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 2019

This document is the accepted version of a manuscript that was published as an article in the *Journal of Policy Analysis & Management*
(April 2020, vol. 39, issue 2, pp. 348–375, https://doi.org/10.1002/pam.22198)

ABSTRACT
Remediation has long been a costly way to address the misalignment between K-12 and higher education. In 2011, the California State University (CSU), the nation's largest public four-year university system, enacted Early Start, *requiring students needing remediation to enroll in such courses in the summer before their freshman year. We estimate the impact of Early Start summer remediation relative to both traditional fall remediation and relative to no remediation at all. Our results suggest Early Start summer remediation has not improved student performance or persistence relative to either alternative. As many states move away from remedial courses altogether, there is continued need for both innovation and for evidence in policy and practice to improve college readiness and success.*

ACKNOWLEDGMENTS
The research reported here was undertaken through the Center for the Analysis of Postsecondary Readiness and supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C140007 to Teachers College, Columbia University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education, or of the agencies providing the data. We thank Ed Sullivan, Paco Martorell, and James Benson for their helpful feedback on this work.

MICHAL KURLAENDER is a Professor of Education Policy at the University of California at Davis, School of Education, One Shields Avenue, Davis, CA 95616 (e-mail: mkurlaender@ucdavis.edu).

LESTER LUSHER is an Assistant Professor of Economics at the University of Hawaii at Manoa, Sanders Hall #542, 2424 Maile Way, Honolulu, HI 96822 (e-mail: lrlusher@hawaii.edu).

MATTHEW CASE is the Assistant Director of Policy Analytics and Institutional Research in the Office of the Chancellor at California State University, 401 Golden Shore, Long Beach, CA 90802 (e-mail: mcase@calstate.edu).

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 60 percent of students who enter a public baccalaureate-granting institution obtain a degree within six years (U.S. Department of Education, 2018). 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 to be 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 a developmental (remedial) course, with some estimates of participation in remediation closer to 50 to 60 percent or higher (Scott-Clayton, Crosta, & Belfield, 2014).[1]

In 2011, the California State University (CSU), the nation's largest public four-year university system, enacted *Early Start*, requiring students needing remediation to enroll in such courses in the summer before their freshman year. Facing enormous numbers of

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

students requiring remedial coursework in mathematics and English (at the time of adoption, less than 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]

In this paper, we investigate 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. On the other hand, students requiring remediation in the Early Start era at CSU have not had better performance or persistence outcomes than similar students who 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

---

[2] See http://asd.calstate.edu/performance/combo/2010/Combo_Prof_Sys_fall2010.htm.
[3] Memo from CSU Chancellor Charles Reed to CSU Presidents on June 11, 2010. See https://www.calstate.edu/eo/EO-1048.pdf.

and unit accumulation outcomes when compared to those not requiring remediation. Thus, although Early Start is successfully remediating students in the summer before the fall term, we do not find that this earlier remediation improves later student outcomes (a finding that corroborates with much of the extant literature on college remediation).

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 either because 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 or format 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 who 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.

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 (Kurlaender et al., 2016); assessment types and placement decisions (Ngo & Melguizo, 2016; Scott-Clayton et al., 2014; Scott-Clayton & Rodriguez, 2015); and types of developmental course options, including co-requisite and stretch courses (Logue, Watanabe, & Douglas, 2016). The Early Start policy offers an opportunity to evaluate remediation practices at scale, mainly the shift in the timing of remediation to the summer prior to freshman entry. 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.

The rest of the paper is organized as follows. In the next section, we summarize the literature on postsecondary remediation efforts, focusing on recent policy changes. The third section describes the CSU system and the Early Start policy. In the fourth section, we describe our difference-in-differences and regression discontinuity research designs, and the fifth section details the results. Finally, in the sixth section, we discuss our findings, including a discussion of campus-level differences, and conclude by describing CSU's most recent remediation reform to eliminate its primary remediation placement exams as well as pre-collegiate credits altogether.

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 estimated costs to students are over $1 billion annually (Jimenez et al., 2016; Scott-Clayton, Crosta, & Belfield, 2014). 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 arrived as freshmen. This resulted in thousands of students who were 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 factors that make these students different (e.g., weaker skills from K-12 schooling, 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, 1999; Attewell et al., 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 (Bettinger & Long, 2009; Calcagno & Long, 2008; Lesik, 2007; Martorell & McFarlin, 2011; Scott-Clayton & Rodriguez, 2015). 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 (Calcagno & Long, 2008; Martorell & McFarlin, 2011; Scott-Clayton & Rodriguez, 2015). 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. Notably, 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 the 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 cutoffs) 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 or are substantially reforming it as a co-requisite experience (Complete College America, 2017). For example, CUNY Start, is a pre-collegiate experience for intensive preparation (i.e., a semester of full-time 25 hours/week, with part-time, options) in academic reading/writing and math, along with broader college success advising for students entering CUNY with significant remedial needs.[4] The effectiveness of CUNY Start (a substantially more intensive remediation experience for community college students) is currently under investigation by MDRC and the Community College Research Center. 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 several important ways. 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 across 23

[4] For a description of CUNY Start see: http://www1.cuny.edu/sites/cunystart/program/cuny-start/. For information on the evaluation in progress see: https://www.mdrc.org/project/addressing-students-remedial-needs-evaluation-cuny-start-and-other-strategies#overview.

campuses. Specifically, we test an important reform to collegiate remediation—requiring summer enrollment to satisfy developmental coursework—adopted at scale across the CSU system. And, we employ multiple strategies to evaluate the causal impacts of this reform on several important outcomes, offering multiple counterfactual comparisons to remediation status.

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, enrolling over 400,000 undergraduate students and serving the top-third of California high school graduates (as dictated by California's Master Plan for Higher Education).[5] 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 diverse, comprehensive, and largely broad-access, save for a few campuses that are slightly more selective and accept only about one-third of eligible applicants. CSU, by comparison to "peer institutions" (based on selectivity and average SAT scores), has historically had a higher SAT or ACT cutoff for exemption from

[5] California's Master Plan for Higher Education, adopted in 1960, defines specific roles for each segment of the State's higher education system: the most selective University of California (UC) is reserved for the top 12.5 percent of California's eligible high school graduates; the California State University (CSU) is reserved for the top 33.3 percent of California's eligible high school graduates; and the California Community Colleges (CCC) are "open to any student capable of benefitting from instruction" (Douglas, 2000).

remediation placement.[6] Moreover, CSU sets admissions and remediation policies at the system level, not at the campus level.

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 40 percent of high school students complete the set of courses that make them CSU-eligible.[7] 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).

[Insert Figure 1 about here]

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),[8] 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

---

[6] Given the diversity of the 23-campus CSU system (see Table A1), it is difficult to identify specific "peer institutions," but using CSU system average SAT scores (~1000) and admissions rate (60 percent) these 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 (based on National Center for Education Statistics, IPEDS peer institution comparisons: https://nces.ed.gov/ipeds/datacenter/).

[7] See University Eligibility Study: http://www.opr.ca.gov/docs/RTI_Eligibility_Report_071417_FINALtoOPR.pdf.

[8] 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.

incredibly low. Only 13 percent of California's 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 who 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 twelfth grade in both math and ELA. Finally, students who either do not take or take but do not pass these courses in the twelfth 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 who pass these tests are considered proficient or college ready, while students who 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.[9]

In 2011, the CSU enacted a policy known as *Early Start*, requiring incoming students who do not demonstrate readiness for college-level math 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

[9] Prior to 2011, the English Placement Test threshold for exemption from remediation was higher, at 151.

achieved proficiency in English and/or mathematics on or before the end of their first year of enrollment at a CSU campus.[10]

The goals of Early Start are to better prepare students in math and English *before* their first semester at CSU.[11] Specifically, Early Start requires all incoming students who do not meet the threshold on the Entry Level Math (ELM) and English Placement Test (EPT) proficiency requirements to take a designated developmental education course in the summer before their freshman year.[12] 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 to meet 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.) 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 who were 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 in need) for

[10] See https://www.calstate.edu/eo/EO-1048.pdf.
[11] See http://www.csusuccess.org/earlystart/early-start-faq for additional information.
[12] 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.

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.

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 to 2011), and four cohorts after (2012 to 2015).[13] 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.[14] To measure whether Early Start led to its intended goals, we investigate several outcome measures.

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 (in non-remedial, credit-bearing courses). To determine if Early Start changed the number and types of

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

[14] 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. We conduct additional subgroup analyses by the subset of students who were affected by the English policy change (EPT scores between 147 and 151); results remain consistent and are available upon request from the authors.

courses students enrolled in, we examine both whether a student attempted 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.[15] 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 3.33 high school GPA, and with average composite SAT scores at 1001. Importantly, we note that 54 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 who took the ELM, about one-third pass. For our analytical sample, we focus exclusively on ELM test takers (who comprise 54 percent of incoming first-time freshmen) since they are the students who are potentially subject to additional remediation upon entry into CSU.

[Insert Table 1 about here]

To investigate the impacts of the Early Start policy (i.e., the impacts of requiring summer prerequisite coursework), we employ two strategies. The first strategy, a

[15] We do not observe individual course titles, just course types (remedial, lower versus upper division, and total units), so these course type measures can include courses from any discipline, and may or may not include the mathematics general education requirement.

difference-in-differences approach, 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 years) to the outcomes of students ineligible to participate because the program was not yet available (in 2009 to 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:

$$Y_{itc} = \gamma 0 + \gamma 1 RemedNeed_i + \gamma 2 ESPolicy_t * RemedNeed_i + \beta X_i + \lambda_c + \delta_t + e_{itc} \quad (1)$$

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 $RemedNeed_i$, $\gamma 2$, 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 groups would not change. If this identifying assumption holds, then $\gamma 2$ 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). Including campus fixed effects also controls for time invariant differences in student outcomes across the 23 campuses of the CSU system. Additional specifications estimate separate coefficients of $\gamma 2$ for students who are in the first, second, or third cohorts exposed to the policy reform (post years one, two, or three). 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 3.12, while it is 3.25 among those who pass; however, both groups witnessed a small increase in high school GPA (.05) between the pre- and post-Early Start period. 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 about here]

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 (Hahn, Todd, & Van der Klaauw, 2001; Imbens & Angrist, 1994). 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 *k*. The quantity $\tau_{FRD} = (\tau^{Y}_{RF} / \tau_{FS})$ is estimated, where the numerator τ^{Y}_{RF} is the reduced-form effect of scoring below *k* on *Y*, while the denominator τ_{FS} is the first stage effect of scoring below *k* on Early Start enrollment *R*. Assuming that unobservable correlates of *Y* and *R* are continuous through *k*, τ^{Y}_{RF} and τ_{FS} are identified and equivalent to the discontinuities in *Y* and *R* at *k*, respectively; these continuity assumptions imply:

$$\tau^{Y}_{RF} = \lim_{S \to k-} E(Y|S) - \lim_{S \to k+} E(Y|S) \text{ and}$$

$$\tau_{FS} = \lim_{S \to k-} E(R|S) - \lim_{S \to k+} E(R|S).$$

Additional assumptions are needed for the estimation of τ_{FRD} (Imbens & Angrist, 1994). First, scoring below *k* 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 *k* 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 *k* with an increase in enrollment for students who scored above the cutoff *k*; this condition holds since almost no students enroll in Early Start if they scored above *k*. 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 & Angrist, 1994).

For estimation, we use the robust local linear estimator proposed by Calonico, Cattaneo, and Titiunik (2014) across four different bandwidths (+/- 8, 12, 16, and 20). That is, for our first stage, we estimate the specification as follows:

$$EarlyStart_{itc} = \alpha + f(S_i) + \varphi FailedELM_i + g(S_i)*FailedELM_i + \beta X_i + \lambda_c + \delta_t + e_{itc}. \quad (2)$$

where $EarlyStart_{itc}$ is an indicator for whether student *i* who entered in year *t* on campus *c* enrolled in Early Start, and $FailedELM_i$ is an indicator of whether the student scored below the required cutoff *k*. Our specifications include controls for student gender, race, high school GPA, SAT scores, campus fixed effects, and year fixed effects. Though these controls are not necessary for identification with the FRD, they should boost statistical power. Using predicted values from the first stage, we then estimate:

$$Y_{itc} = \alpha + j(S_i) + \phi hat(EarlyStart_{itc}) + k(S_i)*FailedELM_i + \beta X_i + \lambda_c + \delta_t + e_{itc}. \quad (3)$$

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 from 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.[16]

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 A2 in the Appendix.) 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 A1, which plots the distribution of observations by the running variable with 95 percent 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

[16] 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 on 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.

distribution of test scores just above the cutoff, which otherwise would suggest some type of manipulation out of the Early Start requirement.

RESULTS

We begin by describing the trends in student inputs throughout the study period. Figure 2 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 status from the additional ELM placement test; entering students who are required to take the ELM and pass the remediation threshold; and entering students who take the ELM and fail the remediation threshold and, thus, are required to enroll in subsequent developmental coursework. Not surprisingly, students who enter CSU exempt from any additional mathematics remedial placement exam have 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. First year GPA at CSU along with second year persistence rates have also been steadily rising over the past decade, but not quite as clearly for students who enter CSU less prepared (as measured by the ELM).

[Insert Figure 2 about here]

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 3 presents findings from the difference-in-differences models. Column 1 includes models with campus and year fixed effects; in column 2, we add demographic controls (race and gender); and in column 3, we also include controls for prior academic achievement (high school GPA and SAT scores). For all specifications, standard errors are clustered at the campus level using wild cluster bootstrapping (Cameron, Gelbach, & Miller, 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 to 0.04 of a grade point average and only statistically significant at the 10 percent level for the fully specified model. Early Start did not have an effect on the likelihood that 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, though the effect is attenuated with the addition of the high school inputs in column 3. 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 find a positive statistically significant effect of Early Start on total units (a magnitude of approximately three units). 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 4, 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 just one year. Persistence rates are higher each year (relative to 2009), but based on results presented in Table 3, 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 3 & 4 about here]

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 3 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 who 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 3 about here]

The remaining five graphs of Figure 3 can be interpreted as unconditional reduced form estimates of $\tau^{Y}_{RF,}$ 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 3, 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 two 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 5 presents regression results from our FRD. Each panel presents 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 below the cutoff on the ELM test were roughly 30 percentage points more likely to enroll in Early Start than students who scored just barely above 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, and statistically significant at the 10 percent level for three of the four considered bandwidths.[17]

[Insert Table 5 about here]

DISCUSSION/CONCLUSION

[17] As an additional robustness check, we also parametrically implement the fuzzy regression discontinuity design using a standard instrumental variables (IV) regression with cubic terms of test scores on either side of the cutoff. These results, presented across our initial bandwidths and a bandwidth of four (the optimal bandwidth as calculated by Imbens & Kalyanaraman, 2012), are presented in Table A3. These results are largely consistent with those in Table 5.

Today, more than ever, colleges and universities are focused on improving persistence and degree completion rates. A well-established tension exists in collegiate remediation: on the one hand, remediation may be a necessary tool for skill 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 (Bettinger, Boatman, & Long, 2013; Bettinger & Long, 2007; Boatman & Long, 2018; Deil-Amen & Rosenbaum, 2002; Scott-Clayton & Rodriguez, 2015). 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 as a system-wide policy 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). Specifically, our two strategies allow us to investigate: first, whether students who were required to enroll in additional developmental coursework had better outcomes when that coursework was in the summer prior to their first year of enrollment. The answer to that question is no, save for a slight increase in upper division enrollment, suggesting that the “early start” in developmental coursework may have, at best, altered students’ course choices, albeit to a small degree. Second, we are able to compare the outcomes of students required to take summer developmental coursework on the basis of just missing the exemption cutoff with those who just met the exemption cutoff, and find negative effects on persistence. Thus, overall, we conclude that this policy did not contribute to improving student performance in the first year or for subsequent persistence rates; and that the summer remediation timing was still not preferable to no remediation when comparing students near the exemption cutoff. This finding is consistent with several other prior studies exploring remediation effects using similar analytic strategies in several institutional contexts (Boatman & Long, 2018; Calcagno & Long, 2008; Martorell & McFarlin, 2011; Scott-Clayton & Rodriguez, 2015).

Of course, regression discontinuity techniques do not allow for generalization beyond the range of the cutoff, and, in this case and many others, we are often interested in the lowest performing students (i.e., those who may have the greatest developmental needs). In supplementary analyses we conduct using DID for a reduced sample, comparing only the lowest performers (based on ELM test score), we find that moving to summer remediation—instead of fall—had a bigger positive magnitude on total units accumulated, and, importantly, a positive effect on persistence to the third year of college, which is more promising (see Table A4).

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. 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. Although we do not find a statistically significant effect on grades, coefficients on GPA were consistently negative across models; and, since students often care a lot about their grades, we'd need additional data to better understand course selection in order to more fully tease this out. Nevertheless, if indeed summer remedial enrollment leads to a boost of human capital accumulation (as measured by more units), but at the small expense of GPA, this may be worth the expense.

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 4 plots the difference-in-differences results for four outcomes at the campus level (the campuses are not ordered in any 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., one or three units).

[Insert Figure 4 about here]

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 (Education Commission of the States, 2018). 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, respectively), and by making all remedial coursework co-requisite (effective for students entering CSU 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 have implemented these changes to remediation. Nevertheless, Early Start stands to play an even greater role as the primary form of pre-collegiate developmental offering for CSU students.

Our analysis of Early Start suggests that summer remediation is no better than fall remediation, and—consistent with prior work—worse than no remediation (when comparing students near the remediation exemption cutoff). Currently, only students demonstrating very low proficiency (based on high school grades and 11th-grade test scores[18]) will be required to take Early Start, and only one subject must be taken in the summer. Moreover, many of the Early Start courses across campuses have been redesigned

[18] The CSU has discontinued the use of the ELM as part of Executive Order 1110.

to offer credit-bearing courses in writing and quantitative reasoning.[19] These changes, along with the broader changes to CSU's remediation policies, may allow many students to bypass the remedial track that slowed them down, and may better target developmental coursework for those students who really need it. The longer-term effects of these changes, however, remain to be settled.

[19] See https://www2.calstate.edu/csu-system/why-the-csu-matters/graduationinitiative-2025/files/academic-preparation-faq.pdf.

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., Boatman, A., & Long, B. T. (2013). Student supports: Developmental education and other academic programs. The Future of Children, 23, 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, 29–58.

Calcagno, J. C., & Long, B. T. (2008). 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., Gelbach, J., & Miller, D. (2008). Bootstrap-based improvements for inference with clustered errors. The Review of Economics and Statistics, 90, 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, R., & Rosenbaum, J. E. (2002). The unintended consequences of stigma-free remediation. Sociology of Education, 75, 249–268.

Douglas, J. A. (2000). The California idea and American higher education. Stanford, CA: Stanford University Press.

Education Commission of the States. (2018). 50-State comparison: Developmental education policies. See: https://www.ecs.org/50-state-comparison-developmental-education-policies/.

Fain, P. (2013, September 13). Going to the root of the problem. Insider Higher Ed. Available at 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, 726–748.

Imbens, G., & Angrist, J. (1994). Identification and estimation of local average treatment effects. Econometrica, 62, 467–475.

Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. The Review of Economic Studies, 79, 933–959.

Jackson, J., & Cook, K. (2016, May). Improving college graduation rates: A closer look at California State University. Sacramento, CA: Public Policy Institute of California. Available at 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, 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, 436–454.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. Journal of Econometrics, 142, 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, 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, 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, 85–109.

Sparks, D., & Malkus, N. (2013). First-year undergraduate remedial coursetaking: 1999-2000, 2003-04, 2007-08. Statistics in brief. NCES 2013-013. Washington, DC: National Center for Education Statistics.

U.S. Department of Education, National Center for Education Statistics. (2018). Integrated Postsecondary Education Data System (IPEDS), Winter 2017–18, Graduation Rates component. See Digest of Education Statistics 2018,Table 326.10.

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.

TABLES AND FIGURES

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

N=403,268 students	Mean	S.D.
Male	0.43	
Hispanic	0.41	
Black	0.05	
Asian	0.17	
White	0.27	
High school GPA	3.33	0.43
SAT composite	1001.15	172.15
Took EPT test	0.50	
Took ELM test	0.54	
Passed conditional on taking ELM test	0.33	
Outcome: Math exempt prior to fall	0.75	
Outcome: Total units at CSU	76.31	59.32
Outcome: Took an upper-division course in year 1	0.11	
Outcome: First term GPA	2.83	0.88
Outcome: Upper-division units in year 1	0.57	2.07
Outcome: Total units at CSU	76.31	59.32
Outcome: Persist to year 2	0.85	
Outcome: Persist to year 3	0.77	

Notes: Summary statistics for the population of fall first-time Freshmen in the California State University system. Sample size drops for first term GPA outcome to 386,974 students since not all students enrolled in the fall receive a term GPA (e.g., students may drop out, or take courses without letter grading). Sample sizes for persistence rate to year 2 and to year 3 drops to 338,286 and 274,611 students, respectively, 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)				
	Pre-ES (2009-2011)		Post-ES (2012-2015)		Pre-ES (2009-2011)		Post-ES (2012-2015)		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	DiD [*p*-value]
Mathematics (ELM) Test Takers	(*n=60,803*)		(*n=86,107*)		(*n=31,192*)		(*n=40,778*)		
Male	0.31		0.31		0.46		0.45		0.01 [0.092]*
White	0.19		0.14		0.28		0.23		0.01 [0.269]
Hispanic	0.49		0.60		0.38		0.47		0.01 [0.436]
Asian	0.12		0.11		0.19		0.18		0.00 [0.407]
High school GPA	3.12	0.37	3.17	0.36	3.25	0.40	3.30	0.39	0.01 [0.200]
SAT Composite	836.95	115.93	848.88	110.68	992.42	96.92	977.96	91.51	1.676 [0.455]
ELM test score	-13.90	8.28	-14.76	8.48	7.74	6.90	6.87	6.48	0.10 [0.575]
Outcome: Math exempt prior to fall	0.08		0.49		1.00		1.00		0.42 [0.000]***
Outcome: First term GPA	2.64	0.97	2.70	0.93	2.73	0.91	2.82	0.84	-0.03 [0.156]
Outcome: Upper-div. units in Y1	0.53	1.90	0.42	2.05	0.71	2.36	0.54	2.24	0.06 [0.020]**
Outcome: Took upper-div. course in Y1	0.13		0.10		0.16		0.13		-0.00 [0.936]
Outcome: Total units at CSU	92.52	63.15	44.15	37.86	104.13	61.23	55.06	40.32	0.13 [0.895]
Outcome: Persist to Y2	0.80		0.81		0.86		0.86		-0.00 [0.950]
Outcome: Persist to Y3	0.71		0.71		0.78		0.77		0.01 [0.154]

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. Main difference-in-differences results.

	(1)	(2)	(3)
Outcome: Math satisfied prior to Fall			
First year 2012+ X Below test cutoff	0.466***	0.466***	0.465***
	(0.022)	(0.022)	(0.022)
Observations	218880	218880	218880
Outcome: First term GPA			
First year 2012+ X Below test cutoff	-0.034	-0.035	-0.042*
	(0.024)	(0.024)	(0.025)
Observations	208804	208804	208804
Outcome: Upper division units in Y1			
First year 2012+ X Below test cutoff	0.057**	0.058**	0.036
	(0.029)	(0.029)	(0.027)
Observations	218880	218880	218880
Outcome: Took an upper division course			
First year 2012+ X Below test cutoff	0.002	0.002	0.001
	(0.004)	(0.004)	(0.004)
Observations	218880	218880	218880
Outcome: Total units			
First year 2012+ X Below test cutoff	2.700***	2.616***	2.456***
	(0.887)	(0.871)	(0.840)
Observations	218880	218880	218880
Outcome: Persist to Y2			
First year 2012+ X Below test cutoff	-0.000	-0.000	-0.001
	(0.005)	(0.005)	(0.005)
Observations	187439	187439	187439
Outcome: Persist to Y3			
First year 2012+ X Below test cutoff	0.005	0.005	0.004
	(0.004)	(0.004)	(0.004)
Observations	154217	154217	154217
Gender/Race Controls		X	X
HS GPA/SAT Controls			X
Campus/Entry Year Fixed Effects	X	X	X

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

Table 4. Difference-in-differences results by post Early Start year interactions.

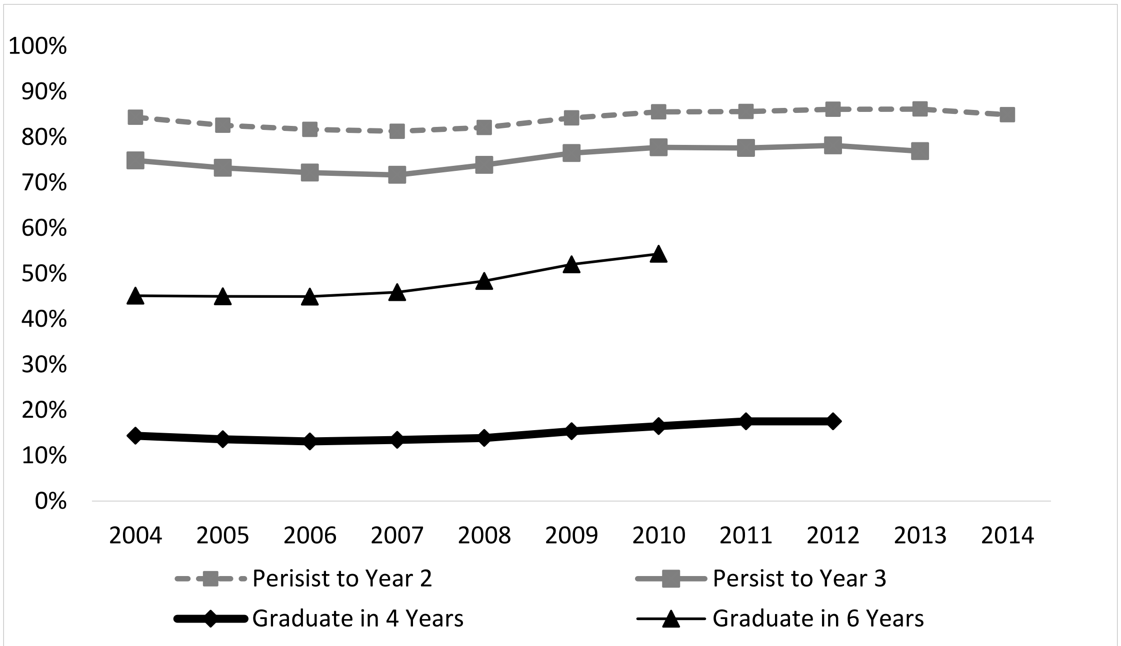
	First term GPA	Upper div. units	Persist to year 2	Persist to year 3
First year 2009 X Below test cutoff	-0.008	-0.047	-0.017**	-0.020**
	(0.045)	(0.046)	(0.008)	(0.009)
First year 2010 X Below test cutoff	0.006	-0.064**	0.003	0.008
	(0.018)	(0.027)	(0.007)	(0.008)
First year 2012 X Below test cutoff	-0.029	-0.039	-0.011	-0.002
	(0.019)	(0.034)	(0.007)	(0.007)
First year 2013 X Below test cutoff	-0.022	-0.017	-0.005	0.000
	(0.027)	(0.035)	(0.007)	(0.006)
First year 2014 X Below test cutoff	-0.047	0.056*	-0.001	
	(0.032)	(0.033)	(0.008)	
First year 2015 X Below test cutoff	-0.079**	-0.008		
	(0.036)	(0.049)		
Observations	208804	218880	187439	154217
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 are not reported. Omitted interaction is “First year 2011 X Below test cutoff.” Standard errors are clustered at the campus level using wild cluster bootstrapping. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 5. Fuzzy regression discontinuity results—student outcomes (local linear).

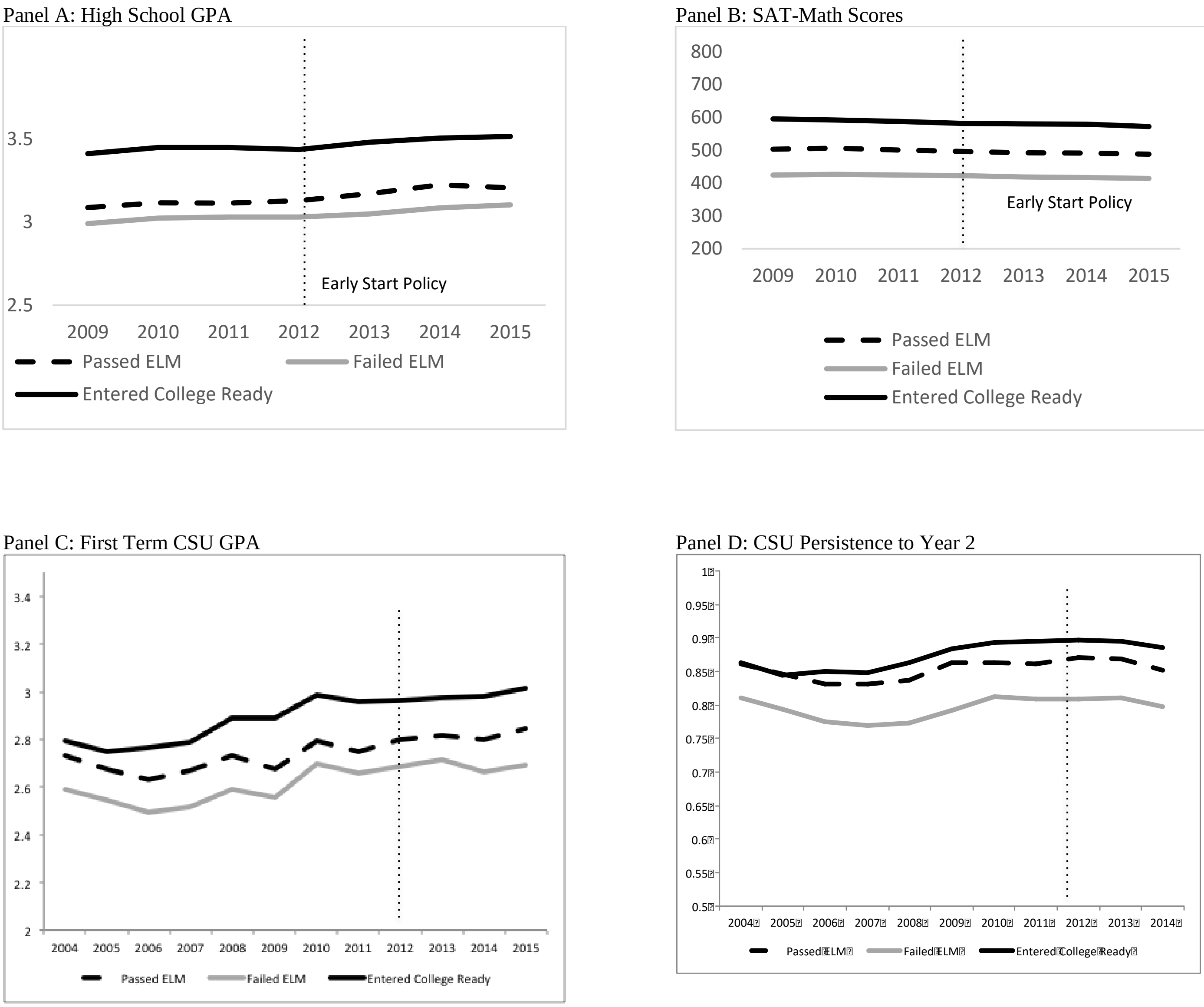
	(1)	(2)	(3)	(4)
First stage: Early Start enrollment				
First-stage discontinuity (failed ELM)	0.300***	0.295***	0.294***	0.293***
	(0.017)	(0.019)	(0.021)	(0.025)
Observations	13006	11035	8807	6338
Mean above cutoff	0.000	0.000	0.001	0.001
Outcome: First term GPA				
Enrolled in Early Start	0.001	-0.006	0.013	-0.100
	(0.107)	(0.115)	(0.117)	(0.135)
Observations	12399	10526	8411	6068
Mean above cutoff	2.868	2.858	2.851	2.852
Outcome: Upper div. units in Y1				
Enrolled in Early Start	0.283	0.290	-0.037	-0.213
	(0.260)	(0.295)	(0.328)	(0.508)
Observations	13006	11035	8807	6338
Mean above cutoff	0.493	0.476	0.475	0.474
Outcome: Took an upper div. course				
Enrolled in Early Start	-0.010	-0.011	-0.049	-0.085
	(0.029)	(0.033)	(0.038)	(0.063)
Observations	13006	11035	8807	6338
Mean above cutoff	0.088	0.087	0.086	0.087
Outcome: Total units				
Enrolled in Early Start	-8.643***	-9.400***	-10.627***	-10.221***
	(2.150)	(2.571)	(2.959)	(3.679)
Observations	13006	11035	8807	6338
Mean above cutoff	31.850	31.727	31.492	31.291
Outcome: Persist to Y2				
Enrolled in Early Start	-0.041*	-0.049*	-0.049	-0.061*
	(0.024)	(0.029)	(0.031)	(0.033)
Observations	6440	5494	4440	3156
Mean above cutoff	0.864	0.862	0.858	0.859
Bandwidth	+/- 20	+/- 16	+/- 12	+/- 8
Gender/Race Controls	X	X	X	X
HS GPA/SAT controls	X	X	X	X
Campus FE	X	X	X	X

Notes: Each cell in the first panel reports the first-stage estimated effect of failing the ELM test on Early Start enrollment from a local linear (local polynomial of degree one) regression discontinuity following Calonico et al. (2014), while the remaining panels report estimates from a local linear fuzzy regression discontinuity, where Early Start enrollment is instrumented for by the test score cutoff. Standard errors are clustered at the campus level. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.



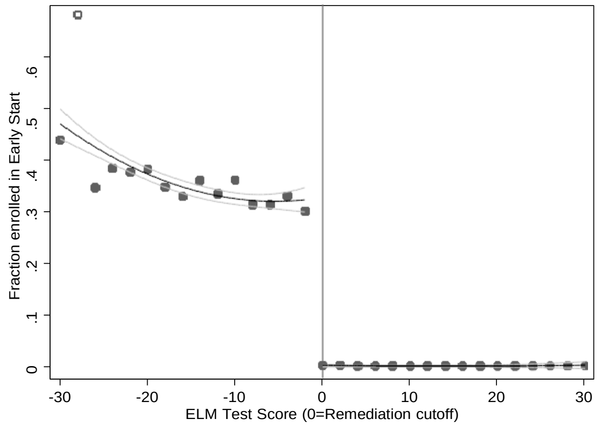
Notes: Authorsistence and Co.
Source: CSU Chancellor's Office Analytics Studies.

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

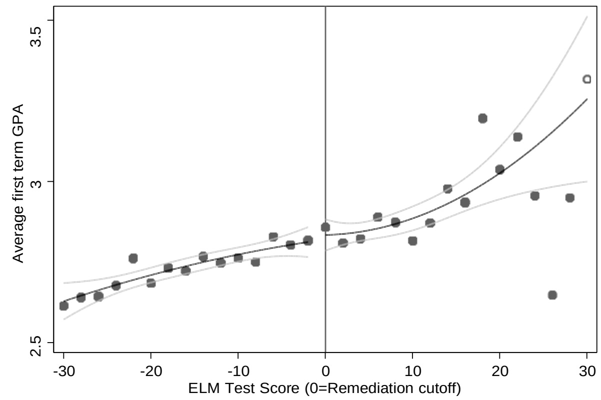


Notes: Authors' calculations.
Source: CSU Chancellor's Office Analytics Studies.

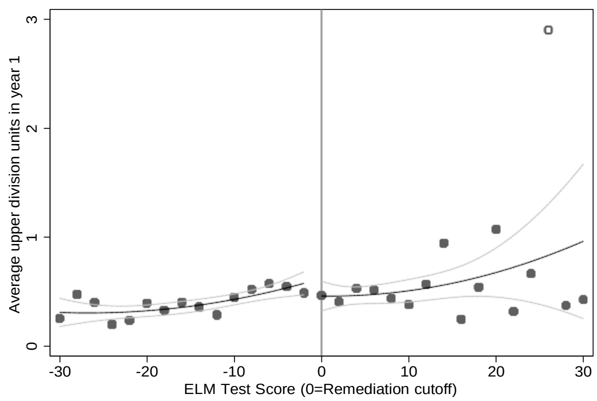
Figure 2. High School Characteristics and CSU Outcomes Over Time by CSU Mathematics Remediation Status at Entry.



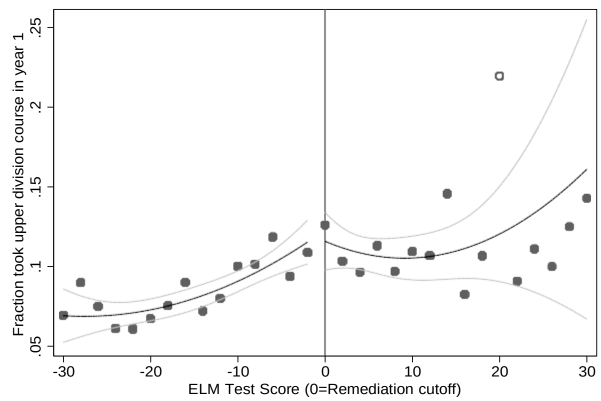
(a) Enrollment in Early Start



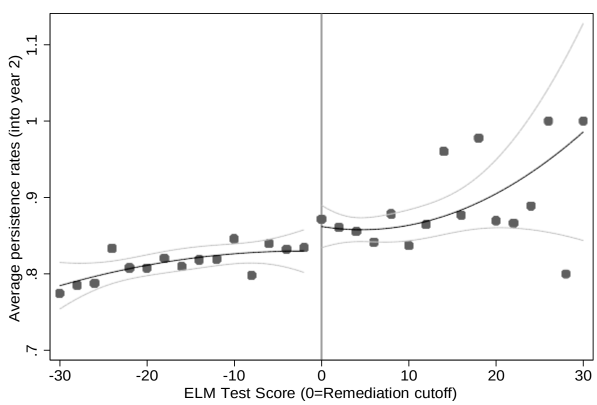
(b) First term GPA



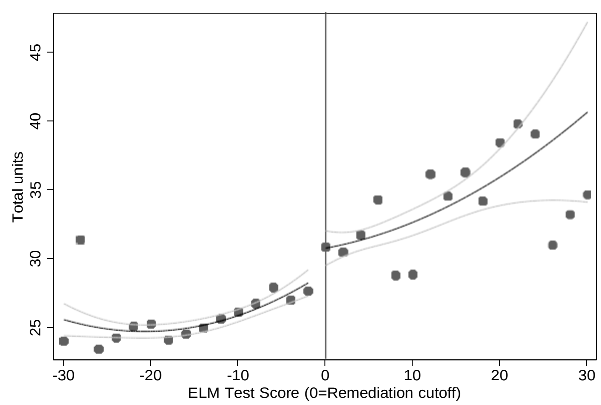
(c) Upper division units in year 1



(d) Enrolled in upper division course in year 1



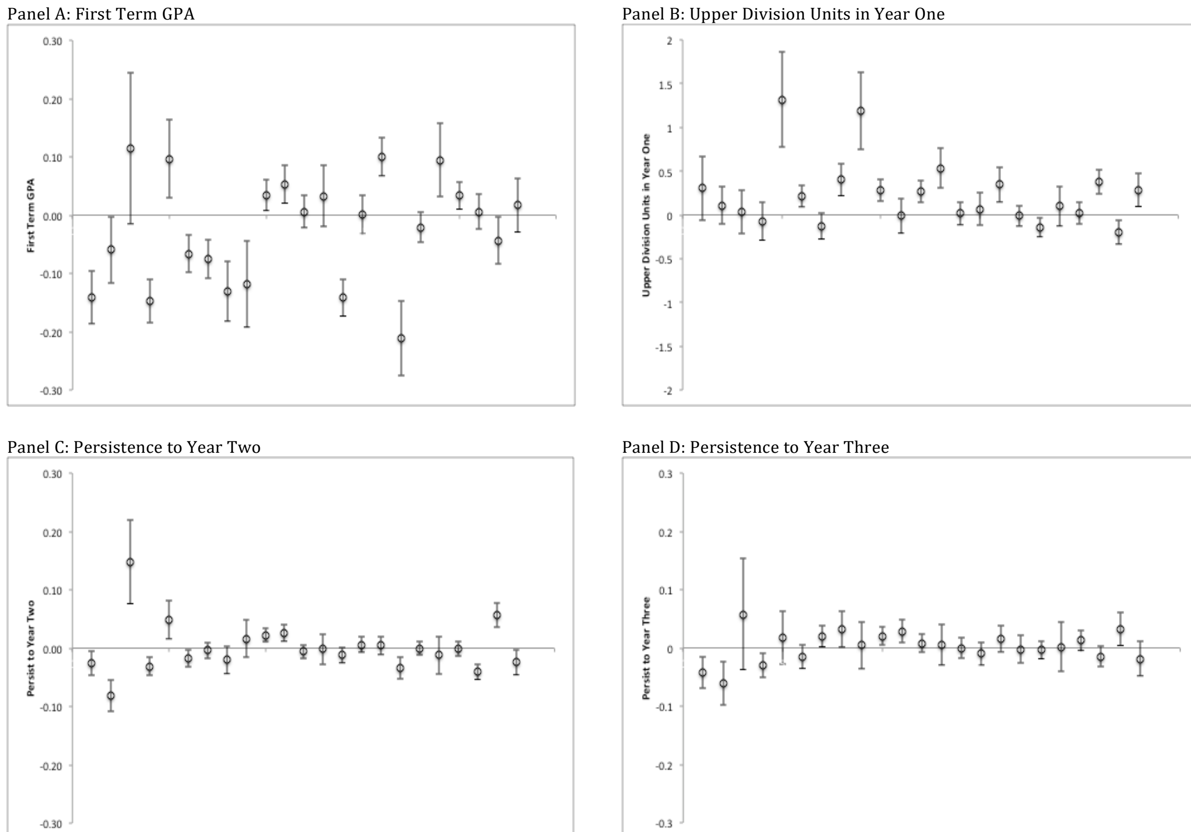
(e) Persist to year 2



(f) Total units

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

Figure 3. Averages of Outcomes by ELM Test Score.



Notes: The panels show the fully specified model, including student controls and year fixed effects, estimated by campus.

Figure 4. Difference-in-Differences Results by Campus.

APPENDIX

Table A1. Characteristics of California State University campuses (2016).

	Admit Rate	# First Time Frosh	SAT-Reading/Writing 25th %ile	SAT-Reading/Writing 75th %ile	SAT-Math 25th %ile	SAT-Math 75th %ile
Bakersfield	73%	1,357	440	540	440	540
Channel Islands	75%	1,010	480	600	480	570
Chico	67%	2,762	500	590	490	580
Dominguez Hills	51%	1,299	420	500	380	490
East Bay	71%	1,596	450	440	550	540
Fresno	53%	3,302	460	560	450	550
Fullerton	49%	4,426	500	600	510	590
Humboldt	77%	1,295	490	590	470	570
Long Beach	32%	4,253	510	610	510	620
Los Angeles	66%	3,830	450	540	440	540
Maritime Academy	73%	241	530	610	540	640
Monterey Bay	35%	802	490	590	480	580
Northridge	49%	4,499	460	570	450	550
Pomona	59%	4,204	500	610	510	620
Sacramento	73%	3,760	470	570	470	570
San Bernardino	59%	2,791	460	550	450	540
San Diego	31%	5,077	550	640	540	650
San Francisco	70%	3,642	480	580	470	570
San Jose	55%	3,208	500	600	500	610
San Luis Obispo	27%	4,279	600	680	600	700
San Marcos	52%	2,152	480	570	470	560
Sonoma	78%	1,806	500	590	480	580
Stanislaus	76%	1,389	460	560	450	540

Sources:
CSU Analytic Studies (https://www.calstate.edu/as/stat_reports/2016-2017/apps_f2016_res.htm),
IPEDS, National Center for Education Statistics.

Table A2. Fuzzy regression discontinuity results—balance test (local linear).

	(1)	(2)	(3)	(4)
Outcome: Male student				
Passed ELM (skipped Early Start)	0.020	0.032	0.020	0.008
	(0.088)	(0.097)	(0.107)	(0.145)
Observations	13,006	11,035	8,807	6,338
Outcome: URM student				
Enrolled in Early Start	0.091	0.084	0.063	0.006
	(0.148)	(0.152)	(0.153)	(0.169)
Observations	13,006	11,035	8,807	6,338
Outcome: White student				
Enrolled in Early Start	-0.075	-0.052	-0.036	0.017
	(0.139)	(0.144)	(0.147)	(0.163)
Observations	13,006	11,035	8,807	6,338
Outcome: High school GPA				
Enrolled in Early Start	-0.042	-0.040	-0.083	-0.112
	(0.125)	(0.129)	(0.129)	(0.140)
Observations	13,006	11,035	8,807	6,338
Outcome: SAT composite score				
Enrolled in Early Start	-23.791	-31.599	-42.352	-48.734
	(37.432)	(40.036)	(43.592)	(54.812)
Observations below cutoff	8,773	6,946	4,978	3,175
Observations above cutoff	4,233	4,089	3,829	3,163
Bandwidth	+/- 20	+/- 16	+/- 12	+/- 8

Notes: Each cell reports an estimate from a local linear (local polynomial of degree one) fuzzy regression discontinuity following Calonico et al. (2014), where Early Start enrollment is instrumented for by the test score cutoff. Standard errors are clustered at the campus level. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A3. Additional regression discontinuity results—parametric estimation.

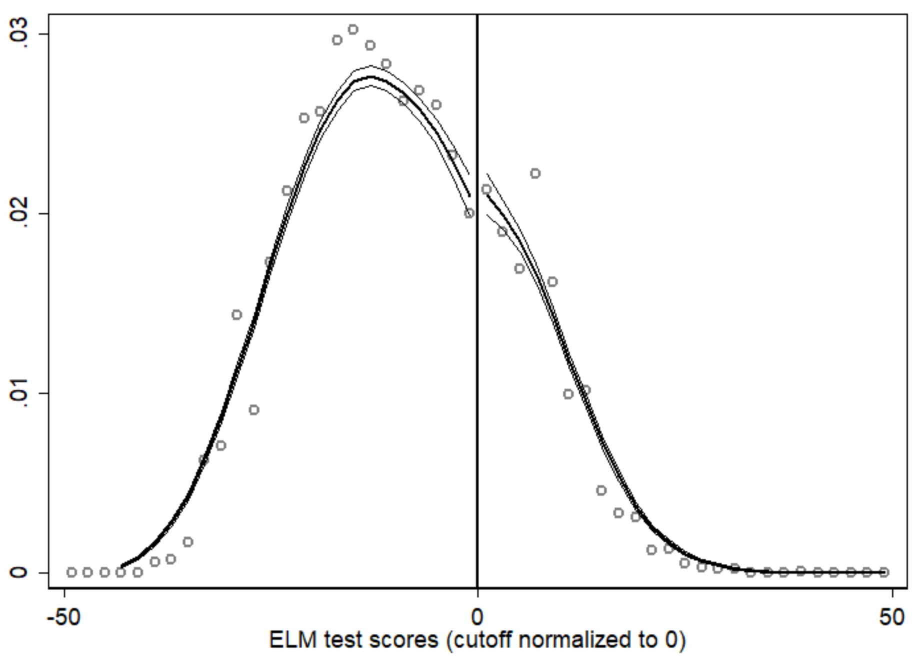
	(1)	(2)	(3)	(4)	(5)
First stage: Early Start enrollment					
First-stage discontinuity (failed ELM)	0.301***	0.304***	0.296***	0.305***	0.269***
	(0.020)	(0.021)	(0.026)	(0.024)	(0.027)
Observations	13006	11035	8807	6338	3322
Outcome: First term GPA					
Enrolled in Early Start	0.055	0.002	0.013	0.053	-0.032
	(0.105)	(0.110)	(0.119)	(0.144)	(0.257)
Observations	12399	10526	8411	6068	3194
Outcome: Upper div. units in Y1					
Enrolled in Early Start	0.264	0.315	0.447	-0.137	-0.134
	(0.258)	(0.334)	(0.398)	(0.392)	(0.735)
Observations	13006	11035	8807	6338	3322
Outcome: Took an upper div. course					
Enrolled in Early Start	0.010	-0.004	0.008	-0.065*	-0.047
	(0.027)	(0.034)	(0.041)	(0.037)	(0.103)
Observations	13006	11035	8807	6338	3322
Outcome: Total units					
Enrolled in Early Start	-8.625***	-7.572***	-8.925***	-11.090***	-10.181**
	(1.771)	(2.086)	(3.290)	(3.320)	(4.826)
Observations	13006	11035	8807	6338	3322
Outcome: Persist to Y2					
Enrolled in Early Start	-0.025	-0.033	-0.050	-0.024	-0.055
	(0.024)	(0.024)	(0.032)	(0.034)	(0.065)
N	6440	5494	4440	3156	1644
Bandwidth	+/- 20	+/- 16	+/- 12	+/- 8	+/- 4
Gender/Race Controls	X	X	X	X	X
HS GPA/SAT/Major controls	X	X	X	X	X
Campus FE	X	X	X	X	X

Notes: Each cell in the first panel reports the first-stage estimated effect of failing the ELM test on Early Start enrollment from a regression discontinuity, with separate cubic terms fitted on either side of the threshold. The remaining panels report estimates from a fuzzy regression discontinuity, where Early Start enrollment is instrumented for by the test score cutoff, and with separate cubic terms fitted on either side of the threshold. Standard errors are clustered at the campus level. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A4. Difference-in-differences results with only lowest performing ELM scores in treatment group.

	(1)	(2)	(3)
Outcome: Math satisfied prior to Fall			
First year 2012+ X Below test cutoff	0.483***	0.482***	0.481***
	(0.020)	(0.020)	(0.020)
Observations	147851	147851	147851
Outcome: First term GPA			
First year 2012+ X Below test cutoff	-0.021	-0.026	-0.035
	(0.029)	(0.030)	(0.031)
Observations	141058	141058	141058
Outcome: Upper division units in Y1			
First year 2012+ X Below test cutoff	0.070*	0.069*	0.047
	(0.036)	(0.036)	(0.034)
Observations	147851	147851	147851
Outcome: Took an upper division course			
First year 2012+ X Below test cutoff	0.003	0.003	0.002
	(0.005)	(0.005)	(0.005)
Observations	147851	147851	147851
Outcome: Total units			
First year 2012+ X Below test cutoff	4.483***	4.344***	4.077***
	(1.312)	(1.267)	(1.253)
Observations	147851	147851	147851
Outcome: Persist to Y2			
First year 2012+ X Below test cutoff	0.008	0.008	0.007
	(0.006)	(0.006)	(0.006)
Observations	126571	126571	126571
Outcome: Persist to Y3			
First year 2012+ X Below test cutoff	0.013**	0.012**	0.011*
	(0.006)	(0.006)	(0.006)
Observations	104001	104001	104001
Gender/Race Controls		X	X
HS GPA/SAT Controls			X
Campus/Entry Year Fixed Effects	X	X	X

Notes: Each cell reports a difference-in-differences coefficient from a single regression. The treatment group includes those who scored below the median ELM score among those who failed the ELM test. Standard errors are clustered at the campus level using wild cluster bootstrapping. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent levels, respectively.



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

Figure A1. Density of ELM Test Scores with 95 percent McCrary Confidence Intervals.