White	0.32	0.26	0.06	0.00
Other	0.09	0.09	0.01	0.49
International	0.00	0.00	-0.00	0.73
Observations	1,866	694		

Note: Sample includes spring-semester matriculants to the College of Letters and Science, UC Berkeley, separated by FPF participation. API stands for California Academic Performance Index for public schools only. Source: UC Berkeley.

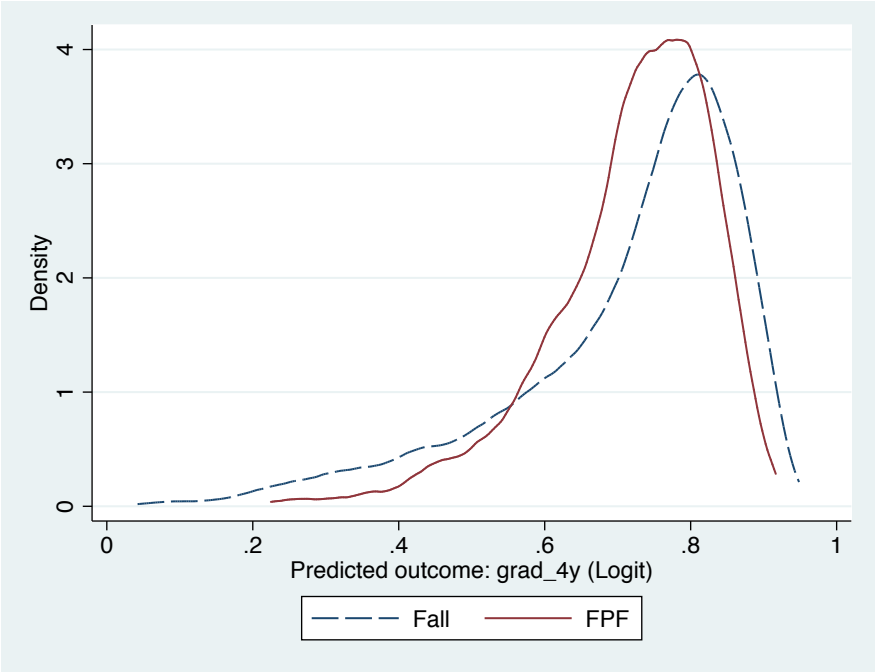
Table A5: Predicted student outcomes

	FPF	Spring-only	Difference	p-value
Graduate in 4 years (Logit)	0.72	0.68	0.04	0.00
Graduate in 6 years (Logit)	0.92	0.91	0.02	0.00
UC Berkeley GPA at graduation (OLS)	3.38	3.33	0.05	0.00
Observations	1,866	694		

Note: Sample includes spring-semester matriculants to the College of Letters and Science, UC Berkeley, separated by FPF participation. Regression model in parentheses. Predictions based on regressing outcome on admission variables for Fall students, then predicting outcomes for FPF and Spring-only students. Source: UC Berkeley.

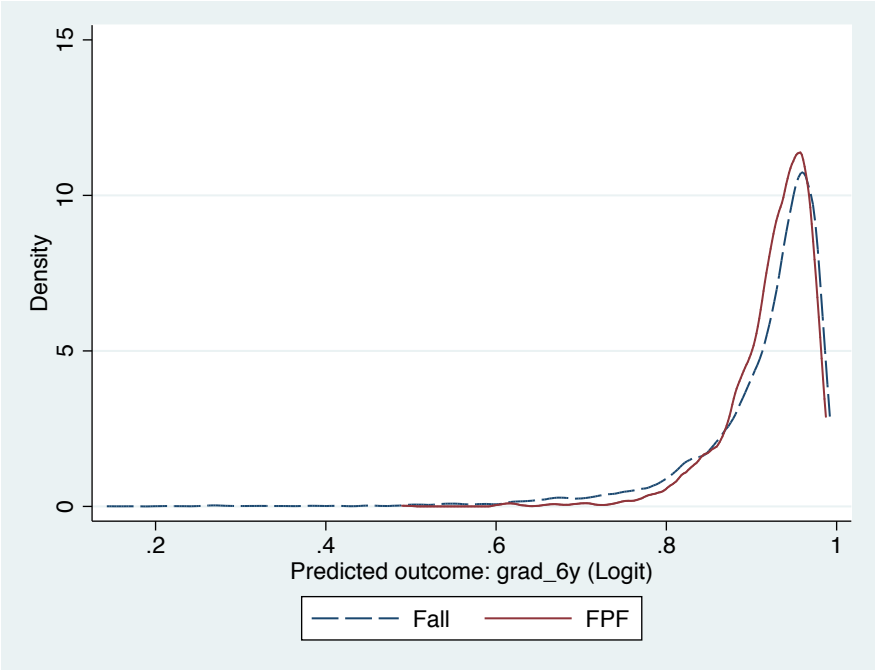
C Additional figures

Figure A2: Prediction for graduation in four years



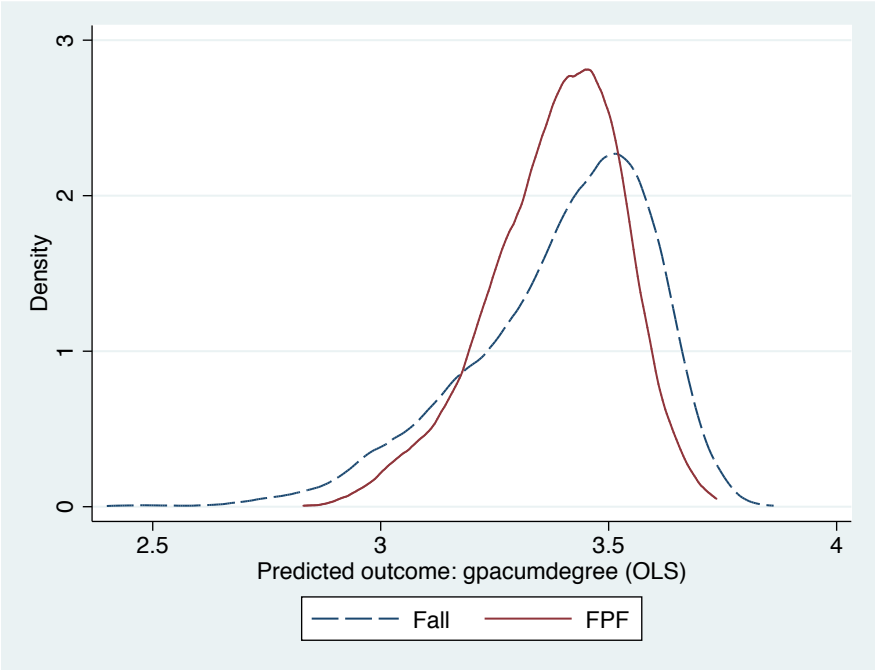
Note: Fall students and FPF participants in College of Letters and Science only. Source: UC Berkeley.

Figure A3: Prediction for graduation in six years



Note: Fall students and FPF participants in College of Letters and Science only. Source: UC Berkeley.

Figure A4: Prediction for UC Berkeley GPA at graduation



Note: Fall students and FPF participants in College of Letters and Science only. Source: UC Berkeley.