

Is Early Start A Better Start?

Evaluating California State University's Early Start Remediation Policy

Michal Kurlaender
University of California, Davis

Lester Lusher
University of Hawaii at Manoa

Matthew Case
California State University Chancellor's Office

December 2018

Abstract:

In 2011, the California State University (CSU), the largest public four-year university system in the nation, enacted a policy reform known as *Early Start*, requiring students identified in need of remediation to enroll in developmental education courses in the summer before their freshmen year. In this paper, we evaluate the CSU's effort to improve college readiness through Early Start and investigate whether this policy reform had an impact on student success (performance and persistence) at CSU. Our results suggest that, although students shifted remedial course taking to the summer before freshman year, Early Start as a policy has not met its intended goals. Overall, students requiring remediation today at CSU have not had better performance or persistence outcomes than similar students that required remediation in the years prior to Early Start, though there is some evidence of more advanced course taking behavior. Moreover, we find no impact of remediation placement on student persistence and achievement outcomes under the Early Start regime.

Note: We thank the CSU Chancellor's Office for their collaboration in this work and for data access. The research reported here was undertaken through the Center for the Analysis of Postsecondary Readiness at Teachers College, Columbia University, and supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C140007. The opinions expressed are those of the authors and do not represent views of the Institute of Education Sciences, the U.S. Department of Education, or the California State University Chancellor's Office.

IS EARLY START A BETTER START?

EVALUATING CALIFORNIA STATE UNIVERSITY'S EARLY START REMEDIATION POLICY

1. Introduction

Today, public colleges and universities are grappling with considerable pressures to improve degree outcomes. Despite increasing college participation, college completion has not kept up, and is particularly low at open-access, or less selective, postsecondary institutions. Nationally, only about two-thirds of students that enter a public baccalaureate-granting institution obtain a degree within six years (NCES, IPEDS 2016). Public postsecondary systems—many facing increasing state accountability pressures for college outcomes—are closely examining policies, practices, or programs that may aid in improving degree completion rates.

One of the main culprits for weak degree completion is college readiness. Many students arrive at college deemed unprepared for college-level work; low levels of college readiness are particularly evident at broad access two- and four-year institutions where nearly 90 percent of all U.S. postsecondary students are enrolled (Sparks & Malkus, 2013). Prior research suggests that nearly one in four first-year undergraduates at broad access institutions report taking remedial courses, with some estimates of participation in remediation closer to 50-60 percent or higher (Scott-Clayton, Crosta, & Belfield, 2014).¹ Weak evidence on the effectiveness of remediation efforts, along with the high costs associated with remediation, have led to important reform efforts across many institutions and state systems of higher education. Remediation reform efforts have addressed all aspects of the college readiness agenda: K-12 alignment efforts

¹ The differences in the rates for remediation need are in part due to measurement; transcript based reports are typically higher than self-reports.

(Kurlaender, Grodsky, Howell, & Jackson, 2016); assessment types and placement decisions (Scott-Clayton et al., 2014; Scott-Clayton & Rodriguez, 2015; Ngo and Melguizo, 2016); and types of developmental course options, including co-requisite and stretch courses (Logue, Watanabe, & Douglas, 2016). Remediation reform efforts demand the close scrutiny of policy researchers in order to inform both higher education leaders and policymakers eager to improve college attainment in an ever-changing labor market.

In 2011, the California State University (CSU), the largest public university system in the nation, enacted one such reform—a policy known as *Early Start*. Facing enormous numbers of students requiring remedial coursework in mathematics and English (at the time of adoption, only about half of all students were deemed exempt from remediation at the point of entry),² the primary goal of Early Start is to have students enter the CSU having met their developmental education needs in the summer before their freshman year.³ This important innovation offers another opportunity to evaluate remediation practices at scale, mainly the shift in the timing of remediation to the summer prior to freshman entry.

The theory of action for the Early Start policy reform was straightforward: students would remediate sooner than they did before (the summer before their first year). This would then result in improved academic preparation to facilitate more credit bearing course enrollment, improved performance in the first year, and ultimately higher persistence, degree attainment and shorter time to degree. Thus, in the short-term, if Early Start were to be effective, policymakers believed it would lead to greater credit accumulation, persistence, and higher performance among students identified for remediation. However, Early Start may not result in desired outcomes because

² http://asd.calstate.edu/performance/proficiency/2010/Prof_Sys_fall2010.htm

³ Memo from CSU Chancellor Charles Reed to CSU Presidents, June 11, 2010. See: <https://www.calstate.edu/eo/EO-1048.pdf>

either students did not comply with the Early Start policy and did not remediate in the summer before their first freshman term, or because the timing of developmental coursework may not matter, or may actually matter in a negative way (for example because it may be delivered poorly in the summer). To the extent that remediation may also have a discouraging effect on students, Early Start may lead some students that ultimately choose not to enroll at CSU to make that choice earlier, or to weaken persistence rates since students may opt to drop out of the CSU more quickly.

In this paper, we evaluate the CSU's effort to improve college outcomes through Early Start. We leverage two empirical strategies to identify the impact of the Early Start reform on student success at CSU (e.g. satisfying remediation, first term performance, unit accumulation, and persistence). First, we investigate the impact of Early Start by comparing student outcomes for those identified for remediation before and after the adoption of the policy against those not identified for remediation using a difference-in-differences approach (i.e. comparing summer enrollment versus fall enrollment in remediation). Second, we use the remediation placement exam to investigate the impact of Early Start by evaluating students identified for remediation under Early Start to otherwise similar students not identified for remediation using a fuzzy regression discontinuity approach (i.e. comparing remediation in the summer to no remediation at all).

Our results suggest that Early Start as a policy has by and large not met its intended outcomes. On the one hand, we identify a strong first-stage effect where students are meeting remediation requirements prior to the fall term as a result of the policy. We also find evidence of more advanced (i.e. upper division) course taking behavior in the first year as a result of Early Start. On the other hand, students requiring remediation today at CSU have not had better

performance or persistence outcomes than similar students that required remediation in the years prior to Early Start. Moreover, from our fuzzy regression discontinuity design, we find that students needing to remediate under Early Start experienced weaker persistence and unit accumulation outcomes when compared to those not requiring remediation (a finding that corroborates with much of the extant literature on college remediation). Thus, although Early Start is successfully remediating students before the fall term, that in and of itself is not enough to improve later outcomes.

The paper is organized as follows. In Section 2, we summarize the literature on postsecondary remediation efforts, focusing on recent policy changes. Section 3 describes the CSU system, and the Early Start policy. In Section 4, we describe our difference-in-differences and regression discontinuity research designs, and Section 5 details the results. Finally, in Section 6 we discuss our findings, looking more closely at the campus level, and conclude by describing CSU's most recent remediation reform to eliminate its primary remediation placement exams, as well as pre-collegiate credits altogether.

2. Postsecondary Remediation: Prior Evidence and Policy Developments

Remediation has long been a costly way to address the misalignment between K-12 and higher education; local, state and federal costs are estimated at approximately \$7 billion annually, and estimates to students are over \$1 billion annually (Scott-Clayton, Crosta, & Belfield, 2012; Center for American Progress (Jimenez, Sargrad, Morales, & Thompson, 2016). Driven by the unfortunate reality that many college students do not have the academic skills necessary to meet the demands of college-level coursework, rates of developmental or remedial course taking have been very high across many of the nation's non-selective postsecondary

institutions (U.S. Department of Education, National Center for Education Statistics, 2011). At the CSU system, for years, nearly two-thirds of all students required some remedial coursework in either math or English when they arrive as freshmen. This resulted in thousands of students required to enroll in non-credit bearing coursework, leaving them nearly a term behind their peers in degree progress.

Despite the investments, the research base on the effectiveness of remedial education programs is inconclusive at best. Part of the difficulty in assessing the impacts of remediation on collegiate outcomes is that students who require remediation are different from those who do not, making it difficult to isolate the effect of remediation on college outcomes from the other things that make these students different (e.g. weaker skills, or less motivation). In research that controls for students' academic skills and other demographic characteristics, students in developmental courses at some colleges do as well as students who never participate in developmental education (Adelman, 1998; Attewell, Lavin, Domina, & Levey, 2006; Shields & O'Dwyer, 2017).

In more recent years, studies have utilized quasi-experimental methods to isolate a causal effect of participating in remedial course-work in college (Scott-Clayton & Rodriguez, 2015; Martorell & McFarlin, 2011; Bettinger & Long, 2009; Calcagno & Long, 2008). The advantage of these studies is that they are able to overcome the main obstacle in evaluating remediation—a viable comparison group. Most of these studies utilize regression discontinuity methods given the strict cutoff on assessments used for remediation placement at many colleges and universities. Of course, the evidence on whether remediation “works” or “does not work” from these studies is restricted to students at the margin of needing it in the first place. Nevertheless, this research yields our best guess about whether remediation policies benefit students in need of

extra skills. Across multiple states' postsecondary systems (in both two-year and four-year institutions), studies comparing students who score just below and just above proficiency on the state-mandated placement test find that students requiring remediation did not have better odds of passing subsequent courses or improved degree performance (Scott-Clayton & Rodriguez, 2015; Martorell & McFarlin, 2011; Calcagno & Long, 2008). Most recently, Boatman and Long (2018) explored remediation placement at two-year and four-year institutions in Tennessee, finding important differences based on students' level of preparation. Importantly, their study extended the prior literature by exploring multiple cutoffs for different placements (i.e. for students at a wider range of developmental needs), and found that students in need of less remediation fare worse when compared to similar students who passed proficiency threshold. However, for students farther below proficiency, remediation actually improved outcomes, particularly persistence through college. These results suggest that remedial and developmental courses may function differently depending on students' level of academic preparedness, and therefore policies that may be beneficial for some students with different levels of academic preparedness may not be for others.

Other studies have utilized variation in policies or practices to causally estimate the effects of remediation placement or policies on student outcomes. For example, Bettinger and Long (2009) explored two-year and four-year colleges in Ohio, taking advantage of the fact that Ohio public institutions have different policies (test cut-offs) for demonstrating proficiency. They found that placement into remediation increased the probability of college persistence when comparing academically-similar peers who were and were not required to take remedial courses. In a study of one large community college system, Ngo and Melguizo (2016) find that altering the mode of the placement assessment (to computer adaptive) resulted in more

placement errors, but that raising the placement cutoffs had no effect. Finally, in a critical study implementing random assignment into a remedial algebra course or a co-requisite college-level statistics course (at three community colleges of the City University of New York (CUNY) system), students enrolled in the co-requisite statistics course were more likely to pass that course when compared to students in remedial algebra, and to have accumulated more college credits (Logue, Watanabe, & Douglas, 2016).

In recent years many states have moved away from remedial courses altogether in their public four-year colleges (Complete College America, 2017), but we are unaware of any similar reform efforts around remediation timing, specifically enrollment in summer school prior to entry, which has been tested. In this paper we extend the existing literature in three important ways. First, we explore remediation in the nation's largest four-year public higher education system—the California State University (CSU) system, which educates an incredibly diverse set of students. Second, we test an important reform to collegiate remediation—requiring summer enrollment to satisfy developmental coursework—adopted at scale across the 23 campuses of CSU system. Finally, we employ multiple strategies to evaluate the causal impacts of this reform on several important outcomes, offering multiple counterfactual comparisons to remediation status.

3. Institutional Setting and Early Start

The California State University system with 23 campuses is the largest public four-year higher education system in the country, educating nearly half a million undergraduate students and serving the top-third of California high school graduates. CSU serves students from a

tremendous range of ethnic and socioeconomic origins. These students come from urban, suburban, and rural areas and attended public high schools that are both among the best and among the worst in the nation. While California may not be a typical state, it reflects the student populations of other states in the U.S., and the mainstream public colleges that educate them, very well. The CSU system campuses are largely broad-access, save for a few campuses that are slightly more selective accepting only about one-third of eligible applicants. As a system, CSU's "peer institutions" include, for example: Cleveland State University, the University of Wisconsin-Milwaukee, Arizona State University, SUNY Albany, University of Maryland-Baltimore County, University of Nevada-Reno, Georgia Southern, University of New Mexico, and Winston-Salem State. Interestingly, all of these institutions have lower SAT or ACT cutoffs for remediation placements.

To meet CSU eligibility requirements, California high school students are required to enroll in a set of pre-approved courses and to take a college entrance exam (ACT or SAT). Approximately 32 percent of high school students complete the set of courses that make them CSU eligible. The four-year degree completion rate at CSU is 18 percent for the most recent cohort, with six-year completion rates at 54 percent. Completion rates have also been increasing steadily in recent years (Figure 1). The four-year completion rates are considerably lower at CSU relative to peer institutions across the nation, yet the CSU's six-year completion rates are comparable to those at peer institutions (Jackson & Cook, 2016).

[Figure 1]

A great number of students who enter the CSU are considered “not ready for college level” work. The CSU uses multiple measures to determine college readiness. First, students may demonstrate college readiness through test scores, both college entrance exam scores (ACT or SAT),⁴ or through meeting a set proficiency standard on California’s 11th grade state assessments. In fact, CSU, in a bold effort to better align to K-12, began partnering with the California Department of Education in 2004 to inform students of their college readiness levels as part of the State’s 11th grade assessments (Howell, Kurlaender, & Grodsky, 2010). College readiness levels based on the 11th grade assessments are incredibly low. Only 13 percent of California 11th grade students meet CSU college readiness standards in mathematics, and 26 percent meet these standards in English Language Arts (ELA). However, an additional 20 percent in mathematics, and 33 percent in ELA, are “conditionally college ready.” This designation represents a second way in which students can demonstrate their college readiness—12th grade coursework. Students that receive a conditionally ready designation through the 11th grade state assessments are able to satisfy their condition through a set of CSU approved courses in the 12th grade in both math and ELA. Finally, students that either do not take and/or take but do not pass these courses in the 12th grade need to take CSU system-wide placement tests, the Entry Level Math (ELM) assessment and the English Placement Test (EPT), respectively, upon enrollment. Students that pass these tests are considered proficient or college ready, while students that do not pass these tests are required to enroll in remedial/developmental coursework. The CSU sets a system-wide threshold for demonstrating college readiness: an ELM score of 50 or higher for identification in math, and a minimum EPT score of 147 for English.⁵

⁴ CSU students are exempt from any additional remedial placement assessments and placement with an English SAT score of 500 and ACT score of 22; for math an SAT score of 550 and an ACT score of 23.

⁵ Prior to 2011 the English Placement Test threshold to exempt from remediation was higher at 151.

In 2011, the CSU enacted a policy known as *Early Start*, requiring incoming students who do not demonstrate readiness for college-level math and/or English to complete remediation during the summer before entering CSU. Executive Order 1048 mandated by the California State University Board of Trustees on May 2010 established as follows:

A program for CSU admitted freshmen who have not demonstrated proficiency in mathematics and/or English as established by CSU faculty. As of summer 2012, incoming freshmen who have not demonstrated proficiency in English and/or mathematics will be required to begin remediation prior to the term for which they have been admitted, e.g., summer prior to fall. All students will be required to have achieved proficiency in English and/or mathematics on or before the end of their first year of enrollment at a CSU campus.⁶

The goals of Early Start are to better prepare students in math and English *before* their first semester at CSU.⁷ Specifically, Early Start requires all incoming students who do not meet the threshold on the Entry Level Math (ELM) and/or English Placement Test (EPT) proficiency requirements to take a designated developmental education course in the summer before their freshman year.⁸ Courses cost the same per unit as regular semester costs for tuition and fees, and students who qualify can also receive financial aid. Early Start courses are largely run by CSU faculty and are meant to expose students to what it's like to attend college. Thus, the mechanisms by which we might expect Early Start to impact student outcomes is twofold: first, to improve preparation, as is the goal of most remedial coursework (e.g. developing academic skills college expectations, time management, etc.), only prior to the freshman year; and second, to facilitate college exposure (including academic coursework, college expectations, time management, etc.)

⁶ <https://www.calstate.edu/eo/EO-1048.pdf>

⁷ See <http://www.csusuccess.org/earlystart/early-start-faq> for additional information.

⁸ Early Start math and English courses are meant to be available at every CSU campus, at a few community colleges, as well as online. Financial aid is available for those who demonstrate need.

thereby potentially expediting the decision about entry (and persistence) versus exit sooner in students' college careers.

We evaluate the impact of this policy for California State University students in two ways. First, in a difference-in-differences framework, we look at this policy relative to the prior policy, which did not require summer enrollment for students identified for developmental coursework. Second, in a fuzzy regression discontinuity framework, we look at the impact of being identified in need of remediation (versus not) for students under the Early Start regime. Together, the two counterfactuals provide strong causal evidence on the effects of the program, and provide suggestive evidence on mechanisms—albeit indirectly.

4. Data and Analysis Plan

We employ data from the California State University Chancellor's Office for the census of CSU freshmen enrolled in 2009 through 2015, three cohorts before the Early Start reform (2009-2011), and four cohorts after (2012-2015).⁹ The Chancellor's Office of the CSU system collects individual level data on all applicants and enrollees, including credits completed, grades by term, and a variety of background information from the application file. We focus on remediation in mathematics given changes in the English remediation cutoff and placement policies for English that occurred in the same years as the Early Start reforms.¹⁰ To measure whether Early Start led to its intended goals, we investigate several outcome measures.

⁹ We start with the 2009 academic year (and not earlier) because test score requirements for mathematics remediation changed after 2008.

¹⁰ CSU changed the SAT threshold for demonstrating college readiness in English in the same year as the Early Start implementation, as well as relaxed Early Start participation for students with different English Placement Test results.

Measures

To measure whether students requiring remediation enrolled in Early Start, we first investigate whether CSU entering freshmen satisfied the remediation requirements prior to fall entry in the Early Start period. Then, to measure short-term impacts of summer remediation on academic performance, we examine first term GPA. To determine if Early Start changed the number and types of courses students enrolled in, we examine both whether a student enrolled in an upper division course and the number of upper division units enrolled in year one, as well as total units accumulated for the student through their time at the CSU. Finally, we examine impacts on persistence rates to year two and year three.

We include several key control variables: race/ethnicity (Underrepresented Minority, which includes Black, Latino/Hispanic or American Indian; White; Asian; Other/Missing), gender, high school GPA, and SAT scores. Table 1 includes the summary statistics of the full population of CSU first-time freshmen during our sample years. At CSU, males make up 43 percent of first-time freshmen, and the student population is considerably diverse with historically underrepresented minority students (African American, Latino or American Indian) at 46 percent of the first-time freshmen student population. CSU students are, on average, B+ high school students, entering with a 3.25 high school GPA, and with average composite SAT scores at 895. Importantly, we note that 56 percent of entering CSU students are required to take the ELM upon entry (i.e. they did not demonstrate college readiness via high school performance or SAT/ACT exam performance that would exempt them from any further remedial placement assessments). Among those that took the ELM, about one-third pass. For our analytical sample, we focus exclusively on ELM test takers (i.e. the 56% of incoming first-time freshmen) since they are those who are potentially subject to additional remediation upon entry into CSU.

[Insert Table 1]

To investigate the impacts of the Early Start policy (i.e. the impacts of requiring summer pre-requisite coursework), we employ two strategies. The first strategy, a difference-in-differences, exploits the plausibly exogenous timing of the passage and implementation of the reform, comparing students who need to enroll in remediation in the summer as a result of Early Start to a counterfactually similar group of pre-Early Start students who took remediation in the fall. The second strategy, a fuzzy regression discontinuity, exploits the cutoff for identification of remediation at CSU (the Entry Level Mathematics placement test), to compare students in the Early Start regime who failed the placement test and had to do summer remediation, to students who just passed it and did not need remediation.

Strategy 1: Difference-in-Differences

By taking advantage of the temporal disjuncture in the implementation of Early Start, we can compare the outcomes for students exposed to Early Start by virtue of the year they entered CSU (between the 2012 and 2015 school year) to the outcomes of students ineligible to participate because the program was not yet available (in 2009-2011). More specifically, using a difference-in-differences strategy, the effect of the Early Start policy on student outcomes is identified by deviations from the pre-treatment trend of a “control” group. Specifically, we estimate the following equation:

$$(1) \quad Y_{itc} = \gamma_0 + \gamma_1 \text{RemedNeed}_i + \gamma_2 \text{ESPolicy}_i * \text{RemedNeed}_i + \beta X_i + \lambda_c + \delta_t + e_{itc}$$

where the dependent variable is a measure of college success (e.g. persistence to year two or GPA) for student i who entered in year t in campus c . The $ESPolicy_t$ variable is an indicator for whether the student entered CSU in the years after Early Start policy implementation; $RemedNeed_i$ identifies whether students did not meet proficiency standards and were identified for remediation (in any year), X_i is a vector of individual-level student controls, and λ_c and δ_t represent CSU campus and year fixed effects, respectively. Given that we focus strictly on students who took the ELM test, $RemedNeed_i$ is an indicator equal to one (zero) if the student scored below (above) the proficiency cutoff. The coefficient on the interaction of $ESPolicy_t$ and $RemedNeed_i$, γ_2 , captures any post-Early Start change in outcomes for the treatment group (i.e. students identified for remediation) relative to the comparison group. The identifying assumption of this difference-in-differences specification is that, absent the policy change, differences in outcomes between the treatment and comparison group would not change. If this identifying assumption holds, then γ_2 captures the causal impact of Early Start on persistence and performance outcomes of students identified for developmental education at CSU.

An important assumption of the difference-in-differences model is that nothing else changed at the time of the policy enactment to differentially affect students in the treatment and control groups. We focus on math because the placement assessment (ELM) threshold for remediation placement and Early Start policy were consistent throughout this time period (whereas English placement cutoffs and decision rules were less clear in this time period). We know of no other change in remediation policies, admissions decisions, or other related CSU policies that may have altered identification of proficiency/remedial placement. We check for any preexisting trends in educational attainment prior to the Early Start policy implementation, and examine how the background characteristics of students change after the policy

implementation. Given the potentially different implementation strategies of Early Start across the system, we include campus fixed effects and interpret these results as an overall Intent-To-Treat effect (i.e. the overall impact of CSU's Early Start Policy given differences in implementation and compliance across the 23-campus system). Additional specifications estimate separate coefficients of γ_2 for students who are in the first, second, or third cohorts exposed to the policy reform (post years 1, 2, or 3). In doing so, we measure how the policy reform impacts evolve over time. In addition, we present campus-specific difference-in-differences results in the discussion section.

Table 2 displays the summary statistics of the analytic sample of ELM test takers, by both remediation status and the pre- and post-Early Start periods. Here we note that although the levels are different by passing status, the descriptive differences between these groups in the pre- and post- Early Start period are not. For example, the high school GPA among those who fail the ELM in the pre-Early Start period is 3.02, while it is 3.11 among those that pass, however, both groups witnessed a small increase in high school GPA between the pre and post-Early Start period: 0.05 for the group that failed the ELM at entry, and 0.07 for the group that passed. Foreshadowing results from the models, there is some indication of non-parallel trends on a couple outcomes, specifically preparedness prior to fall and upper division units.

[Insert Table 2]

Strategy 2: Regression Discontinuity

Our institutional setting also grants us the opportunity to implement a fuzzy regression discontinuity (FRD) design by exploiting the fact that exemption from remediation at CSU is based on a strict system-wide test score cutoff on the Entry Level Math (ELM) assessment. The underlying assumption is that students who score just above and just below the placement threshold required for remediation status are similar in all ways relevant to observed outcomes save for their likelihood of enrolling in the Early Start program. Thus, the causal impact of enrollment in Early Start is identified by comparing students who scored just below the threshold, and subsequently experienced an increased likelihood of enrolling in Early Start, against students who scored above the threshold and thus were exempt from Early Start. The FRD design utilizes the test score cutoff as an instrumental variable for Early Start enrollment (Imbens & Angrist, 1994; Hahn et al., 2001). More formally, we are interested in identifying the effect of Early Start remediation (R) on an outcome Y . Enrollment in Early Start is triggered by whether the student's score (S) on their ELM test fell below the cutoff c . The quantity $\tau_{FRD} = (\tau_{RF}^Y / \tau_{FS})$ is estimated, where the numerator τ_{RF}^Y is the “reduced-form” effect of scoring below c on Y , while the denominator τ_{FS} is the “first stage” effect of scoring below c on Early Start enrollment R . Assuming that unobservable correlates of Y and R are continuous through c , τ_{RF}^Y and τ_{FS} are identified and equivalent to the discontinuities in Y and R at c , respectively; these continuity assumptions imply:

$$\tau_{RF}^Y = \lim_{S \rightarrow c^-} E(Y|S) - \lim_{S \rightarrow c^+} E(Y|S) \text{ and}$$

$$\tau_{FS} = \lim_{S \rightarrow c^-} E(R|S) - \lim_{S \rightarrow c^+} E(R|S).$$

Additional assumptions are needed for the estimation of τ_{FRD} (Imbens and Angrist, 1994).

First, scoring below c must actually increase the likelihood of enrollment in Early Start so that $\tau_{FS} \neq 0$. Moreover, the “exclusion restriction” must hold such that scoring below the cutoff c must only impact outcomes Y through its impact on Early Start enrollment R . This assumption is likely satisfied since a student’s test score is used solely to determine their remedial eligibility in the CSU system. A final assumption is one of “monotonicity”, where there must not be decreases in Early Start enrollment for students who scored below the cutoff c with an increase in enrollment for students who scored above the cutoff c ; this condition holds since almost no students enroll in Early Start if they scored above c . Under these assumptions, τ_{FRD} will be equal to the average effect of Early Start enrollment at the cutoff among “compliers”, or those who enrolled in Early Start because they scored below the cutoff (Imbens and Angrist, 1994).

For estimation, we use the robust local linear estimator proposed by Calonico et al. (2014) across four different bandwidths (+/- 8, 12, 16, and 20). Our specifications include controls for student gender, race, high school GPA, SAT scores, and campus fixed effects. Though these controls are not necessary for identification with the FRD, they should boost statistical power.

The FRD hinges on the assumption that students do not sort around the cutoff in such a way that students who score just above/below the cutoff differ in ways that relate to their academic outcomes. Such sorting would violate the aforementioned continuity assumptions. Fortunately, in our setting, sorting due to direct manipulation of the test scores or of the test score process is extremely unlikely since the exam is administered centrally by the CSU system. Moreover, for our FRD analysis, we reduce our sample of students to those whose final test attempt was in March (i.e. relatively early in the year post-admission). This is done because students had the opportunity to retake the ELM if they were unsatisfied with their score;

continuity assumptions would be violated if we used students who scored just below the cutoff and retook the test and subsequently scored above the cutoff, and if these students differed in unobservable ways than those students whose highest test score (regardless of number of attempts) was just below the cutoff. Institutional knowledge suggests that students whose final test attempt was made in March are students who were very likely to have taken the test only once, and so focusing on this sample of students improves the likelihood of the validity of the FRD identification strategy.¹¹

We provide several pieces of empirical evidence to further support the validity of the FRD design. First, we implement the FRD using student covariates as the outcome variable. Finding no discontinuous changes in observable student characteristics at the cutoff should ease concerns that there may be unobservable characteristics that discontinuously change at the cutoff. Results from this analysis are presented in Table 3. Across the four bandwidths considered, we estimate no statistically significant relationship between Early Start enrollment and student gender, race, high school GPA, or SAT score. Additional empirical support for the FRD comes from Figure 2, which plots the distribution of observations by the running variable with 95% McCrary confidence intervals (McCrary, 2008). One can see that the majority of students who take the test fail, but more importantly, we do not observe a significant jump in the distribution of test scores just above the cutoff, which otherwise would suggest some type of manipulation out of the Early Start requirement.

[Insert Table 3 and Figure 2]

¹¹ Ideally, we'd focus strictly on students' scores on their first attempt (regardless of number of attempts). Unfortunately, our data only include the student's highest test score and the date of their final test attempt. By focusing our sample to students whose final test attempt was on the first major administrative month (March), the observed test scores in the data likely correspond to the student's only test attempt.

5. Results

We begin by describing the trends in student inputs throughout the study period. Figure 3 displays the high school characteristics (GPA and SAT scores) by math remediation status for students entering CSU under three conditions: demonstrating college readiness at entry (based on high school examinations), and thus exempt from the additional ELM placement test; entering students that are required to take the ELM and pass the remediation threshold; and entering students that take the ELM and fail the remediation threshold and, thus, are required to enroll in subsequent developmental coursework. Not surprisingly, students that enter CSU exempt from any additional mathematics remedial placement exam have a higher GPA and SAT scores than those required to take an additional placement exam. Overall trends also suggest that, for all groups, high school GPA has been steadily on the rise among CSU entrants, while SAT math scores have witnessed a steady modest decline.

[Insert Figure 3]

Our first strategy to test the impact of Early Start employs a difference-in-differences approach, comparing changes in outcomes for students needing remediation (before and after the Early Start regime, when requirements shifted from the fall to the summer) against changes in outcomes for students who did not need remediation. Table 4 presents findings from the difference-in-differences models. Column 1 includes models with no controls, in column 2 we add demographic controls (race and gender), in column 3 we also include controls for prior academic achievement (high school GPA and SAT scores), and column 4 represents our fully specified model where we include campus and year fixed effects. For all specifications, standard

errors are clustered at the campus level using wild cluster bootstrapping (Cameron et al., 2008).

Results reveal that Early Start did lead to a significant “first-stage” effect of increasing the likelihood a student satisfied their remediation requirements prior to the fall. Thus, there is evidence of compliance with the summer requirement for remediation. Early Start seems to be associated with a reduction in first term GPA among students requiring remediation, though the magnitude is quite small at about 0.03 of a grade point average and not statistically significant. Early Start did not have an effect on the likelihood a student took an upper division course (extensive margin of upper division course taking), but did have a positive effect on total upper division courses taken (intensive margin of upper division course taking). In other words, Early Start did not seem to induce more students to take upper division courses, but it did induce more upper division unit accumulation among those who were already taking upper division courses. We also look at Early Start impacts on total units (constructed as total units accumulated through the sample), which need not be positive despite the upper division results since students may simply reach their target units more quickly; we only find a positive effect for the fully specified model. Finally, we do not observe an effect of Early Start on persistence to year two or year three.

We test the effects of the Early Start policy by year for the main outcomes in Table 5, and include several pre-Early Start years. Results are consistent overall, but suggest the negative effects on GPA are greater in more recent years, and the positive effects on upper division units are concentrated in the second and third year of the Early Start reform efforts. Persistence rates are higher each year (relative to 2009), but based on results presented in Table 4, we know differences in persistence rates, on average, are not significantly different in the post-Early Start years when compared to the pre-Early Start years. Additional years of data may further

illuminate persistence rates for additional cohorts and longer-term impacts on completion and time to degree.

[Insert Tables 4 & 5]

Fuzzy regression discontinuity results:

Next, we identify the impacts of Early Start on student outcomes by utilizing a fuzzy regression discontinuity (FRD) design. Identification comes from comparing students during the Early Start years who scored just below the cutoff on their ELM test, who then were subsequently required to enroll in Early Start, against students who scored just above the cutoff and thus were exempt from needing Early Start. Recall that this identification strategy effectively compares students who enrolled in Early Start remediation against those who did not require *any* remediation, while the difference-in-differences strategy, in an intent-to-treat framework, compares Early Start summer remediation students against fall remediation students.

We start in Figure 4 by plotting averages of various outcome variables across ELM test scores. In the first graph (top-left), we plot the fraction of students who enrolled in Early Start against their ELM test score. Unsurprisingly, we find virtually no enrollment in Early Start among students that passed the ELM test, while students who scored below the required cutoff were significantly more likely to enroll in Early Start. This corroborates our earlier finding from the difference-in-differences results that summer remedial completion rates increased during the Early Start regime.

[Insert Figure 4]

The remaining five graphs of Figure 4 can be interpreted as unconditional “reduced form” estimates of τ_{RF}^Y i.e. differences in outcomes experienced by students who scored just below versus just above the cutoff. Starting with first term GPA, we first see that as expected, higher ELM test scores correlate with higher GPA. However, students who scored just below the cutoff do not appear to have received a higher GPA in their first term when compared to students just above the cutoff, despite being “similar” in all ways to the control group except for an increased likelihood in enrolling in Early Start as a result of their ELM score. Thus, this analysis seems to suggest that Early Start remediation was not effective in improving students’ preparedness, when compared to similar students who received no remediation.

Moving to the remaining four graphs of Figure 4, which consider more “downstream” outcomes, we again observe that students who score higher on the ELM assessment tend to experience more “positive” outcomes: they enroll in more upper division units in their first year, they finish the sample with more accumulated units, and they persist into year 2 at a higher rate. The only two outcomes for which we plausibly observe discontinuities are for total unit accumulation (through the student’s entire tenure at the CSU) and persistence rates into the student’s second year. For both of these outcomes, it appears that students who scored below the cutoff were *less* likely to persist into year two and accumulated *fewer* units than students who scored just above the cutoff. This suggests that Early Start may have nudged students out of the CSU system, perhaps because students identified as needing Early Start (and potentially as a result of subsequently enrolling in Early Start) conclude that college is not for them.

Table 6 presents regression results from our FRD. Each panel considers a separate outcome variable and each column considers a different bandwidth such that each cell reports an estimated discontinuity in an outcome variable. Each regression controls for student gender, race,

high school GPA, SAT test taking and scores, and campus and year fixed effects. The first panel reports the conditional first stage discontinuity in Early Start enrollment rates (i.e. conditional estimates for τ_{FS}); across all four specifications, we find that students who just barely scored above the cutoff on the ELM test were roughly 28 percentage points less likely to enroll in Early Start than students who scores just barely below the cutoff. Results from our remaining five outcome variables report conditional estimates for τ_{FRD} , and can be interpreted as the impact of enrolling in Early Start on the outcome for “compliers”. These results largely reflect the results from reduced-form graphical analyses. Namely, we estimate no statistically significant changes in first term GPA or upper division unit accumulation in year one for students who enrolled in Early Start. We do observe significant reductions in total unit accumulation through the student’s tenure. Results for persistence into year two also remain negative, but are statistically significant at the 10% level for only three of the four considered bandwidths.

[Insert Table 6]

6. *Discussion/Conclusion*

Today, more than ever, colleges and universities are focused on improving persistence and degree completion for their students. A well-established tension in collegiate remediation is, on the one hand, remediation as a necessary tool for skills development to address college readiness gaps from K-12, and, on the other hand, a costly program that potentially sends a negative signal and a force of discouragement among entering college students (Scott-Clayton & Rodriguez, 2015; Deil-Amen & Rosenbaum, 2002; Boatman & Long, 2018). The Early Start reform speaks to both of these competing hypotheses: first, by allowing students to develop necessary college readiness skills before they actually begin college-level courses; and second, through potential

discouragement by having to do “summer school.” Importantly, Early Start did not require reform to what remediation actually consisted of across CSU campuses. That is, it did not mandate any particular practices in developmental education that should occur in the summer. This suggests that many CSU campuses simply offered existing remedial courses in the summer instead of (or in addition to) the fall.

Our results present a less than optimistic picture of simply doing remediation earlier (in the summer before freshman year). Overall, we do not find that this policy contributed to improving student performance in the first year or subsequent persistence rates. This finding is consistent with several other prior studies exploring remediation effects using similar analytic strategies in several institutional contexts (Boatman & Long, 2018; Scott-Clayton & Rodriguez, 2015; Martorell & McFarlin, 2011; Calcagno & Long, 2008). However, we do find positive evidence that Early Start contributed to more upper division enrollment, particularly for some students already inclined to enroll in more advanced coursework. Moreover, the slight penalty associated with Early Start on GPA is quite small. By having students satisfy their remedial requirements in the summer before their freshman year, perhaps more courses, including upper division courses, became possible—structurally or psychologically—to students. Thus, if indeed summer remedial enrollment leads to a boost of human capital accumulation (as measured by more units), but at the expense of GPA, this may be worth the expense. Of course students often care a lot about their grades, and we’d need additional data to better understand course selection in order to more fully tease this out.

Our goal in evaluating the Early Start policy is to determine whether this sweeping requirement of summer remedial enrollment had an impact on CSU student outcomes. Although the policy was enacted system-wide, campuses likely varied in their implementation and format

for Early Start. A full accounting of these differences is beyond the scope of this paper, however, we nevertheless can evaluate these policy effects at the campus level. Figure 5 plots the difference-in-differences results for four outcomes at the campus level (the campuses are not ordered in particular way to protect identification). From these figures we note, first, that a majority of campuses saw no impact on first term GPA, save for a few campuses with a sizeable (0.10 of a GPA) positive and statistically significant effect. The positive effect on upper division units is largely concentrated at two campuses, with an additional several campuses with statistically significant positive effect on upper division unit accumulation—the magnitudes for all of these are very small (less than one unit). Despite overall null findings on persistence, we note that there are several campuses with positive effects of Early Start on second year persistence, including one of sizeable magnitude (0.15 in percentage point units), and a couple with small negative effects on persistence to the second year. Finally, we see no effect of Early Start on persistence to the third year for any campus, though these results should be interpreted with some caution as they are based on the first cohorts to have been under the Early Start regime. In sum, we note considerable variation across campuses. Subsequent work may interrogate these differences further, investigating differences in Early Start compliance, and whether and how campuses implemented different instructional models for summer remediation (e.g. online versus face-to face), as well as credit types (e.g. 1 or 3 units).

[Insert Figure 5]

Increased attention to the over- and mis-placement of students into remedial coursework has contributed to greater scrutiny of remediation policies and procedures, and to a more widely held sentiment that remediation is “perhaps the biggest barrier to improving the nation’s college graduation rates” (Fain, 2013). Consequently, several states and institutions have taken steps to alter, reduce, or even eliminate remedial course offerings in public colleges and universities (Ross, 2014). Most recently, CSU has dramatically altered remediation practices with executive order 1110 eliminating both CSU’s longstanding homegrown assessments in English and mathematics (ELM and EPT), and by making all remedial coursework co-requisite (signed this past summer and effective for students entering in Fall 2018 as freshmen). At the time of this writing, the dust has not quite settled on how campus administrators and faculty across the system received this latest policy nor how it was to be implemented. Nevertheless, Early Start stands to play an even greater role as the primary form of pre-collegiate developmental course offering for CSU students. As such, CSU leaders should take a close look at our existing evidence of its impact.

REFERENCES

- Adelman, C. (1999). *Answers in the toolbox: Academic intensity, attendance patterns, and bachelor's degree attainment*. Washington, DC: U.S. Department of Education Office of Educational Research and Improvement.
- Attewell, P., Lavin, D., Domina, T., & Levey, T. (2006). New evidence on college remediation. *Journal of Higher Education*, 77, 886–924.
- Bettinger, E. P., Boatman, A., & Long, B. T. (2013). Student supports: Developmental education and other academic programs. *The Future of Children*, 23(1), 93–115.
- Bettinger, E., & Long, B. T. (2007). Institutional responses to reduce inequalities in college outcomes: Remedial and developmental courses in higher education. In S. Dickert-Conlin & R. Rubenstein (Eds.), *Economic inequality and higher education: Access, persistence, and success* (pp. 69–100). New York, NY: Russell Sage Foundation.
- Bettinger, E., & Long, B. T. (2009). Addressing the needs of underprepared students in higher education: Does college remediation work? *Journal of Human Resources*, 44, 736–771.
- Boatman, A., & Long, B. T. (2018). Does remediation work for all students? How the effects of postsecondary remedial and developmental courses vary by level of academic preparation. *Educational Evaluation and Policy Analysis*, 40(1), 29-58.
- Calcagno, J. C., & Long, B. T. (2008, July). *The impact of postsecondary remediation using a regression discontinuity approach: Addressing endogenous sorting and noncompliance* (NBER Working Paper No. 14194). Cambridge, MA: National Bureau of Economic Research.
- Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, 82, 2295-2326.
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. (2008) Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics* 90.3: 414-427.
- Complete College America. (2017). Remediation: Higher Education's Bridge to Nowhere. Retrieved from: <https://completecollege.org/wp-content/uploads/2017/11/CCA-Remediation-final.pdf>
- Deil-Amen, Regina, and James E. Rosenbaum. (2002). The Unintended Consequences of Stigma-Free Remediation. *Sociology of Education* 75(3):249-68.

- Fain, P. (2013, September 13). Going to the root of the problem. *Insider Higher Ed*. Retrieved from: <https://www.insidehighered.com/news/2013/09/13/promising-remedial-math-reform-tennessee-expands>
- Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, 69, 201-209.
- Howell, J., Kurlaender, M., & Grodsky, E. (2010). Postsecondary Preparation and Remediation: Examining the Effect of the Early Assessment Program at California State University. *Journal of Policy Analysis and Management*, 29(4): 726–748.
- Imbens, G & Angrist, J. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62, 467-475.
- Jackson, J. & Cook, K. (2016, May). *Improving college graduation rates: A closer look at California State University*. Sacramento, CA: Public Policy Institute of California. Retrieved from: http://www.ppic.org/content/pubs/report/R_516JJR.pdf
- Jimenez, L., Sargrad, S., Morales, J., & Thompson, M. (2016). *Remedial education: The cost of catching up*. Washington, DC: Center for American Progress.
- Kurlaender, M., Grodsky, E., Howell, J., & Jackson, J. (2016). Ready or Not? California's Early Assessment Program and the Transition to College.
- Lesik, S. (2007). Do developmental mathematics programs have a causal impact on student retention? An application of discrete-time survival and regression-discontinuity analysis. *Research in Higher Education*, 48, 583–608.
- Logue, A. W., Watanabe-Rose, M., & Douglas, D. (2016). Should Students Assessed as Needing Remedial Mathematics Take College-Level Quantitative Courses Instead? A Randomized Controlled Trial. *Educational Evaluation and Policy Analysis*, 38(3), 578-598.
- Martorell, P., & McFarlin Jr, I. (2011). Help or hindrance? The effects of college remediation on academic and labor market outcomes. *The Review of Economics and Statistics*, 93(2), 436-454.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2), 698-714.
- Ngo, F., & Melguizo, T. (2016). How can placement policy improve math remediation outcomes? Evidence from experimentation in community colleges. *Educational Evaluation and Policy Analysis*, 38(1), 171-196.

- Scott-Clayton, J., Crosta, P. M., & Belfield, C. R. (2014). Improving the targeting of treatment: Evidence from college remediation. *Educational Evaluation and Policy Analysis*, 36(3), 371-393.
- Scott-Clayton, J., & Rodriguez, O. (2015). Development, discouragement, or diversion? New evidence on the effects of college remediation. *Education Finance and Policy* 10, 4-45.
- Shields, K. A., & O'Dwyer, L. M. (2017). Remedial Education and Completing College: Exploring Differences by Credential and Institutional Level. *The Journal of Higher Education*, 88(1), 85-109.
- Sparks, D., & Malkus, N. (2013). *First-year undergraduate remedial coursetaking: 1999-2000, 2003-04, 2007-08*. Statistics in Brief. NCES 2013-013. National Center for Education Statistics.
- U.S. Department of Education, National Center for Education Statistics. (2011). *The condition of education 2011 (NCES 2011-033)*. Washington, DC: U.S. Government Printing Office.

Table 1: Summary statistics for population of CSU first-time freshmen 2009-2015

Variable	Mean	Std. Dev.	N
Male	0.43		415393
Hispanic	0.41		415393
Black	0.05		415393
Asian	0.17		415393
White	0.27		415393
High school GPA	3.25	0.64	415393
SAT composite	894.6	346.25	415393
Took EPT test	0.23		415393
Took ELM test	0.56		415393
Passed conditional on taking ELM test	0.33		231005
Outcome: Math exempt prior to fall	0.75		415393
Outcome: First term GPA	2.83	0.88	398500
Outcome: Upper-division units in year 1	1.02	3.80	415393
Outcome: Took an upper-division course in year 1	0.14		415393
Outcome: Total units at CSU	77.67	59.63	415393
Outcome: Persist to year 2	0.85		348789
Outcome: Persist to year 3	0.77		283387

Notes: Summary statistics for the population of fall first-time Freshmen in the California State University system. Sample size drops for the first term GPA outcome since not all students enrolled in the fall receive a term GPA (e.g. students may drop out, or take course without letter grading). Sample size for persistence rate outcomes drops due to censoring with sample ending in 2015.

Table 2: Summary statistics for Mathematics College Readiness (ELM) test takers

	Failed placement test (Remediation needed)				Passed placement test (No remediation needed)				DiD [p-value]	
	Pre-ES (2009-2011)		Post-ES (2012-2015)		Pre-ES (2009-2011)		Post-ES (2012-2015)			
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		
<i>Mathematics (ELM) Test Takers</i>	<i>(n=64,161)</i>				<i>(n=90,943)</i>				<i>(n=32,977)</i>	
Male	0.31		0.31		0.46		0.45		0.01 [0.065]*	
White	0.19		0.14		0.28		0.23		0.01 [0.264]	
Hispanic	0.49		0.60		0.37		0.46		0.01 [0.377]	
Asian	0.12		0.11		0.19		0.18		0.00 [0.550]	
High school GPA	3.02	0.66	3.07	0.65	3.11	0.78	3.18	0.71	-0.02 [0.172]	
SAT Composite	746.45	315.07	758.27	282.22	874.20	334.04	886.84	297.39	-0.705 [0.911]	
ELM test score	-13.93	8.31	-14.77	8.49	7.90	7.28	6.98	6.61	0.17 [0.418]	
Outcome: Math exempt prior to fall	0.11		0.51		1.00		1.00		0.42 [0.000]***	
Outcome: First term GPA	2.65	0.96	2.70	0.93	2.74	0.91	2.82	0.84	-0.03 [0.202]	
Outcome: Upper-div. units in Y1	1.31	4.53	1.33	4.96	1.47	4.66	1.30	4.60	0.21 [0.003]***	
Outcome: Took upper-div. course in Y1	0.17		0.14		0.20		0.16		0.00 [0.367]	
Outcome: Total units at CSU	94.26	63.20	48.04	41.71	105.68	61.38	58.46	43.48	0.62 [0.523]	
Outcome: Persist to Y2	0.81		0.81		0.86		0.86		0.00 [0.900]	
Outcome: Persist to Y3	0.71		0.70		0.77		0.76		0.00 [0.228]	

Notes: Each cell under “Mean” reports the average for the covariate for one of four subgroups: before/after Early Start’s passing for those who passed/failed the ELM test. The final column reports the unconditional difference-in-differences estimate for each covariate as the dependent variable. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 3: Fuzzy Regression Discontinuity Results - Balance Test (local linear)

	(1)	(2)	(3)	(4)
<u>Outcome: Male student</u>				
Enrolled in Early Start	0.029 (0.092)	0.041 (0.101)	0.031 (0.111)	0.027 (0.149)
Observations	13679	11609	9264	6666
<u>Outcome: URM student</u>				
Enrolled in Early Start	0.124 (0.153)	0.123 (0.156)	0.109 (0.157)	0.047 (0.170)
Observations	13679	11609	9264	6666
<u>Outcome: White student</u>				
Enrolled in Early Start	-0.097 (0.140)	-0.077 (0.145)	-0.064 (0.146)	-0.008 (0.160)
Observations	13679	11609	9264	6666
<u>Outcome: High school GPA</u>				
Enrolled in Early Start	-0.035 (0.185)	-0.036 (0.188)	-0.068 (0.183)	-0.078 (0.181)
Observations	13679	11609	9264	6666
<u>Outcome: SAT composite score</u>				
Enrolled in Early Start	-38.663 (39.873)	-49.899 (44.017)	-63.563 (50.560)	-73.495 (64.435)
Observations	13679	11609	9264	6666
Bandwidth	+/- 20	+/- 16	+/- 12	+/- 8

Notes: Each cell reports the estimated effect of passing the ELM test (i.e. not needing to enroll in Early Start) from a fuzzy regression discontinuity design, where Early Start enrollment is instrumented for by the test score cutoff. Standard errors clustered at the campus level. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 4: Main difference-in-differences results

	(1)	(2)	(3)	(4)
<u>Outcome: Math satisfied prior to Fall</u>				
First year 2012+ X Below test cutoff	0.422*** (0.019)	0.421*** (0.019)	0.420*** (0.019)	0.446*** (0.021)
Observations	231005	231005	231005	231005
<u>Outcome: First term GPA</u>				
First year 2012+ X Below test cutoff	-0.026 (0.024)	-0.026 (0.025)	-0.029 (0.026)	-0.036 (0.025)
Observations	220330	220330	220330	220330
<u>Outcome: Upper division units in Y1</u>				
First year 2012+ X Below test cutoff	0.191*** (0.070)	0.193*** (0.072)	0.086** (0.039)	0.087** (0.040)
Observations	231005	231005	231005	231005
<u>Outcome: Took an upper division course</u>				
First year 2012+ X Below test cutoff	0.006 (0.005)	0.007 (0.005)	0.002 (0.004)	0.001 (0.003)
Observations	231005	231005	231005	231005
<u>Outcome: Total units</u>				
First year 2012+ X Below test cutoff	1.010 (0.888)	0.986 (0.880)	0.695 (0.906)	2.676*** (0.812)
Observations	231005	231005	231005	231005
<u>Outcome: Persist to Y2</u>				
First year 2012+ X Below test cutoff	-0.000 (0.005)	-0.000 (0.005)	-0.000 (0.005)	0.000 (0.005)
Observations	197942	197942	197942	197942
<u>Outcome: Persist to Y3</u>				
First year 2012+ X Below test cutoff	0.004 (0.004)	0.004 (0.004)	0.004 (0.005)	0.005 (0.005)
Observations	162993	162993	162993	162993
Gender/Race Controls		X	X	X
HS GPA/SAT Controls			X	X
Campus/Entry Year Fixed Effects				X

Notes: Each cell reports a difference-in-differences coefficient from a single regression. Standard errors clustered at the campus level. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 5: Difference-in-differences results by post Early Start year interactions

	First term GPA	Upper div. units	Persist to year 2	Persist to year 3
First year 2010 X Below test cutoff	0.016 (0.040)	0.002 (0.067)	0.020*** (0.006)	0.027*** (0.006)
First year 2011 X Below test cutoff	0.011 (0.043)	0.032 (0.064)	0.016** (0.008)	0.018* (0.009)
First year 2012 X Below test cutoff	-0.018 (0.045)	-0.050 (0.066)	0.005 (0.006)	0.019*** (0.007)
First year 2013 X Below test cutoff	-0.003 (0.047)	0.186*** (0.061)	0.015* (0.008)	0.020** (0.009)
First year 2014 X Below test cutoff	-0.030 (0.049)	0.171** (0.073)	0.015 (0.009)	
First year 2015 X Below test cutoff	-0.064 (0.051)	0.073 (0.096)		
Observations	220330	231005	197942	162993
Gender/Race Controls	X	X	X	X
HS GPA/SAT Controls	X	X	X	X
Campus/Entry Year Fixed Effects	X	X	X	X

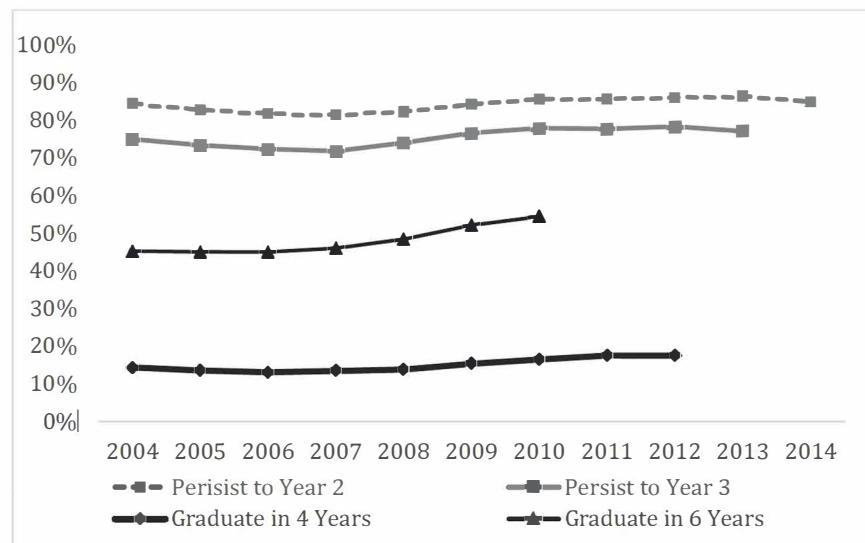
Notes: Each column reports coefficients from a single regression. Base components for each interaction not reported. Omitted interaction is “First year 2009 X Below test cutoff”. Standard errors clustered at the campus level. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 6: Fuzzy Regression Discontinuity Results - Student Outcomes (local linear)

	(1)	(2)	(3)	(4)
<u>First stage: Early Start enrollment</u>				
First-stage discontinuity (passed ELM)	-0.285*** (0.016)	-0.281*** (0.018)	-0.281*** (0.020)	-0.279*** (0.025)
Observations	13679	11609	9264	6666
Mean above cutoff	0.000	0.000	0.000	0.001
<u>Outcome: First term GPA</u>				
Enrolled in Early Start	-0.000 (0.102)	-0.013 (0.107)	0.017 (0.111)	-0.083 (0.133)
Observations	13039	11071	8844	6379
Mean above cutoff	2.871	2.861	2.855	2.857
<u>Outcome: Upper division units in Y1</u>				
Enrolled in Early Start	0.608 (0.446)	0.278 (0.573)	-0.344 (0.787)	-0.592 (1.066)
Observations	13679	11609	9264	6666
Mean above cutoff	1.323	1.282	1.323	1.338
<u>Outcome: Took an upper division course</u>				
Enrolled in Early Start	-0.003 (0.029)	-0.017 (0.036)	-0.061 (0.046)	-0.094 (0.071)
Observations	13679	11609	9264	6666
Mean above cutoff	0.128	0.127	0.126	0.127
<u>Outcome: Total units</u>				
Enrolled in Early Start	-8.699*** (2.461)	-10.881*** (2.840)	-13.388*** (3.485)	-14.710*** (5.103)
Observations	13679	11609	9264	6666
Mean above cutoff	35.491	35.318	35.001	34.876
<u>Outcome: Persist to Y2</u>				
Enrolled in Early Start	-0.043* (0.023)	-0.049* (0.028)	-0.045 (0.032)	-0.060* (0.031)
Observations	6770	5777	4666	3320
Mean above cutoff	0.867	0.864	0.860	0.861
Bandwidth	+/- 20	+/- 16	+/- 12	+/- 8
Gender/Race Controls	X	X	X	X
HS GPA/SAT/Took EPT controls	X	X	X	X
Campus FE	X	X	X	X

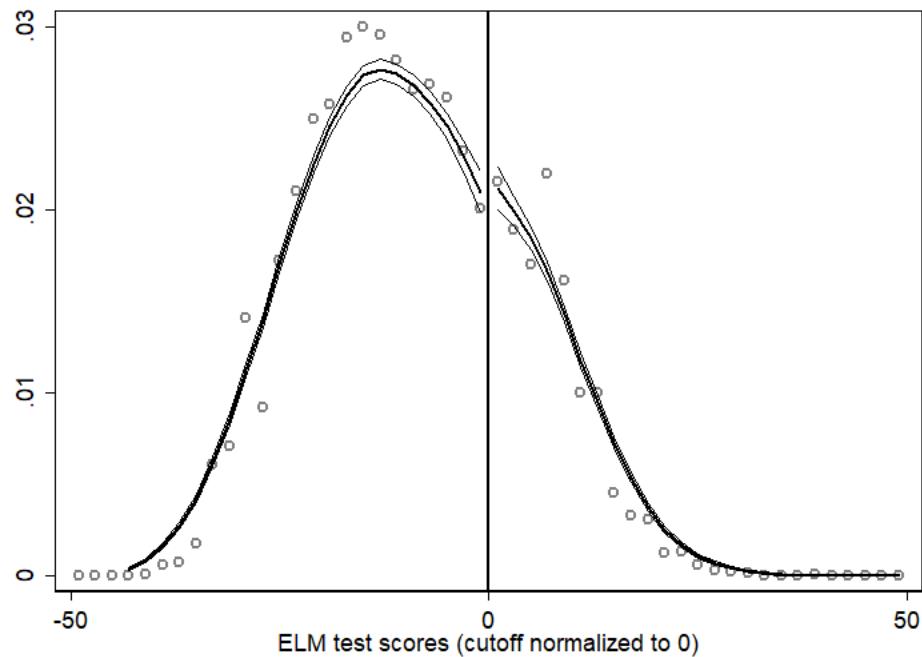
Notes: Each cell reports the estimated effect of passing the ELM test (i.e. not needing to enroll in Early Start) from a fuzzy regression discontinuity design, where Early Start enrollment is instrumented for by the test score cutoff. Standard errors clustered at the campus level. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Figure 1: CSU Persistence and Completion Outcomes for First Time Freshmen Cohorts



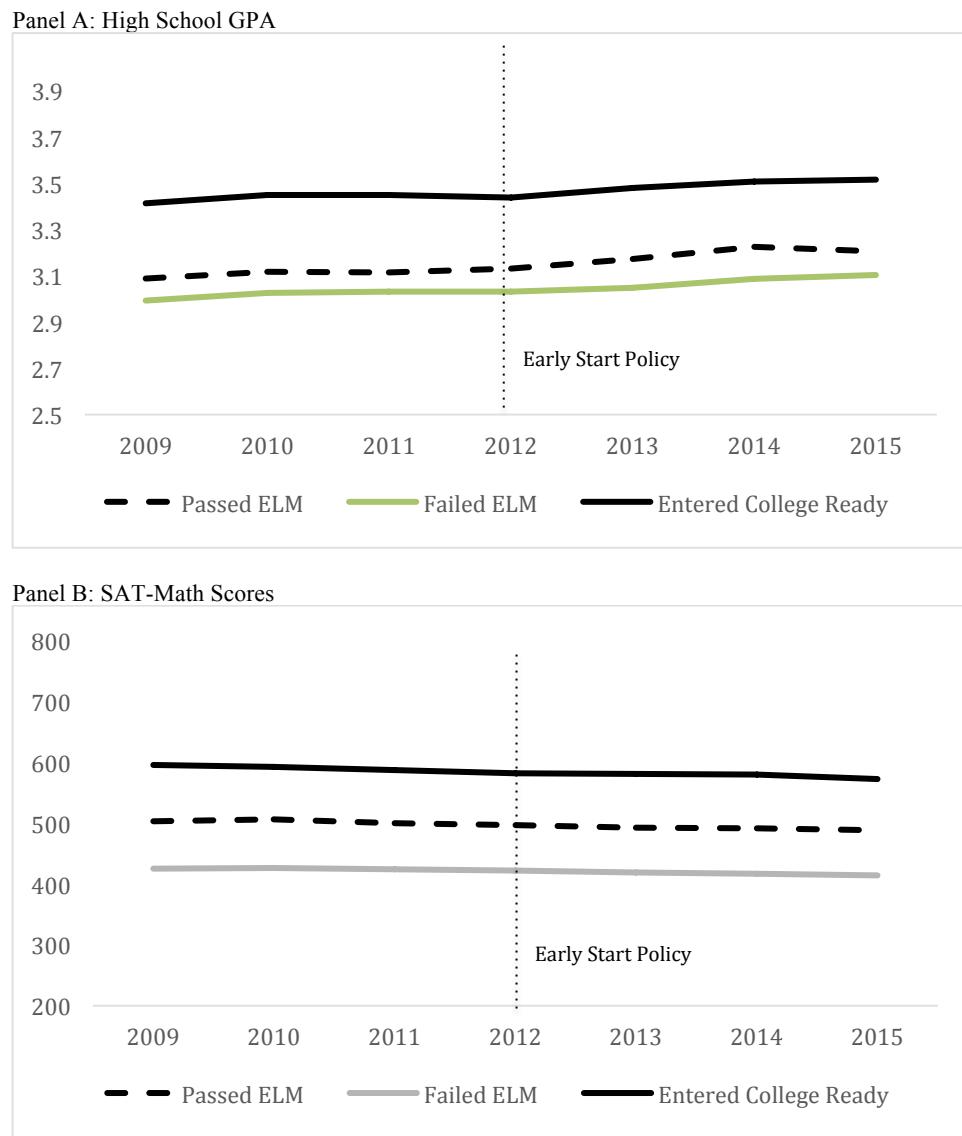
Notes: Authors' calculations, Source: CSU Chancellor's Office Analytics Studies.

Figure 2: Density of ELM Test Scores with 95% McCrary Confidence Intervals



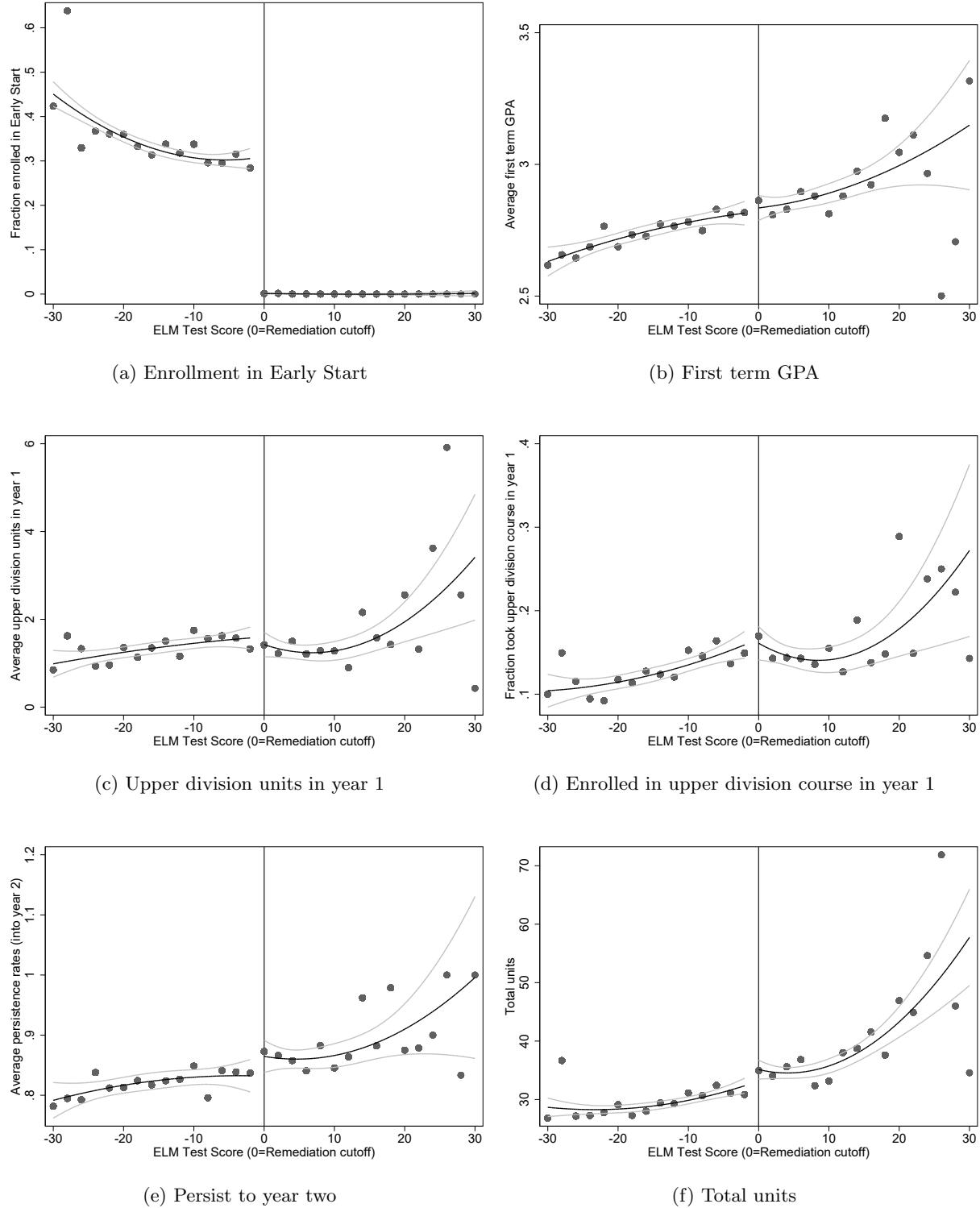
Notes: Figure depicts distribution of ELM test scores (rescaled so passing is equal to zero). Estimated distribution and 95% confidence intervals estimated using McCrary (2008).

Figure 3: High School Characteristics Over Time by CSU Mathematics Remediation Status at Entry



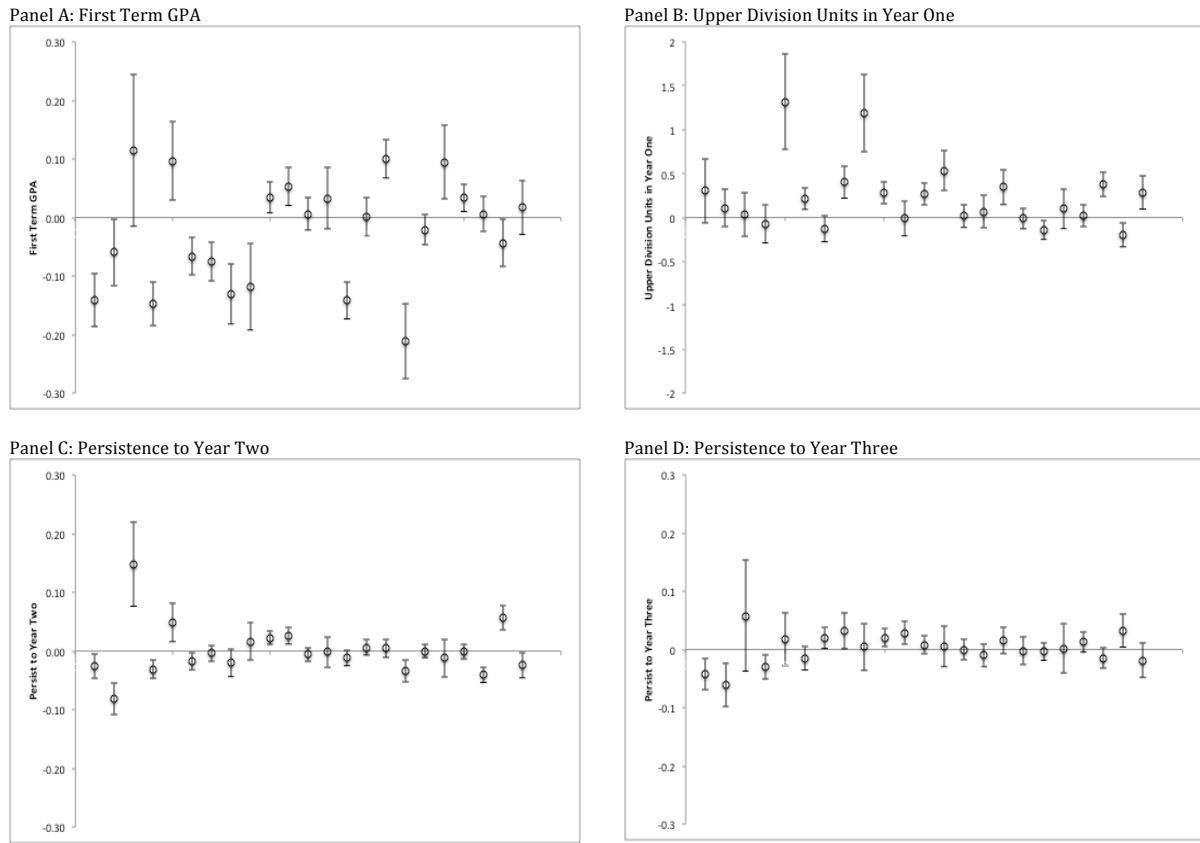
Notes: Authors' calculations, Source: CSU Chancellor's Office Analytics Studies.

Figure 4: Averages of Outcomes by ELM Test Score



Notes: Each graph plots averages of an outcome variable against ELM test scores. Quadratic curves with 95% confidence intervals are estimated on both sides of the cutoff.

Figure 5: Difference-in-differences Results by Campus



Notes: Fully specified model, including student controls and year fixed effects, estimated by campus.