Does Reclassifying English Learners Affect Their Postsecondary Outcomes?

Julian Betts[1], Laura Hill[2], Karen Bachofer[3], Joseph Hayes[2], Andrew Lee[2] and Andrew Zau[3]

This draft: September 30, 2020
Preliminary and incomplete. Please do not cite without permission.

Abstract

In California 38% of students are either current or former English Learners (ELs). A crucial decision in their educational trajectory is when to reclassify ELs. Upon reclassification, ELs cease to receive language supports, but have more opportunities to take the same courses as fluent English speakers. This paper uses regression discontinuity in California's second largest school district to ask: "Are ELs being reclassified at the right time?" In most cases the district was reclassifying students appropriately based on the California Standards Test, with no discontinuity for outcomes related to post-secondary outcomes. Some discontinuities emerged for reclassification based on the test of language proficiency, but further analysis is needed given that for that experiment we have found discontinuities in baseline characteristics.

1: Department of Economics, UC San Diego and National Bureau of Economic Research
2: Public Policy Institute of California
3: Department of Economics, UC San Diego

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through R305A170288 and by the W.T. Grant Foundation. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Introduction

English Learner (EL) status is meant to be temporary—when students demonstrate sufficient English language proficiency, their official designation changes to Reclassified Fluent English Proficient (RFEP). If ELs are reclassified too soon, their academic performance may falter because their lack of mastery in English may impede their understanding of course materials. But if ELs are reclassified too late, their academic progress may also falter. In this case the language supports the ELs are receiving could create an opportunity cost in the form of reduced opportunities to take the same courses as native English speakers. This paper uses student-level data secured thought special arrangements with one large urban California school districts to ask "Are ELs being reclassified at the right time?" In this paper, in order to understand if English Learner support services are being removed at the right time, we examine post-secondary outcomes that matters for labor market outcomes: do reclassified students fare better or worse when it comes to enrolling in, and graduating from, college?"

It is important to study reclassification because it touches upon the lives of so many students. In total, more than 40 percent of the students in California's public schools speak a language other than English at home. In the 2016–17 school year, 21 percent (or more than 1.3 million) of all students were English Learners. When students who were formerly English Learners are added in, the population of "ever ELs" expands to 38 percent of all K–12 students in the state. Nationwide, 10 percent of all K-12 students are English Learners.[1]

A second reason reclassification is important concerns the recent passage of the Every Student Succeeds Act (ESSA). ESSA now requires states to standardize reclassification policies across all of their school districts. In addition, California has recently introduced two new tests to assess the overall English language arts proficiency of all students and the language mastery of ELs specifically. We must understand the impact of reclassification policies that relied on the old tests in order to create a roadmap for developing the new reclassification policies.

Students entering K–12 schools in California are classified as English Learners if they speak a language other than English at home and score below a proficiency threshold on the California English Language Development Text (CELDT).[2] English Learners are meant to have English language development instruction, either in stand-alone English language development classes ("designated") or as part of regular instruction ("integrated"). EL students do receive core subject instruction, but English language development instruction may mean they have less of it. After they are reclassified, these students no longer receive English language development instruction and take core subjects without additional support.

In order to be reclassified, students must demonstrate English language proficiency on assessments administered only to EL students and demonstrate basic skills in English that are

[1] National Center for Education Statistics (2019). English Language Learners in Public Schools.
[2] Starting in the 2018-19 school year, students were assessed at school entry using the English Language Proficiency Assessment for California (ELPAC).

comparable to native English speakers. Individual school districts make the decision to reclassify EL students based on a number of criteria recommended by the state.

Student outcome data has long revealed that EL students do not fare as well as their non-EL peers. However, reclassified students perform much better than current EL students, and they sometimes perform better than students who do not speak a foreign language at home. Murillo and Lavadenz (2020) find that California EL students who have completed high school are much less likely to attend college than high school completers in general (47 percent versus 68 percent). A national study finds that only 19 percent of ELs that have completed high school enroll in four year colleges, whereas the same is true for 35 percent of non-native English speakers who are fluent in English (Kanno and Cromley 2015). Yet this does not imply any causal effect of reclassification. Without clarity about causality, it is difficult to design reclassification policies that optimize the duration of English language support. Our earlier research examines elementary and secondary school outcomes (Hill, Betts, Bachofer, Hayes, Lee, and Zau 2019). In this paper, we examine postsecondary outcomes, including enrollment, persistence, and graduation.

This paper uses student-level data in California's second-largest school district (San Diego Unified), to evaluate the effect of reclassification on academic outcomes for former ELs. [3]

We use a Regression Discontinuity design (RD) to college enrollment, persistence, and graduation for students just above and just below the cutpoints on various tests used to make reclassification decisions. Because San Diego Unified has used two different reclassification policies over the period we consider, we can not only identify any causal relationship between duration of English language support and post-secondary student outcomes, but also identify which of the reclassification standards comes closest to reclassifying a student at the appropriate time. If a policy reclassifies a student too soon, we detect a positive effect of an extra year of EL support on post-secondary outcomes. If a policy reclassifies too late, we detect a negative effect. And if we found no significant difference, it would suggest that students are being reclassified appropriately.

This is an important time to be thinking about how we measure when ELs are prepared to fully integrate into academic courses without English language support. The implementation of the common core standards and new English language development standards are underway, but it is uneven across California (Warren and Murphy 2014; McLaughlin, Glaab and Carrasco 2014). The new statewide standards test, the SBAC test, was administered for the first time in 2014-15, and the test results for the first year were particularly alarming for EL students (Hill and Ugo, 2016). The replacement for the CELDT, the English Language Proficiency Assessment for California (ELPAC), was fully implemented in the 2017-18 school year. Further, some research suggests it may be beneficial to use just one criterion, such as the ELPAC, to assess whether EL students are ready for reclassification (Umansky et al 2015 and Hill et al 2014).

[3] As is the case across the United States, as well as in California, the primary language of most ELs is Spanish. The districtincludes elementary, middle, and high schools, newcomer schools, and teachers with a variety of credentials. Sixty-one percent of SDUSD students are low income, and SDUSD's students' race/ethnic distribution is close to that of the state's student population. Overall, SDUSD has a very diverse student populations and represent the growing heterogeneity of students in the state and the nation.

Background on English Learner Reclassification Policies

We focus on EL students and reclassification decisions made in SDUSD between 2004 and 2012.[4] In California, the state issues guidance about reclassification policy, and district staff (EL program administrators and teachers) may modify it and implement reclassification decisions.[5] The policies and practices for identifying EL students, assigning them to instructional programs, and reclassifying them are clearly articulated in SDUSD's *Master Plan for English Learners*. Although policies vary somewhat between the districts, the policies are guided by state law, the California Department of Education, and the State Board of Education.

California's reclassification guidelines require the use of four criteria (California Education Code, Section 313(f)). During the period we study, the reclassification guidelines included:

1. The California English Language Development Test (CELDT), with recommended overall and subtest scores;
2. A test of basic skills in English, with a recommendation of the California Standards Test (CST) of English Language Arts (ELA) and a minimum score on the test;
3. Teacher evaluation; and
4. Parent consultation.

School districts in California, as in many other states (cited in Kim and Herman, 2012), are allowed to determine their own reclassification policies as long as they follow minimum suggested guidelines issued by the state. However, most states (30) in the pre-ESSA era relied only on an English Language proficiency assessment (Linquanti and Cook 2015). In California before January 2019, districts decided how to implement the four state reclassification criteria by setting local policy, but since January 2019, the English proficiency criteria have been standardized statewide.[6] The policies and practices for identifying EL students, assigning them to instructional programs, and reclassifying them are articulated in SDUSD's *Master Plan for English Learners*.[7]

A 2013 survey found most California school districts had developed more rigorous reclassification standards than those recommended by the state (Hill, Weston, and Hayes, 2014). For example, many districts required higher cut scores on the CST ELA or on the CELDT, required the CST math as another measure of basic skills, or required course marks for teacher evaluation. Results from the survey suggest parental consultation was of limited importance.

Like most other California districts, SDUSD has more rigorous standards than the minimum state guidelines (*Master Plan,* SDUSD 2009).

SDUSD has had its English language instructional programs in place for well over a decade, and the policies for determining reclassification in both districts have remained the same since 2006 through the final year in which we study reclassification, 2014. Different policies were in place in earlier years. For example, SDUSD raised its basic skills reclassification criterion for students

4 Through 2014 in SDUSD.

[5] Decision-making and implementation are the responsibility of the Office of Language Acquisition in SDUSD and the Multilingual and Multicultural Education Department in LAUSD.

[6] CDE's current guidance, as of January 2019, is found in the Updated Reclassification Guidance for 2018-19. The new guidance specifies that an ELPAC Overall Performance Level (PL) 4 is required for reclassification. In order to assess whether an EL meets the ELA basic skills requirement, districts may use either a local assessment or the Smarter Balanced Summative Assessment (grades 3-8, 11).

[7] Decision-making and implementation are the responsibility of the Office of Language Acquisition in SDUSD.

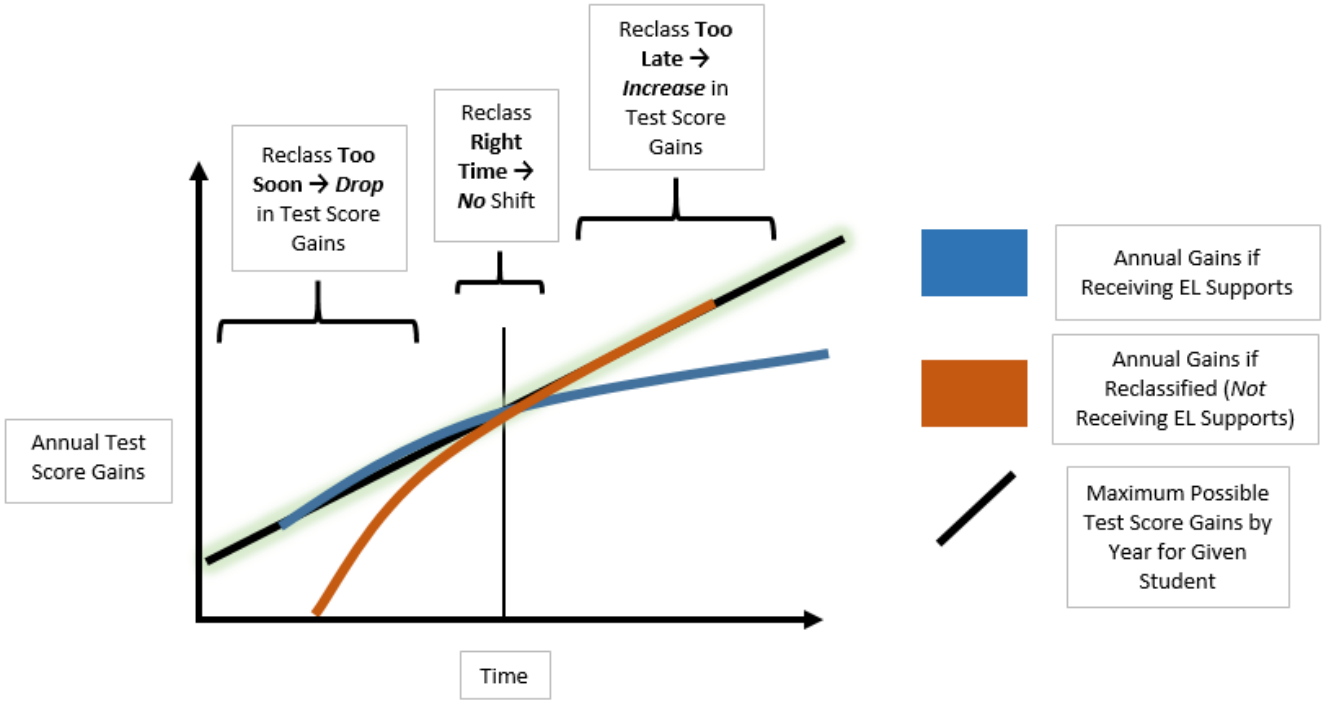
at all grade levels in 2005-06 and then lowered the threshold to the current level for secondary students in 2006-07. Perhaps even more important, the state's rescaling of its language test (which made it more difficult) in the 2006-07 school year.

Why Reclassification Could Produce Better Student Outcomes

There is a clear rationale for providing additional language support for English Learners. Classrooms designed for native English speakers are designed with the presumption that students already have a certain ability to comprehend and speak English. A further presumption is that students are roughly at grade level in reading and writing. The basis for EL support programs is that the rate of language acquisition is assumed to be slower for an EL student placed in a traditional class than for one placed into a special class with additional language supports because facility with English is a prerequisite for understanding teachers and curricular materials. However, as ELs' proficiency with English grows, at some juncture they may actually perform as well or better in classes which native English speakers typically take.

Figure 1 illustrates hypothetical annual test score gains depending on whether a student receives EL support. The blue line shows gains for the student if the student receives EL support in all years. The brown line shows the gains if the student does not get EL support. To the left of the vertical line, a student needs EL support to maximize test score gains. To the right of the vertical line, a student should be reclassified to maximize test score gains. Where the blue and brown lines intersect, a hypothetical student should transition from receiving EL services to being reclassified because annual test score gains become identical whether or not students get EL support. Thus, the optimal time to reclassify a student is shown by the vertical black line. In our experiments, if we see that students see a dip in their test scores post reclassification, this suggests that they were reclassified too soon and that reclassification criteria were too loose. If we observe that test scores jump post reclassification, this indicates students were reclassified later than would have been optimal, possibly holding them back from mainstream instruction or additional academic content. If we find no statistically significant difference in test score gain post reclassification, then reclassification criteria were optimally set for student progress.

Figure 1 Test Score Gains Rise, Fall, Or Stay The Same Depending On Whether Reclassification Occurs At The Right Time



SOURCE: Authors.

NOTE: Hypothetical effects of EL reclassification

Relevant Literature

Ideally, students are reclassified at the moment when EL support no longer benefits them and they are prepared to undertake an English-only instructional program without that support. Because reclassification policies vary widely across the state, it is unlikely that all school district policies are pegged to the moment of most benefit for an individual student. Capitalizing on our ability to precisely identify the reclassification policies used to decide when EL support should be removed, we estimate the causal relationship between continued EL instructional support and postsecondary outcomes for students who are performing near reclassification cutoffs under twodifferent reclassification policies.

How does being reclassified affect student outcomes? Prior research has also found that reclassified ELs are among the best performing students on a variety of academic measures (Hill et al., 2014; Saunders and Marcelletti, 2013; Gándara and Rumberger, 2006; EdSource, 2008; Flores, Painter, and Pachon, 2009). And prior research focusing on the rigor of reclassification criteria finds that students reclassified under more rigorous criteria in elementary school grades usually have slightly better outcomes than those reclassified under less rigorous criteria (Hill et al., 2014; Kim and Herman, 2012), but that those reclassified at older ages with more rigorous criteria may not (Hill, Weston, and Hayes, 2014). Again, these are observational studies and say

nothing about causation. The mere act of increasing standards for reclassification should mechanically improve outcomes for reclassified students by removing the left tail of the original distribution of reclassified students.

We hypothesize that the decision to reclassify could influence college-going behavior of English learners and reclassified English Learners, notably college enrollment, persistence, and graduation. There are several moderating factors that we will ultimately consider in addition to the rigor of the reclassification criteria, which can vary across districts (Parrish et al., 2006; Hill, Weston, and Hayes, 2014). For example, prior research has demonstrated that grade at reclassification is an important correlate of outcomes (Hill, Weston, and Hayes, 2014; Hill et al., 2014).

Ours is one of several studies that have had access to the student-level data needed to establish causality using and RD approach. We use the data to examine whether reclassified students do better or worse *as a result of reclassification* than their counterparts who just miss the reclassification cutoff. It is, as far as we know, the first to study the impact of the reclassification decision on college enrollment, persistence, *and* degree attainment. We use the data to examine whether reclassified students do better or worse *as a result of reclassification* than their counterparts who just miss the reclassification cutoff.

There are two prior studies that use regression discontinuity to examine college enrollment (Johnson 2020 and Carlson and Knowles 2015). In one California school district, just missing the reclassification cutoff resulted in higher chances of college enrollment than scoring above the cutoff (Johnson 2020). A Wisconsin study found the opposite – scoring just above the cutpont in 10th grade resulted in higher chances of college enrollment (Carlson and Knowles 2015). The present paper is the first to ask this question about college enrollment in San Diego.

Others have studied reclassification and elementary and secondary outcomes in large California districts including many in Los Angeles Unified School District. Robinson (2011) studied reclassification policy at a single point in time in an unnamed medium-sized California district, finding that the district had appropriate reclassification criteria for elementary and middle school students, but may have reclassified high school students too soon. Pope (2016), examining LAUSD reclassification decisions from 2002–03 to 2003–04, looked at the impact of being just above or below the CELDT cutpoint.[8] Robinson-Cimpian and Thompson (2016) investigated two LAUSD reclassification policy eras. This study looked at only one of five reclassification criteria (the CST, which measures ELA basic skills), and therefore did not fully evaluate whether the overall policy in these periods was set appropriately.

Cimpian et al. (2017) present quite a different context—they use longitudinal data on two separate states' EL students to examine the effect of reclassification upon subsequent achievement, while comparing districts. They present results first for their estimates of statewide average effects, and then for inter-district variability within each state. The state-level analysis yields negligible estimates of reclassification's benefits to subsequent achievement. But intra-state variation in district policies is large; in some districts, they find a large, significant negative effect upon graduation; in others, a large and significant positive effect.

[8] Pope found that 2nd and 4th grade students just above the CELDT cutpoint had improved English test scores and GPA relative to those just below, and but no statistically significant differences for high school students just above or below the CELDT cutpoint.

Reyes and Hwang (2019) use an RD approach to examine a specific event—reclassification by the end of 8th grade in an unnamed southern Californian district. Non-causal results indicated that reclassification was associated with higher standardized test scores and better behavioral outcomes. But using the RD approach they find null effects of reclassification on each of the subsequent outcomes examined: CST-ELA scores, results on the California High School Exit Examination (CAHSEE), math course placement in high school, absences, and suspensions. The authors conclude reclassification thresholds in their district are appropriately assigned.

Johnson (2020) also uses a regression discontinuity approach to examine outcomes for students in a large unnamed California district during a time frame similar to ours. In this case, the event under consideration is reclassification in the 8th grade. Johnson is similarly unable to reject the null hypothesis of no effect of 8th grade reclassification upon all three outcomes: ELA scores in 9th, 10th, and 11th grades; SAT reading scores; and being on-track to graduate in 10th and 11th grades.

Data and Measures

Our sample includes all EL students in grades 3 through 12 in SDUSD for the school years between 2003-04 and 2015-16. Our RD design will control for baseline characteristics from the year before reclassification, with outcomes measured several years after reclassification.

Our outcomes include on-time high school graduation, and a broad array of postsecondary outcomes obtained from the National Student Clearinghouse. The two key variables we use as running variables are the CST ELA test score and the CELDT reading score. California administered the CST to all students (except certain students in special education and some recently arrived immigrant students) in grades 2 through 11 from 2001-02 through 2012-13. On-time graduation can be measured based on the expected year of graduation, based on the grade in which the student is enrolled when his or her records are included in the RD sample for that year.

We also use CST Math scores for grades 3 through 7. At grade 8 and later, students take different CST math tests depending on their courses, and test scores are not comparable. For example, 9th graders may be taking either algebra or geometry.

As in a randomized controlled trial, we can gain precision by controlling for baseline achievement. As mentioned earlier, most students are reclassified part way through a school year (most typically in spring in SDUSD), and thus they receive the treatment for much but not all of the reclassification year. Were therefore set the baseline year to be the year before the reclassification decision. We test for differences in numerous baseline variables between those above and below the cutpoint in each of the two tests used for reclassification.

Methods

We use a regression discontinuity design (RD) that exploits the rules the district has established for reclassification, as presented in Table 1. The assignment variables related to reclassification are the CST test scores and the overall and subtest scores on the CELDT and (in LAUSD) course marks. In both districts, a key distinction between what we call Era 1 and Era 2 is that in the

latter era, a new more rigorous CELDT test was introduced. The other key distinctions are as follows. In San Diego, in elementary schools the CST cutoff was raised relative to Era 1. In Los Angeles, in Era 2 the district eliminated the requirement of reaching a certain grade in math classes. In San Diego, then, in Era 2 the reclassification unambiguously became more rigorous, due to the new CELDT test and, in elementary grades, a higher cutpoint on the state reading test (the CST). In Los Angeles, in Era 2 the two main changes worked in opposite directions, as the district dropped the math grade requirement at the same time that the state introduced a more demanding CELDT test. Thus it remains somewhat ambiguous in Los Angeles whether the overall standards became more or less rigorous.

Two questions must be addressed: What are the treatment and control groups, and is there a meaningful contrast between the experiences of the two groups? Students just below the cutoff point remain in EL support (described earlier), and are our control group. Students at or above the cutoff point are generally reclassified. The act of reclassification is meaningful. By definition, reclassified students will be treated like otherwise similar native English speakers. Thus, the treatment is being reclassified and having English language development classes removed.

Conceptually, there are (at least) two ways to handle RD designs with more than one forcing variable, by combining the various criteria or by studying them separately. The What Works Clearinghouse (U.S. Department of Education, 2014) recommends that RDs for a given outcome, but based on different forcing variables, should be treated separately. We adopt this approach in our main analysis, which allows Us to test whether cutpoints are appropriate separately for the various criteria.

Table 1 Reclassification Criteria In SDUSD, And The Definition Of Eras

		San Diego
2003-06 (Era 1)	**2004-05 (Era 1)**	Basic ELA skills = 300+ English proficiency = "Early Advanced" overall English proficiency = no more than one "Intermediate" on subtests
	2005-06	Increased CST ELA threshold to 333+
		New CELDT test debuts in 2006-07
2006-12 (Era 2)	**2007-14 (Era 2)**	333+ CST ELA for elementary, 300+ grades 6-12

SOURCE: San Diego Unified School District, Master Plan for English Learners, 2009

NOTES: When the new CELDT was introduced in 2006-07, cut scores for proficiency levels were raised across all grades and subtests (CDE, 2007). The two eras in each district are highlighted using lighter and darker shading for Eras 1 and 2 respectively. The range of years listed here and in later figures and tables refers to spring of the given school year. For example, the reference to "2007-2014" means that the school years 2006-2007 through 2013-2014 are used.

Estimation Method

A simple intent-to-treat estimator tests for whether there is a discontinuity in the outcome at the cutoff value of the running variable. Specifically, for the subsample of students who were ever English Learners, let t denote the year in which the reclassification decision is made, let $Y_{is,t+x}$ denote the outcome of student i in school s in school year t+x, which is observed x years after the reclassification decision where $x \geqq 1$. (Apart from graduation on time, our outcomes are at the postsecondary so that x is typically five or ten years.) Let $Z_{i,t-1}$ be a vector of baseline characteristics represent a vector of background variables measured the year before the reclassification decision. We measure the vector Z in period t-1 because in the year of reclassification reclassified students receive treatment for part of the school year before being reclassified. This vector includes indicators for race/ethnicity, language spoken at home, gender, and baseline student achievement characteristics. In terms of statistical expectation, there should be no differences in baseline characteristics of those just above and below the cutoff of the forcing variable, although in finite samples differences will emerge. By controlling for these variables we increase precision.

Consider the RD analysis based on the cutoff score of the ELA CST. The spring CST score is used to make a reclassification decision in the following school year. Thus, to be reclassified in year t, a necessary but not sufficient condition is that $CST_{i,t-1} \geq 0$ where we have rescaled the test score to equal 0 at the level required for reclassification in the given grade. In addition, the student must meet the other cutoffs imposed by the given district on the CELDT. Define the dummy variable $ABOVE_{it} = 1(CST_{i,t-1} \geq 0)$, (thus equaling 1/0 as the CST score is non-negative/negative). We estimate linear models on either side of the cutoff or, equivalently, estimate the two models at the same time by interacting controls with the ABOVE dummy:

$$Y_{is,t+x} = \alpha + \beta CST_{i,t-1} + \delta ABOVE_{it} + \gamma ABOVE_{it} \cdot CST_{i,t-1} + Z_{i,t-1}{}'\Delta + \mu_{i,t+x} \quad (1)$$

If δ is not significantly different from zero, then we retain the null hypothesis of a zero causal impact of meeting the reclassification criterion on the outcome. (In the above model we assume a linear relation between the outcome and the running variable, $CST_{i,t-1}$, while allowing for different slopes on either side of the cutoff. In the main models we assume a more flexible quadratic model, and for robustness we later use higher order polynomials in the running variable as well.)

The intent-to-treat model in (1) estimates the causal effect of meeting the reclassification criterion, but does not tell us the impact of treatment on the treated, that is, the impact of actual reclassification. Because we will have a fuzzy regression discontinuity design, based on Hahn et al. (2001) we can estimate the causal effect of reclassification, using a Two Stage Least Squares (2SLS) strategy. This approach produces a causal estimate of the impact of reclassification, which can be interpreted as a Local Average Treatment Effect (LATE). The coefficient of interest is a consistent estimate of the average causal effect of reclassification for ELs who were close to the cutpoint and who would comply with the reclassification policy.[9]

Our instrument for reclassification is the dummy variable $ABOVE_{it}$. To perform 2SLS, in the first stage we model the actual reclassification decision as

[9] Despite having a fuzzy RD because not all students are reclassified when they should be, we obtain consistent estimates under certain conditions (Angrist et al., 1996). The fact that not all students comply does not raise concerns of bias, but it does raise concerns about external validity (applicability to non-compliers), an issue that also exists in randomized controlled trials.

$$R_{it} = a + b\,ABOVE_{it} + Z_{i,t-1}'\Gamma + \epsilon_{it} \quad (2)$$

where Γ is a vector of coefficients, a and b are coefficients, and ε_{it} is an error term. In the second stage, we model outcomes x years after the reclassification decision, but replace actual reclassification with predicted reclassification $\hat{R}_{it}$. We estimate models that are linear and models that allow for a polynomial in the running variable. The linear model is:

$$Y_{is,t+x} = \theta + \mu CST_{i,t-1} + \pi\hat{R}_{it} + \rho ABOVE_{it} \cdot CST_{i,t-1} + Z_{i,t-1}'\Lambda + \xi_{i,t+x} \quad (3)$$

If π is not significantly different from zero, then we retain the null hypothesis of a zero causal impact of reclassification on test scores.

Tests for Manipulation of the Running Variable(s)

To have a valid RD design, the running variables used to determine reclassification such as the CST ELA score should not be easily manipulated by teachers or other school officials who may take a personal interest in either reclassifying or not reclassifying a given student. Both the CST tests and the CELDT test are statewide tests, and they are graded outside of the given school district, which greatly reduces but does not eliminate the possibility of the scores being manipulated locally.

However, it is also useful to check for discontinuities in the density of the running variable at the cutoff point (McCrary, 2008). It is not required that the distribution be continuous at the cutoff to have a valid RD design (Imbens and Lemieux, 2008), but it increases confidence that no manipulation occurred.

A related check for manipulation of the running variable involves testing for a discontinuity in one or more baseline characteristics at the cutoff value of the running variable in the year before the reclassification decision is made.

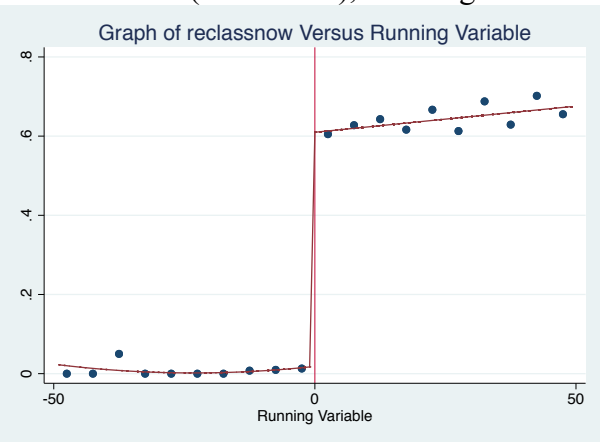
Results

Results of RD Validity Checks

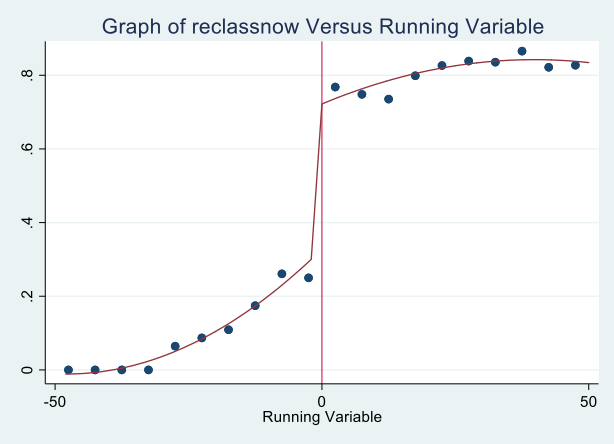
We start by establishing whether the regression discontinuity design applies for the running variable (ELA CST scores), by checking for a discontinuity in the probability that a student is reclassified at the cutpoint. Figure 2 gives three graphical examples of the jump in reclassification rates from SDUSD, for elementary schools. The first two graphs show jumps in the probability of reclassification of about 50 to 60 percentage points as students move above the CST reading cutpoint for Eras 1 and 2. The third graph shows a smaller but still large jump, about 35 percentage points, as students move above the CELDT reading (Early Advanced cutpoint in Era 2. in the first era in each district. Tables 1A and 1B show the results from the underlying models of the reclassification probability for the CST RD experiments in Eras 1 and 2 and for the CELDT Reading experiment in Era 2, respectively. The lines in the graph show the regression fit using a quadratic polynomial estimated independently on the two sides of the cutoff.

FIGURE 2 Probability of EL Reclassification at The CST and CELDT Reading Cutpoints, Grades 3-5, by Era

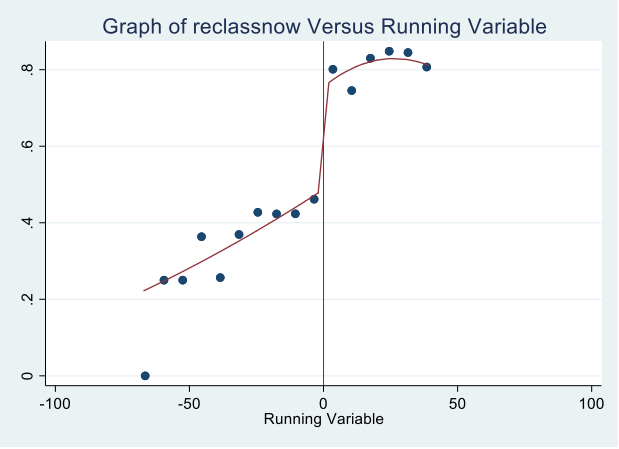
SDUSD Era 1 (2004-2005), Running Variable: CST



SDUSD Era 2 (2007-20014), Running Variable: CST



SDUSD Era 2 (2007-2014), Running Variable CELDT Reading



SOURCE: Authors' estimates

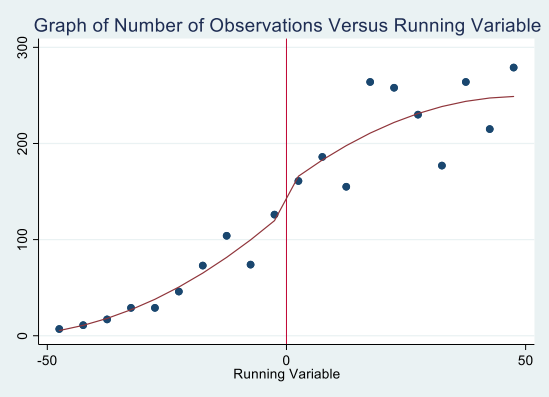
NOTE: The lines in the graphs show the regression fit using a quadratic polynomial estimated independently on the two sides of the cutoff. The dots represent bin sizes of 5. The range of years listed here and in later figures and tables refers to spring of the given school year.

We next turn to checks on the validity of the RD design. We begin by performing the McCrary test for manipulation of the running variable. We expect to see no discontinuity in the frequency of students above versus below the cutpoint. Table 2 shows results for this test for the sample used to model the impact of reclassification on graduation on time and on whether the student obtains a Bachelor's degree within five years of high school graduation with section a pertaining to the CST experiment and section b pertaining to the CELDT experiment. (Subsequent tables are split in the same way.) [10] The coefficient on ABOVE is not significantly different from zero in either subsample, for any of the three grade spans. Figure 3 gives three examples, showing the distribution of students versus the running variable CST for elementary grades in Era 1 and 2. The bottom graph shows the number of observations above and below the CELDT reading cutpoint in elementary grades in Era 2. In each case there is no clear break in the number of students at the cutpoint. The distribution of students for the CELDT is not particularly smooth, though.

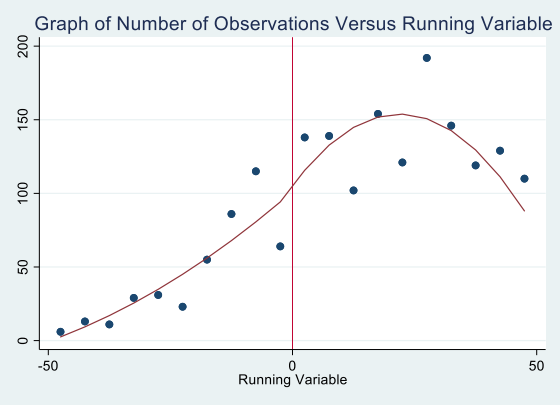
[10] Performing this twice on the earliest observed and latest observed outcome is useful because if there is differential attrition that occurs in the data after grade 12 the McCrary test may show a balance in numbers for the earlier outcome, high school graduation, but an imbalance for the outcome observed five years later.

Figure 3 Frequency of Observations at the Cutoff, for Elementary Students, Era 1 and Era 2, CST Running Variable, and Era 2, CELDT Reading Running Variable

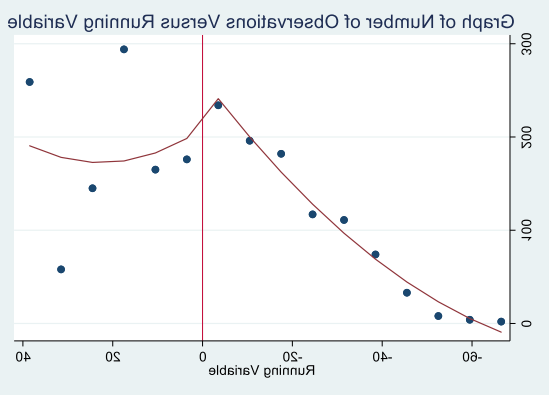
Era 1, CST Running Variable



Era 2, CST Running Variable



Era 2, CELDT Reading Running Variable



Notes: See notes to Figure 2.

While not strictly required for the RD design to be valid, a finding that there is not any discontinuous jump in the mean value of each background variable at the CST cutoff would provide reassurance that the treatment and control groups are similar. Table 3 shows the coefficient on ABOVE in models where the dependent variable is one of a number of student background characteristics. We present these models for three different samples: students in the model for graduation time, those in the sample for any postsecondary enrollment in the first year after grade 12, and those in the sample for whether the student has a Bachelor's degree within five years of grade 12. Any differences in results could reflect non-random attrition.

For these models of the baseline values of background variables, based on the CST experiments, the discontinuity variable is almost always statistically insignificant, with two minor exceptions. Table 3a shows the CST results. Academic year and grade level were statistically significant in the samples for Era 2 middle school and in Era 1 high school respectively. Both of these differences occurred in the sample for the outcome of high school graduation on time. When we instead focused on the sample for which we measured obtaining a Bachelor's degree in five years or less, shown in the second panel of the table, none of the baseline characteristics showed any discontinuity in baseline variables.

Overall, it appears that SDUSD has adhered to its stated reclassification policies quite closely, and that the fuzzy RD design is appropriate. But we can still improve on precision by controlling for baseline characteristics.

For the CELDT experiment in Era 2, tests for discontinuities in baseline characteristics appear in Table 3b. Unlike for the CST experiments, there were several cases in which we found significant discontinuities. For this reason, although we report results for this experiment, we will need to devote more attention to understanding the CELDT cutpoint experiment before placing full confidence in it. We return to this point in the conclusion.

CST Experiments

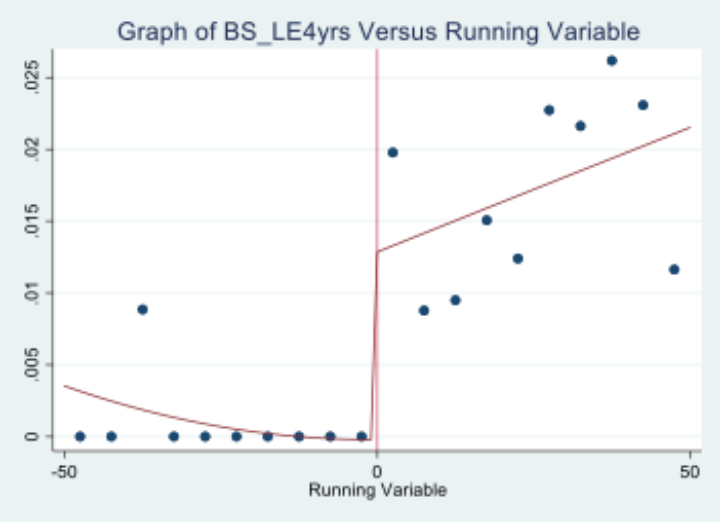
Turning now to the estimated impacts, Tables 4 and 5 show sample sizes for the CST experiments and summary statistics for the outcome variables and baseline characteristics respectively. Table 6 shows the coefficient on the indicator for meeting the CST cutoff, where the dependent variable is one of the outcomes of interest. We show intent to treat estimates and the impact of treatment on the treated on the right. We also show four specifications, with the running variable entered as a linear, quadratic and third and fourth order polynomials, estimated separately on each side of the cutoff. Visual inspection led us to focus on the quadratic (second order) models due to slight linearities. These columns have a yellow header.

In brief, for the CST cutoff in both eras, for the most part outcomes display no discontinuity at the CST cutoff, indicating that for outcomes ranging from graduation from high school on time through various postsecondary outcomes, the district was reclassifying students at the right time. There were a few exceptions from era 1. In era 1 in middle schools those meeting the CST cutpoint were more likely to obtain a Bachelor's degree in 4 or 5 years, with the effect rising from 1.39% for the former outcome to 2.04% for the latter outcome. The other exception was

that in era 1 among high school students, those meeting the CST cutpoint were 1% less likely to obtain a Bachelor's degree in 4 years. Notably, however, there was no significant impact on whether these students obtained a Bachelor's degree in five years.

Visual inspection further reinforced the likelihood of a positive middle school effect. Figure 4 shows the quite obvious discontinuity in the likelihood of obtaining a Bachelor's degree. In contrast, the negative impact in high school was not visually obvious and therefore is less credible.

Figure 4 Discontinuity in the Probability of Obtaining a Bachelor's Degree in Four Years or Less, for the CST Experiment, Era 1, Middle School



Thus it is the Era 1 middle school results for obtaining a Bachelor's degree that seem more robust. Those meeting the cutpoint were more likely to obtain a Bachelor's degree, implying that students were being reclassified too late. Interestingly, in era 2 we do not find this pattern. It could be that the more difficult CELDT standard in this later period made a difference.

CELDT Experiments

Results from the experiment involving the CELDT Reading (Early Advanced cutpoint) appear in Table 7. Unlike for the CST experiment, we find that for most postsecondary outcomes the impact of meeting the CELDT cutpoint requirement is positive and meaningfully large. This pattern occurred most often in elementary schools. In an analysis that combined middle and high schools (to give adequate sample sizes), results were more mixed, with positive impacts on community college enrollment in the first two years after high school graduation but a negative impact on obtaining a Bachelor's degree in less than or equal to four years, but no effect five years after high school graduation. The overall pattern of better outcomes for those meeting the CELDT cutpoint, at least in elementary schools, suggests that in Era 2 students were being reclassified too late.

This said, we are in the process of exploring the CELDT data in much more detail. We are concerned that numerous discontinuities emerged in testing for baseline equivalence in this experiment. If those discontinuities cannot be resolved, we cannot place much confidence in this part of the results. Another reason for caution is that many of the impacts that are significant in the quadratic model are not significant in the models that include the running variable only linearly. Part of the pattern of positive impacts appears to be due to overfitting in the quadratic model.

Conclusion

We ask in this paper if English Learner support services are being removed at the right time by examining post-secondary outcomes: do reclassified students fare better or worse when it comes to enrolling in, and graduating from, college? This paper provides an example of how RD designs can help policymakers determine whether their criteria remove the temporary supports at the optimal time. This approach could be used in many different settings.

Overall, in the two reclassification eras we studied in San Diego Unified (the second largest district in California), reclassification criteria appear to largely have been appropriate.

When it comes to examining the cutpoint used on the test of English Language Arts (the CST), there were just a few instances in Era 1 where it appears that middle school students may have been reclassified too soon, and one where it appears that high school students could have benefited from earlier reclassification (i.e. earning a B.A. within four years). When the CELDT reclassification threshold increased in the Era 2, without exception we found that those reclassified performed about as well those just below the CST cutpoint.

We were also able to examine whether the CELDT reading cutscores have been set appropriately, but only in Era 2, and we are less confident in these results than those for the CST. While the diagnostics required for a valid RD experiment were met (e.g. discontinuities at the cut score and the McCrary test), we did observe a number of background variables that were significant just above the cutpoint. Our preliminary conclusions are that elementary ELs may have been subject to reclassification policies that were too rigorous in Era 2, at least when it comes to postsecondary outcomes. In middle and high school, the case is more mixed – there are two post-secondary outcomes for which it appears students should have been reclassified sooner, and one where it appears they are reclassified too late. For the rest of the post-secondary outcomes, it appears that middle and high school students were reclassified at the appropriate time.

An interesting puzzle that arises is how all of the academic outcomes that are associated with reclassification decisions should be balanced. Our earlier paper found that in San Diego, cutpoints were largely set appropriately in both Era 1 and Era 2 when it comes to elementary, middle, and high school outcomes, including graduating on time. Is there a way to further optimize reclassification policy such that these results can hold ***and*** the only effects of reclassification on post-secondary outcomes are neutral? Or might districts be better advised to direct college counseling resources to all students, including current and former English Learners?

Finally, we conclude by pointing out that educators and policymakers have long understood that it makes no sense for similar students to have different EL status depending on where they attend school. Indeed, a survey of California districts by Hill, Weston, and Hayes (2014) revealed considerable variation in reclassification policies among districts. This idea of the need to create a common set of rules for reclassifying students has trickled up to the federal level. Standardization of reclassification policies is required by the federal Every Student Succeeds Act (ESSA), signed into law in 2015. ESSA is likely to create considerable disruption as districts in each state abandon their former policies in favor of new reclassification policies that individual states are now designing.

For the many states with large EL populations it will be important to assess whether states' new standards, which are currently being designed and rolled out, are set appropriately. Monitoring recently reclassified students, as is required by ESSA, is one important way to do so, but regression discontinuity modeling represents an ideal approach to answering the question of whether the policies are the optimal ones.

References

Angrist, Joshua D., Imbens, Guido W., & Rubin, Donald B. (1996). "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*: 91(434), 444-455.

California Education Code, Section 313(f).

Carlson, Deven, and Jared E Knowles. 2016. "The Effect of English Language Learner Reclassification on Student ACT Scores, High School Graduation, and Postsecondary Enrollment: Regression Discontinuity Evidence from Wisconsin." *Journal of Policy Analysis & Management* 35 (3).

EdSource. 2008. *English Learners in California: What the Numbers Say.* Mountain View, CA.

Flores, E., G. Painter, and H. Pachon. 2009. *¿Qué Pasa?: Are English Language Learning Students Remaining in English Learning Classes Too Long?* Los Angeles: Tomás Rivera Policy Institute.

Gándara, P., and R. Rumberger. 2006. "Resource Needs for California's English Learners." Linguistic Minority Research Institute, University of California.

Hahn, J., P. Todd, and W. Van Der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression Discontinuity Design." *Econometrica* 69: 201-209.

Hill, L, J. Betts, M. Hopkins, M. Lavadenz, K. Bachofer, J. Hayes, A. Lee, M. Murillo, T. Vahdani, and A. Zau. 2019. A*cademic Progress of English Learners: The Role of School Language Environment and Course Placement in Grades 6-12*. Public Policy Institute of California.

Hill, L, and I. Ugo. 2016. *High-Need Students and California's New Assessments*. Public Policy Institute of California.

Hill, L., J. Betts, B. Chavez, A. Zau, and K. Bachofer. 2014. *Pathways to Fluency: Identifying Success among Current and Former English Learners at the Elementary School Level*. San Francisco: Public Policy Institute of California.

Hill, L., M. Weston, and J. Hayes. 2014. *Reclassification of English Learner Students in California*. San Francisco: Public Policy Institute of California.

Imbens, G.W., and T. Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics*142: 615-635.

Johnson, Angela. 2020. The Effects of English Learner Classification on High School Graduation and College Attendance." AERA Open , Vol. 5, No. 2: 1-23.

Kanno, Yasuko, and Jennifer Cromley. 2015. English Language Learners' Pathways to Four-Year Colleges. Teachers College Record, Vol. 117.

Kim, J., and J. L. Herman. 2012. "Understanding Patterns and Precursors of ELL Success Subsequent to Reclassification." CRESST Report 818. National Center for Research on Evaluation, Standards, and Student Testing.

Linquanti, R. and H.G. Cook. 2015. Re-examining Reclassification: Guidance from a National Working Session on Policies and practices for Exiting Students for English Learners Status. Council of Chief State School Officers.

Mavrogordato, M. and R.S. White. 2017. "Reclassification Variation: How Policy Implementation Guides the process of exiting Students from English Learner Status." *Educational Evaluation and Policy Analysis*, Vol 39, No. 2.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics*, 142(2): 698–714.

McLaughlin, M., L. Glaab, I. Hilliger Carrasco, June 2014. Implementing Common Core State Standards in California: A Report from the Field. Stanford University and PACE.

Murillo, Marco and Magaly Lavadenz. 2020. Examining English Leaners' College Readiness and Postsecondary Enrollment in California. Education and Policy Brief, No. 8. Loyola Marymount University.

National Center for Education Statistics, 2019. *English Language Learners in Public Schools*.

Parrish, T., A. Merickel, M. Perez, R. Linquanti, M. Socia, A. Spain, C. Speroni, P. Esra, L. Brock, and D. Delancey. 2006. *Effects of the Implementation of Proposition 227 on the Education of English Learners, K-12, Findings from a Five-Year Evaluation*. Palo Alto: American Institutes for Research and WestEd.

Pope, N.G. 2016. "The Marginal Effect of K-12 English Language Development Programs: Evidence for Los Angeles Schools." Economics of education Review: 53 (2016), 311-328.

Robinson, J. P. 2011. "Evaluating Criteria for EL Reclassification: A Causal-Effects Approach Using a Binding-Score Regression Discontinuity Design with Instrumental Variables." *Education Evaluation and Policy Analysis 33*(3), 267–92.

Robinson-Cimpian, J.P. and K.D. Thompson. 2016. The Effects of Changing Test-Based Policies for Reclassifying English Learners. *Journal of Policy Analysis and Management*, 35(2), 279-305.

San Diego Unified School District. 2009. *English Learner Master Plan*.

Saunders, W. M., and D. J. Marcelletti. 2013. “The Gap That Can’t Go Away: The Catch-22 of Reclassification in Monitoring the Progress of English Learners.” Educational Evaluation and Policy Analysis, 35 (2): 139-156.

Umansky, I.M., S. Reardon, K. Hakuta, K.D. Thompson, P. Estrada, K. Hayes, H. Maldonado, S. Tandberg, C. Goldenberg. October 2015. Improving the Opportunities and outcomes of California’s Students learning English: Findings from school District-University Collaborative Partnerships. PACE.

U.S. Department of Education. 2014. *What Works Clearinghouse Procedures and Standards Handbook Version 3.0.* Washington, D.C.: What Works Clearinghouse, Institute for Education Sciences.

Warren, P. and P. Murphy. April 2014. “Implementing the Common Cores State Standards in California,” Public Policy Institute of California.

Table 1a
SDUSD Size of Discontinuity in Reclassification Rates for **CST** Reclassification Criterion, Reclassification Era 1 and Era 2

Reclassification Rate Discontinuity by Outcome		CELDT Era 1			CELDT Era 2		
		Grades 3-5	Grades 6-8	Grades 9-12	Grades 3-5	Grades 6-8	Grades 9-12
Graduation on time	Coef.	0.592**	0.645**	0.763**	0.394**	0.729**	0.494**
	S.E.	(0.0362)	(0.0282)	(0.0386)	(0.0610)	(0.0255)	(0.0311)
Postsecondary 2 or 4 year 1st year	Coef.	0.585**	0.639**	0.737**	0.494**	0.718**	0.501**
	S.E.	(0.0308)	(0.0255)	(0.0374)	(0.0349)	(0.0197)	(0.0288)
Postsecondary 2 or 4 year 2nd year	Coef.	0.585**	0.639**	0.737**	0.494**	0.718**	0.501**
	S.E.	(0.0308)	(0.0255)	(0.0374)	(0.0349)	(0.0197)	(0.0288)
Postsecondary 2 year 1st year	Coef.	0.585**	0.639**	0.737**	0.494**	0.718**	0.501**
	S.E.	(0.0308)	(0.0255)	(0.0374)	(0.0349)	(0.0197)	(0.0288)
Postsecondary 2 year 2nd year	Coef.	0.585**	0.639**	0.737**	0.494**	0.718**	0.501**
	S.E.	(0.0308)	(0.0255)	(0.0374)	(0.0349)	(0.0197)	(0.0288)
Postsecondary 4 year 1st year	Coef.	0.585**	0.639**	0.737**	0.494**	0.718**	0.501**
	S.E.	(0.0308)	(0.0255)	(0.0374)	(0.0349)	(0.0197)	(0.0288)
Postsecondary 4 year 2nd year	Coef.	0.585**	0.639**	0.737**	0.494**	0.718**	0.501**
	S.E.	(0.0308)	(0.0255)	(0.0374)	(0.0349)	(0.0197)	(0.0288)
AS or certification in <=4 years	Coef.	0.585**	0.639**	0.737**	0.494**	0.718**	0.501**
	S.E.	(0.0308)	(0.0255)	(0.0374)	(0.0349)	(0.0197)	(0.0288)
BS in <=4yrs	Coef.	0.585**	0.639**	0.737**	0.494**	0.718**	0.501**
	S.E.	(0.0308)	(0.0255)	(0.0374)	(0.0349)	(0.0197)	(0.0288)
AS or certification in <=5 years	Coef.	0.585**	0.639**	0.737**	0.494**	0.718**	0.501**
	S.E.	(0.0308)	(0.0255)	(0.0374)	(0.0349)	(0.0197)	(0.0288)
BS in <=5yrs	Coef.	0.585**	0.639**	0.737**	0.494**	0.718**	0.501**
	S.E.	(0.0308)	(0.0255)	(0.0374)	(0.0349)	(0.0197)	(0.0288)

Note: Cells in green indicate significant coefficients. In Tables 1-3 we use the model that employs a quadtratic model of the running variable on either side of the cutpoint.

Table 1b
SDUSD Size of Discontinuity in Reclassification Rates for **CELDT** Reclassification Criterion, Reclassification Era 2

Reclassification Rate Discontinuity by Outcome		CELDT Era 2	
		Grades 3-5	Grades 6-12
Graduation on time	Coef.	0.269**	0.207**
	S.E.	(0.0533)	(0.0293)
Postsecondary 2 or 4 year 1st year	Coef.	0.252**	0.227**
	S.E.	(0.0316)	(0.0245)
Postsecondary 2 or 4 year 2nd year	Coef.	0.252**	0.227**
	S.E.	(0.0316)	(0.0245)
Postsecondary 2 year 1st year	Coef.	0.252**	0.227**
	S.E.	(0.0316)	(0.0245)
Postsecondary 2 year 2nd year	Coef.	0.252**	0.227**
	S.E.	(0.0316)	(0.0245)
Postsecondary 4 year 1st year	Coef.	0.252**	0.227**
	S.E.	(0.0316)	(0.0245)
Postsecondary 4 year 2nd year	Coef.	0.252**	0.227**
	S.E.	(0.0316)	(0.0245)
AS or certification in <=4 years	Coef.	0.252**	0.227**
	S.E.	(0.0316)	(0.0245)
BS in <=4yrs	Coef.	0.252**	0.227**
	S.E.	(0.0316)	(0.0245)
AS or certification in <=5 years	Coef.	0.252**	0.227**
	S.E.	(0.0316)	(0.0245)
BS in <=5yrs	Coef.	0.252**	0.227**
	S.E.	(0.0316)	(0.0245)

Table 2a
McCrary Test for discontinuity in the number of students in the sample just above CST cutpoint

Outcome		CELDT Era 1			CELDT Era 2		
		Grades 3-5	Grades 6-8	Grades 9-12	Grades 3-5	Grades 6-8	Grades 9-12
Graduation on time	Coef.	26.41	61.19	-1.539	3.964	-97.99	41.27
	S.E.	(42.46)	(44.93)	(25.63)	(34.25)	(56.34)	(67.49)
Bachelor's in LE 5 years	Coef.	36.13	86.10	9.752	-9.975	-136.1	47.24
	S.E.	(53.11)	(50.51)	(28.12)	(70.89)	(105.8)	(74.89)

Table 2b
McCrary Test for discontinuity in the number of students in the sample just above CELDT cutpoint

Outcome		CELDT Era 2	
		Grades 3-5	Grades 6-12
Graduation on time	Coef.	-54.04	-134.5
	S.E.	(97.31)	(142.3)
Bachelor's in LE 5 years	Coef.	-60.10	-169.7
	S.E.	(184.6)	(156.2)

Table 3a
Checks for discontinuities in background variables above the CST reclassification cutpoint. Sample is based on using the outcome is any postsecondary enrollment in first year after high school graduation.

Baseline Variables		CELDT Era 1			CELDT Era 2		
		Grades 3-5	Grades 6-8	Grades 9-12	Grades 3-5	Grades 6-8	Grades 9-12
Female	Coef.	0.0317	0.0490	0.0703	0.0719	-0.00542	-0.0194
	S.E.	(0.0554)	(0.0422)	(0.0589)	(0.0457)	(0.0347)	(0.0339)
Spanish Home Language	Coef.	0.0534	0.0269	-0.0302	-0.0383	0.00482	0.00809
	S.E.	(0.0405)	(0.0314)	(0.0458)	(0.0396)	(0.0254)	(0.0280)
CST Math Basic	Coef.	0.0448			0.0226		
	S.E.	(0.0535)			(0.0297)		
CST Math Proficient	Coef.	-0.0263			-0.0162		
	S.E.	(0.0485)			(0.0442)		
CELDT Listening Scaled Score	Coef.	-7.298	0.0812	3.358	1.288	2.977	-2.605
	S.E.	(5.513)	(3.430)	(3.805)	(4.029)	(3.890)	(3.332)
CELDT Speaking Scaled Score	Coef.				5.697	-2.854	-2.108
	S.E.				(4.781)	(3.777)	(3.992)
CELDT Reading Scaled Score	Coef.	-3.539	1.322	1.969	3.824	-0.564	-0.893
	S.E.	(2.871)	(1.763)	(3.033)	(2.481)	(1.932)	(2.134)
CELDT Writing Scaled Score	Coef.	0.347	-0.410	-1.931	-3.081	1.112	-0.726
	S.E.	(2.606)	(2.092)	(3.385)	(2.203)	(1.976)	(2.461)
CELDT Overall Scaled Score	Coef.	-4.436	0.273	1.663	1.913	0.148	-1.572
	S.E.	(2.846)	(1.839)	(2.373)	(1.727)	(1.553)	(1.615)
Grade Level	Coef.	-0.0152	0.0864	0.126	-0.0393	0.0161	-0.107
	S.E.	(0.0626)	(0.0646)	(0.0693)	(0.0514)	(0.0519)	(0.0754)
Academic Year	Coef.	-0.0531	0.0110	0.0983	-0.0519	0.230	0.0572
	S.E.	(0.0538)	(0.0413)	(0.0533)	(0.0986)	(0.122)	(0.148)

Table 3a continued:

Table of checks on baseline characteristics (For sample with high school graduation on time)

Baseline Variables		CELDT Era 1			CELDT Era 2		
		Grades 3-5	Grades 6-8	Grades 9-12	Grades 3-5	Grades 6-8	Grades 9-12
Female	Coef.	-0.0220	0.0613	0.0263	-0.00528	-0.0686	-0.0347
	S.E.	(0.0639)	(0.0467)	(0.0620)	(0.0751)	(0.0447)	(0.0361)
Spanish Home Language	Coef.	0.0243	0.00499	-0.0679	-0.150*	0.0360	0.0160
	S.E.	(0.0453)	(0.0345)	(0.0492)	(0.0601)	(0.0335)	(0.0301)
CST Math Basic	Coef.	0.0698			0.0453		
	S.E.	(0.0619)			(0.0529)		
CST Math Proficient	Coef.	-0.0187			-0.00805		
	S.E.	(0.0557)			(0.0744)		
CELDT Listening Scaled Score	Coef.	-2.689	0.215	3.130	0.575	4.075	-3.927
	S.E.	(6.116)	(3.822)	(4.246)	(6.753)	(5.095)	(3.477)
CELDT Speaking Scaled Score	Coef.				3.463	-4.086	-2.823
	S.E.				(8.263)	(4.926)	(4.267)
CELDT Reading Scaled Score	Coef.	-2.457	1.104	2.184	3.610	-0.452	-0.680
	S.E.	(3.362)	(1.947)	(3.170)	(4.296)	(2.509)	(2.255)
CELDT Writing Scaled Score	Coef.	1.588	-1.451	-3.669	-4.716	0.385	-2.770
	S.E.	(2.907)	(2.306)	(3.581)	(3.529)	(2.604)	(2.629)
CELDT Overall Scaled Score	Coef.	-1.559	0.0197	1.155	0.686	-0.0475	-2.547
	S.E.	(3.133)	(2.032)	(2.631)	(2.923)	(2.017)	(1.707)
Grade Level	Coef.	0.00823	0.0611	0.154*	-0.0456	-0.0539	-0.0968
	S.E.	(0.0712)	(0.0704)	(0.0751)	(0.0760)	(0.0659)	(0.0770)
Academic Year	Coef.	-0.0745	0.0153	0.106	0.0268	0.370**	0.0178
	S.E.	(0.0621)	(0.0453)	(0.0557)	(0.112)	(0.133)	(0.154)

.

Table 3a continued:

Table of checks on baseline characteristics (for the sample with the outcome indicating a Bachelor's degree in less than or equal to 5 years)

Baseline Variables		CELDT Era 1			CELDT Era 2		
		Grades 3-5	Grades 6-8	Grades 9-12	Grades 3-5	Grades 6-8	Grades 9-12
Female	Coef.	0.0317	0.0490	0.0703	0.0719	-0.00542	-0.0194
	S.E.	(0.0554)	(0.0422)	(0.0589)	(0.0457)	(0.0347)	(0.0339)
Spanish Home Language	Coef.	0.0534	0.0269	-0.0302	-0.0383	0.00482	0.00809
	S.E.	(0.0405)	(0.0314)	(0.0458)	(0.0396)	(0.0254)	(0.0280)
CST Math Basic	Coef.	0.0448			0.0226		
	S.E.	(0.0535)			(0.0297)		
CST Math Proficient	Coef.	-0.0263			-0.0162		
	S.E.	(0.0485)			(0.0442)		
CELDT Listening Scaled Score	Coef.	-7.298	0.0812	3.358	1.288	2.977	-2.605
	S.E.	(5.513)	(3.430)	(3.805)	(4.029)	(3.890)	(3.332)
CELDT Speaking Scaled Score	Coef.				5.697	-2.854	-2.108
	S.E.				(4.781)	(3.777)	(3.992)
CELDT Reading Scaled Score	Coef.	-3.539	1.322	1.969	3.824	-0.564	-0.893
	S.E.	(2.871)	(1.763)	(3.033)	(2.481)	(1.932)	(2.134)
CELDT Writing Scaled Score	Coef.	0.347	-0.410	-1.931	-3.081	1.112	-0.726
	S.E.	(2.606)	(2.092)	(3.385)	(2.203)	(1.976)	(2.461)
CELDT Overall Scaled Score	Coef.	-4.436	0.273	1.663	1.913	0.148	-1.572
	S.E.	(2.846)	(1.839)	(2.373)	(1.727)	(1.553)	(1.615)
Grade Level	Coef.	-0.0152	0.0864	0.126	-0.0393	0.0161	-0.107
	S.E.	(0.0626)	(0.0646)	(0.0693)	(0.0514)	(0.0519)	(0.0754)
Academic Year	Coef.	-0.0531	0.0110	0.0983	-0.0519	0.230	0.0572
	S.E.	(0.0538)	(0.0413)	(0.0533)	(0.0986)	(0.122)	(0.148)

Table 3b

Checks for discontinuities in background variables above the CELDT reclassification cutpoint.

Baseline Variables		Postsec2_4yr1		Gradontime		BS_LE5yrs	
		Grades 3-5	Grades 6-12	Grades 3-5	Grades 6-12	Grades 3-5	Grades 6-12
Female	Coef.	-0.0124	0.0141	0.0209	0.0302	-0.0124	0.0141
	S.E.	(0.0360)	(0.0256)	(0.0623)	(0.0307)	(0.0360)	(0.0256)
Spanish Home Language	Coef.	-0.0189	0.0652**	-0.0300	0.0748**	-0.0189	0.0652**
	S.E.	(0.0331)	(0.0207)	(0.0563)	(0.0252)	(0.0331)	(0.0207)
CST Math Basic	Coef.	0.0485**		0.0673*		0.0485**	
	S.E.	(0.0182)		(0.0341)		(0.0182)	
CST Math Proficient	Coef.	0.0907**		0.106		0.0907**	
	S.E.	(0.0309)		(0.0551)		(0.0309)	
Grade Level	Coef.	-0.204**	-0.546**	-0.223**	-0.627**	-0.204**	-0.546**
	S.E.	(0.0535)	(0.0932)	(0.0799)	(0.115)	(0.0535)	(0.0932)
Academic Year	Coef.	-0.0976	-0.702**	-0.310**	-0.798**	-0.0976	-0.702**
	S.E.	(0.0778)	(0.0990)	(0.0912)	(0.114)	(0.0778)	(0.0990)

Note: Column headers indicate the outcome variable used to choose the sample.

Table 4: Table of sample sizes for each outcome using CST as the running variable

Outcome	CELDT Era 1			CELDT Era 2		
	G3-5	G6-8	G9-12	G3-5	G6-8	G9-12
Gradontime Baseline	2705	3527	1636	1814	4271	5715
Gradontime ITT	2702	3527	1636	1813	4271	5715
Postsec2yr1 Baseline	3718	4267	1816	4740	7354	6400
Postsec2yr1 ITT	3714	4267	1816	4738	7354	6400
Postsec2yr2 Baseline	3718	4267	1816	4740	7354	6400
Postsec2yr2 ITT	3714	4267	1816	4738	7354	6400
Postsec4yr1 Baseline	3718	4267	1816	4740	7354	6400
Postsec4yr1 ITT	3714	4267	1816	4738	7354	6400
Postsec4yr2 Baseline	3718	4267	1816	4740	7354	6400
Postsec4yr2 ITT	3714	4267	1816	4738	7354	6400
Postsec2_4yr1 Baseline	3718	4267	1816	4740	7354	6400
Postsec2_4yr1 ITT	3714	4267	1816	4738	7354	6400
Postsec2_4yr2 Baseline	3718	4267	1816	4740	7354	6400
Postsec2_4yr2 ITT	3714	4267	1816	4738	7354	6400
AS_cert_le4yrs Baseline	3718	4267	1816	4740	7354	6400
AS_cert_le4yrs ITT	3714	4267	1816	4738	7354	6400
BS_LE4yrs Baseline	3718	4267	1816	4740	7354	6400
BS_LE4yrs ITT	3714	4267	1816	4738	7354	6400
AS_cert_le5yrs Baseline	3718	4267	1816	4740	7354	6400
AS_cert_le5yrs ITT	3714	4267	1816	4738	7354	6400
BS_LE5yrs Baseline	3718	4267	1816	4740	7354	6400
BS_LE5yrs ITT	3714	4267	1816	4738	7354	6400

Table 5 Summary statistics (using gradontime as outcome for the CST experiment)

Variable		Intent to treat					
		CELDT Era 1			CELDT Era 2		
		g3-5	g6-8	g9-12	g3-5	g6-8	g9-12
Graduated on time	# Obs	2595	3309	1384	893	4118	1750
	Mean	0.755684	0.647628	0.610549	0.853303	0.788975	0.724571
	SD	0.429764	0.477781	0.487802	0.354001	0.408085	0.446858
Reclassified English Proficient	# Obs	3414	3903	1513	3740	6540	1931
	Mean	0.521383	0.419933	0.317911	0.462032	0.56789	0.45624
	SD	0.499616	0.493611	0.465819	0.498623	0.495407	0.49821
Spanish language	# Obs	3414	3903	1513	3740	6540	1931
	Mean	0.818395	0.837048	0.827495	0.760428	0.821101	0.812015
	SD	0.385575	0.369369	0.377944	0.426879	0.383297	0.390802
CELDT Listening Scaled Score	# Obs	3414	3903	1513	3740	6540	1931
	Mean	571.3617	560.5837	549.2855	574.4209	633.206	662.7861
	SD	48.35494	42.44957	35.53401	42.20504	57.69468	48.44208
CELDT Reading Scaled Score	# Obs	3414	3903	1513	3740	6540	1931
	Mean	544.4997	548.2121	561.727	563.969	591.8668	623.5422
	SD	28.13981	23.5968	24.93749	28.83008	29.90951	33.45035
CELDT Writing Scaled Score	# Obs	3414	3903	1513	3740	6540	1931
	Mean	553.0164	555.8101	553.5466	563.1457	578.7709	596.3729
	SD	24.92561	26.11882	28.60365	25.64098	30.48658	34.94739
CELDT Speaking Scaled Score	# Obs	3414	3903	1513	3740	6540	1931
	Mean	571.3617	560.5837	549.2855	548.212	585.8309	607.9855
	SD	48.35494	42.44957	35.53401	48.95743	53.7016	56.87191
CELDT Overall Scaled Score	# Obs	3414	3903	1513	3740	6540	1931
	Mean	559.7047	555.9221	553.0826	562.0594	597.0379	622.2947
	SD	25.97738	23.27223	21.37017	19.65197	23.97384	22.81604
CST Math Basic	# Obs	3414	3902	1510	3740	6533	1927
	Mean	0.74546	0.523834	0.22649	0.88877	0.65636	0.290088
	SD	0.435666	0.499496	0.418698	0.314459	0.474959	0.453921
CST Math Proficient	# Obs	3414	3902	1510	3740	6533	1927
	Mean	0.384007	0.160687	0.074834	0.639572	0.25685	0.100156
	SD	0.486431	0.367289	0.263211	0.480189	0.436929	0.300285
Female student	# Obs	3414	3903	1513	3740	6540	1931
	Mean	0.512009	0.476044	0.47191	0.50508	0.480122	0.386846
	SD	0.499929	0.49949	0.499375	0.500041	0.499643	0.487154
School year	# Obs	3414	3903	1513	3740	6540	1931
	Mean	2004.511	2004.531	2004.718	2009.717	2009.592	2009.606
	SD	0.499949	0.49911	0.449909	1.684818	1.61514	1.650733
Grade level	# Obs	3414	3903	1513	3740	6540	1931
	Mean	4.451963	7.085831	9.354263	4.399733	6.984557	9.484205
	SD	0.6440291	0.7878899	0.5027142	0.6200276	0.7912042	0.5172006

Table 5 (continued) Summary statistics (using post-secondary 2 or 4 year in year 1 as outcome)

Variable		Intent to treat					
		CELDT Era 1			CELDT Era 2		
		g3-5	g6-8	g9-12	g3-5	g6-8	g9-12
Graduated on time	# Obs	2595	3309	1384	893	4118	1750
	Mean	0.755684	0.647628	0.610549	0.853303	0.788975	0.724571
	SD	0.429764	0.477781	0.487802	0.354001	0.408085	0.446858
Reclassified English Proficient	# Obs	3414	3903	1513	3740	6540	1931
	Mean	0.521383	0.419933	0.317911	0.462032	0.56789	0.45624
	SD	0.499616	0.493611	0.465819	0.498623	0.495407	0.49821
Spanish language	# Obs	3414	3903	1513	3740	6540	1931
	Mean	0.818395	0.837048	0.827495	0.760428	0.821101	0.812015
	SD	0.385575	0.369369	0.377944	0.426879	0.383297	0.390802
CELDT Listening Scaled Score	# Obs	3414	3903	1513	3740	6540	1931
	Mean	571.3617	560.5837	549.2855	574.4209	633.206	662.7861
	SD	48.35494	42.44957	35.53401	42.20504	57.69468	48.44208
CELDT Reading Scaled Score	# Obs	3414	3903	1513	3740	6540	1931
	Mean	544.4997	548.2121	561.727	563.969	591.8668	623.5422
	SD	28.13981	23.5968	24.93749	28.83008	29.90951	33.45035
CELDT Writing Scaled Score	# Obs	3414	3903	1513	3740	6540	1931
	Mean	553.0164	555.8101	553.5466	563.1457	578.7709	596.3729
	SD	24.92561	26.11882	28.60365	25.64098	30.48658	34.94739
CELDT Speaking Scaled Score	# Obs	3414	3903	1513	3740	6540	1931
	Mean	571.3617	560.5837	549.2855	548.212	585.8309	607.9855
	SD	48.35494	42.44957	35.53401	48.95743	53.7016	56.87191
CELDT Overall Scaled Score	# Obs	3414	3903	1513	3740	6540	1931
	Mean	559.7047	555.9221	553.0826	562.0594	597.0379	622.2947
	SD	25.97738	23.27223	21.37017	19.65197	23.97384	22.81604
CST Math Basic	# Obs	3414	3902	1510	3740	6533	1927
	Mean	0.74546	0.523834	0.22649	0.88877	0.65636	0.290088
	SD	0.435666	0.499496	0.418698	0.314459	0.474959	0.453921
CST Math Proficient	# Obs	3414	3902	1510	3740	6533	1927
	Mean	0.384007	0.160687	0.074834	0.639572	0.25685	0.100156
	SD	0.486431	0.367289	0.263211	0.480189	0.436929	0.300285
Female student	# Obs	3414	3903	1513	3740	6540	1931
	Mean	0.512009	0.476044	0.47191	0.50508	0.480122	0.386846
	SD	0.499929	0.49949	0.499375	0.500041	0.499643	0.487154
School year	# Obs	3414	3903	1513	3740	6540	1931
	Mean	2004.511	2004.531	2004.718	2009.717	2009.592	2009.606
	SD	0.499949	0.49911	0.449909	1.684818	1.61514	1.650733
Grade level	# Obs	3414	3903	1513	3740	6540	1931
	Mean	4.451963	7.085831	9.354263	4.399733	6.984557	9.484205
	SD	0.644029	0.78789	0.502714	0.620028	0.791204	0.517201

Table 6
Robustness of the Estimates to the Order of Polynomials in the Running Variable CST Used, Bandwidth of 50

Outcome	Era	Grade Span	Intent to Treat				Treatment on Treated			
			4th	3rd	2nd	1st	4th	3rd	2nd	1st
Postsecondary 2 or 4yr school, 1st year	CELDT Era 1 ('03-'06)	Elm	-0.0397	-0.0182	-0.0178	0.0381	-0.0690	-0.0321	-0.0303	0.0645
			(0.0953)	(0.0713)	(0.0525)	(0.0353)	(0.166)	(0.125)	(0.0891)	(0.0597)
		MS	-0.0439	-0.0125	-0.0146	0.0202	-0.0730	-0.0203	-0.0227	0.0309
			(0.0717)	(0.0530)	(0.0384)	(0.0258)	(0.119)	(0.0859)	(0.0599)	(0.0393)
		HS	-0.0374	-0.0547	-0.0440	-0.0135	-0.0518	-0.0713	-0.0593	-0.0182
			(0.0900)	(0.0694)	(0.0526)	(0.0362)	(0.124)	(0.0901)	(0.0706)	(0.0488)
	CELDT Era 2 ('07-'12)	Elm	-0.0359	-0.0190	-0.0417	-0.0571	-0.0646	-0.0385	-0.0834	-0.104
			(0.0879)	(0.0631)	(0.0449)	(0.0294)	(0.158)	(0.127)	(0.0900)	(0.0538)
		MS	0.0809	0.0497	0.0144	0.0270	0.120	0.0693	0.0203	0.0411
			(0.0550)	(0.0442)	(0.0334)	(0.0227)	(0.0814)	(0.0614)	(0.0472)	(0.0345)
		HS	0.0792	0.0511	-0.0108	0.0252	0.166	0.102	-0.0210	0.0457
			(0.0561)	(0.0432)	(0.0326)	(0.0222)	(0.118)	(0.0861)	(0.0634)	(0.0402)
Postsecondary 2yr school, 1st year	CELDT Era 1 ('03-'06)	Elm	-0.0689	-0.0602	-0.00414	0.0352	-0.120	-0.106	-0.00704	0.0595
			(0.0911)	(0.0687)	(0.0504)	(0.0339)	(0.159)	(0.121)	(0.0855)	(0.0573)
		MS	-0.0540	-0.0339	-0.0130	0.0163	-0.0896	-0.0550	-0.0203	0.0248
			(0.0695)	(0.0515)	(0.0372)	(0.0249)	(0.115)	(0.0834)	(0.0580)	(0.0380)
		HS	-0.0247	-0.0304	-0.0259	-0.0149	-0.0343	-0.0397	-0.0350	-0.0201
			(0.0886)	(0.0680)	(0.0513)	(0.0354)	(0.122)	(0.0883)	(0.0689)	(0.0477)
	CELDT Era 2 ('07-'12)	Elm	0.0311	0.0703	0.0302	-0.0329	0.0560	0.142	0.0605	-0.0602
			(0.0843)	(0.0607)	(0.0432)	(0.0283)	(0.151)	(0.123)	(0.0866)	(0.0517)
		MS	0.0437	0.0407	0.0199	0.0340	0.0649	0.0567	0.0282	0.0518
			(0.0534)	(0.0430)	(0.0324)	(0.0220)	(0.0791)	(0.0598)	(0.0457)	(0.0333)
		HS	0.0679	0.0400	-0.0236	0.0302	0.143	0.0799	-0.0459	0.0548

			(0.0553)	(0.0426)	(0.0323)	(0.0220)	(0.117)	(0.0850)	(0.0628)	(0.0398)
		Grade	Intent to Treat				Treatment on Treated			
Outcome	Era	Span	4th	3rd	2nd	1st	4th	3rd	2nd	1st
Postsecondary 4yr school, 1st year	CELDT Era 1 ('03-'06)	Elm	0.0173	0.0193	-0.0265	-0.00173	0.0300	0.0340	-0.0451	-0.00293
			(0.0538)	(0.0386)	(0.0288)	(0.0199)	(0.0935)	(0.0680)	(0.0489)	(0.0336)
		MS	-0.00384	0.0281	0.0121	0.0224	-0.00638	0.0456	0.0188	0.0342
			(0.0345)	(0.0253)	(0.0190)	(0.0132)	(0.0571)	(0.0411)	(0.0296)	(0.0201)
		HS	-0.00305	-0.0425	-0.0201	0.0124	-0.00423	-0.0554	-0.0271	0.0168
			(0.0419)	(0.0342)	(0.0268)	(0.0183)	(0.0578)	(0.0446)	(0.0360)	(0.0246)
	CELDT Era 2 ('07-'12)	Elm	-0.0293	-0.0450	-0.0629	-0.0124	-0.0526	-0.0910	-0.126	-0.0227
			(0.0631)	(0.0450)	(0.0323)	(0.0217)	(0.114)	(0.0916)	(0.0655)	(0.0396)
		MS	0.0437	0.00909	0.00440	-0.00608	0.0650	0.0127	0.00623	-0.00925
			(0.0288)	(0.0221)	(0.0169)	(0.0121)	(0.0428)	(0.0307)	(0.0239)	(0.0184)
		HS	0.0413	0.0222	0.0142	-0.00697	0.0869	0.0444	0.0276	-0.0126
			(0.0297)	(0.0223)	(0.0168)	(0.0119)	(0.0622)	(0.0442)	(0.0325)	(0.0215)
Postsecondary 2 or 4yr school, 2nd year	CELDT Era 1 ('03-'06)	Elm	0.0182	0.0298	0.0388	0.0436	0.0316	0.0525	0.0659	0.0738
			(0.0959)	(0.0711)	(0.0523)	(0.0352)	(0.166)	(0.125)	(0.0887)	(0.0595)
		MS	-0.0469	-0.0104	-0.0291	0.00881	-0.0778	-0.0169	-0.0454	0.0135
			(0.0713)	(0.0531)	(0.0386)	(0.0260)	(0.118)	(0.0860)	(0.0601)	(0.0396)
		HS	-0.0246	-0.0545	-0.0637	-0.0575	-0.0341	-0.0711	-0.0860	-0.0778
			(0.0899)	(0.0696)	(0.0528)	(0.0363)	(0.124)	(0.0905)	(0.0711)	(0.0489)
	CELDT Era 2 ('07-'12)	Elm	-0.0869	-0.0257	-0.0241	-0.0529	-0.156	-0.0519	-0.0484	-0.0967
			(0.0874)	(0.0628)	(0.0445)	(0.0292)	(0.159)	(0.127)	(0.0891)	(0.0534)
		MS	-0.0129	-0.0313	-0.0288	-0.00213	-0.0192	-0.0436	-0.0408	-0.00325
			(0.0557)	(0.0442)	(0.0334)	(0.0228)	(0.0827)	(0.0616)	(0.0473)	(0.0346)
		HS	0.0902	0.0670	0.0307	0.0280	0.190	0.134	0.0596	0.0508

			(0.0565)	(0.0432)	(0.0323)	(0.0220)	(0.119)	(0.0861)	(0.0627)	(0.0399)

		Grade	Intent to Treat				Treatment on Treated			
Outcome	Era	Span	4th	3rd	2nd	1st	4th	3rd	2nd	1st
Postsecondary 2yr 2nd year	CELDT Era 1 ('03-'06)	Elm	0.0173	-0.0156	0.0378	0.0261	0.0300	-0.0276	0.0643	0.0442
			(0.0927)	(0.0690)	(0.0505)	(0.0340)	(0.160)	(0.121)	(0.0856)	(0.0574)
		MS	-0.0449	-0.00638	-0.0219	0.0123	-0.0746	-0.0104	-0.0341	0.0187
			(0.0690)	(0.0514)	(0.0372)	(0.0251)	(0.114)	(0.0832)	(0.0580)	(0.0382)
		HS	-0.0274	-0.0278	-0.0429	-0.0639	-0.0379	-0.0363	-0.0579	-0.0865
			(0.0879)	(0.0681)	(0.0517)	(0.0355)	(0.121)	(0.0886)	(0.0696)	(0.0480)
	CELDT Era 2 ('07-'12)	Elm	-0.0745	0.0376	0.0260	-0.0287	-0.134	0.0760	0.0522	-0.0524
			(0.0844)	(0.0606)	(0.0428)	(0.0279)	(0.153)	(0.122)	(0.0855)	(0.0510)
		MS	-0.0243	-0.0275	-0.0177	0.0178	-0.0361	-0.0383	-0.0251	0.0271
			(0.0545)	(0.0434)	(0.0327)	(0.0222)	(0.0810)	(0.0605)	(0.0462)	(0.0337)
		HS	0.0680	0.0516	0.0226	0.0350	0.143	0.103	0.0440	0.0636
			(0.0559)	(0.0428)	(0.0321)	(0.0218)	(0.118)	(0.0853)	(0.0622)	(0.0395)
Postsecondary 4yr 2nd year	CELDT Era 1 ('03-'06)	Elm	0.0117	0.0392	-0.00949	0.0192	0.0204	0.0690	-0.0161	0.0326
			(0.0534)	(0.0386)	(0.0283)	(0.0188)	(0.0927)	(0.0682)	(0.0480)	(0.0319)
		MS	0.00551	0.0141	0.00232	0.00497	0.00914	0.0230	0.00362	0.00759
			(0.0322)	(0.0249)	(0.0184)	(0.0126)	(0.0534)	(0.0403)	(0.0287)	(0.0192)
		HS	0.0343	-0.0154	-0.000370	0.00810	0.0476	-0.0201	-0.0005	0.0110
			(0.0414)	(0.0322)	(0.0252)	(0.0174)	(0.0573)	(0.0418)	(0.0339)	(0.0235)
	CELDT Era 2 ('07-'12)	Elm	0.00211	-0.0434	-0.0497	-0.00918	0.00380	-0.0877	-0.0995	-0.0168
			(0.0583)	(0.0417)	(0.0304)	(0.0206)	(0.105)	(0.0847)	(0.0613)	(0.0376)
		MS	0.00675	-0.00643	-0.00703	-0.0219	0.0100	-0.00896	-0.00994	-0.0333
			(0.0277)	(0.0206)	(0.0159)	(0.0114)	(0.0411)	(0.0287)	(0.0225)	(0.0173)

		HS	0.0298	0.0147	0.0158	-0.00487	0.0626	0.0293	0.0307	-0.00885
			(0.0265)	(0.0203)	(0.0153)	(0.0108)	(0.0558)	(0.0405)	(0.0296)	(0.0197)

		Grade	Intent to Treat				Treatment on Treated			
Outcome	Era	Span	4th	3rd	2nd	1st	4th	3rd	2nd	1st
AS in LE 4yrs	CELDT Era 1 ('03-'06)	Elm	-0.0468	0.0168	0.00626	0.0195	-0.0813	0.0296	0.0106	0.0330
			(0.0399)	(0.0307)	(0.0228)	(0.0144)	(0.0695)	(0.0540)	(0.0386)	(0.0244)
		MS	0.00830	-0.000819	0.00224	-0.00330	0.0138	-0.00133	0.00349	-0.00504
			(0.0191)	(0.0154)	(0.0111)	(0.00777)	(0.0317)	(0.0249)	(0.0173)	(0.0118)
		HS	-0.0225	-0.0385	-0.0259	-0.0177	-0.0311	-0.0503	-0.0350	-0.0240
			(0.0319)	(0.0244)	(0.0160)	(0.0112)	(0.0441)	(0.0318)	(0.0216)	(0.0152)
	CELDT Era 2 ('07-'12)	Elm	0.0229	0.0330	0.00311	0.00662	0.0411	0.0667	0.00622	0.0121
			(0.0254)	(0.0191)	(0.0139)	(0.00942)	(0.0458)	(0.0391)	(0.0278)	(0.0172)
		MS	-0.00824	-0.0169	-0.0102	0.00146	-0.0123	-0.0236	-0.0145	0.00222
			(0.0226)	(0.0165)	(0.0120)	(0.00836)	(0.0336)	(0.0231)	(0.0169)	(0.0127)
		HS	-0.0218	-0.0269	-0.0223	-0.00346	-0.0458	-0.0536	-0.0434	-0.00629
			(0.0207)	(0.0155)	(0.0118)	(0.00782)	(0.0438)	(0.0312)	(0.0231)	(0.0142)
BS in LE 4yrs	CELDT Era 1 ('03-'06)	Elm	0.0225	0.0298**	0.00999	0.00872	0.0391	0.0526**	0.0170	0.0148
			(0.0147)	(0.00997)	(0.00979)	(0.00795)	(0.0260)	(0.0181)	(0.0166)	(0.0135)
		MS	0.0208*	0.0187*	0.0139*	0.0138**	0.0345*	0.0304*	0.0217*	0.0211**
			(0.00830)	(0.00776)	(0.00615)	(0.00429)	(0.0140)	(0.0127)	(0.00961)	(0.00656)
		HS	0.0111	-0.00550	-0.0113*	0.00203	0.0153	-0.00717	-0.0153*	0.00274
			(0.00653)	(0.00470)	(0.00535)	(0.00421)	(0.00906)	(0.00613)	(0.00724)	(0.00567)
	CELDT Era 2 ('07-'12)	Elm	-0.0265	-0.00242	0.00823	0.00993	-0.0476	-0.00489	0.0165	0.0182
			(0.0201)	(0.0139)	(0.00975)	(0.00630)	(0.0367)	(0.0280)	(0.0195)	(0.0115)
		MS	-0.000778	-0.00283	0.00279	0.000513	-0.00116	-0.00395	0.00395	0.000780

			(0.00788)	(0.00592)	(0.00466)	(0.00347)	(0.0117)	(0.00824)	(0.00658)	(0.00527)
		HS	0.00976	0.00272	0.00393	-0.00527	0.0205	0.00543	0.00764	-0.00956
			(0.00683)	(0.00614)	(0.00598)	(0.00401)	(0.0145)	(0.0122)	(0.0116)	(0.00728)

		Grade	Intent to Treat				Treatment on Treated			
Outcome	Era	Span	4th	3rd	2nd	1st	4th	3rd	2nd	1st
AS or cert in LE 5yrs	CELDT Era 1 ('03-'06)	Elm	-0.0468	0.0168	0.00626	0.0195	-0.0813	0.0296	0.0106	0.0330
			(0.0399)	(0.0307)	(0.0228)	(0.0144)	(0.0695)	(0.0540)	(0.0386)	(0.0244)
		MS	0.00830	-0.000819	0.00224	-0.00330	0.0138	-0.00133	0.00349	-0.00504
			(0.0191)	(0.0154)	(0.0111)	(0.00777)	(0.0317)	(0.0249)	(0.0173)	(0.0118)
		HS	-0.0225	-0.0385	-0.0259	-0.0177	-0.0311	-0.0503	-0.0350	-0.0240
			(0.0319)	(0.0244)	(0.0160)	(0.0112)	(0.0441)	(0.0318)	(0.0216)	(0.0152)
	CELDT Era 2 ('07-'12)	Elm	0.0229	0.0330	0.00311	0.00662	0.0411	0.0667	0.00622	0.0121
			(0.0254)	(0.0191)	(0.0139)	(0.00942)	(0.0458)	(0.0391)	(0.0278)	(0.0172)
		MS	-0.00824	-0.0169	-0.0102	0.00146	-0.0123	-0.0236	-0.0145	0.00222
			(0.0226)	(0.0165)	(0.0120)	(0.00836)	(0.0336)	(0.0231)	(0.0169)	(0.0127)
		HS	-0.0218	-0.0269	-0.0223	-0.00346	-0.0458	-0.0536	-0.0434	-0.00629
			(0.0207)	(0.0155)	(0.0118)	(0.00782)	(0.0438)	(0.0312)	(0.0231)	(0.0142)
BS in LE 5yrs	CELDT Era 1 ('03-'06)	Elm	0.00485	0.0529*	0.0251	0.0162	0.00843	0.0932*	0.0427	0.0274
			(0.0380)	(0.0253)	(0.0190)	(0.0130)	(0.0659)	(0.0452)	(0.0323)	(0.0220)
		MS	0.0366**	0.0294*	0.0204*	0.0197**	0.0608**	0.0476*	0.0318*	0.0301**
			(0.0131)	(0.0115)	(0.00924)	(0.00663)	(0.0222)	(0.0188)	(0.0145)	(0.0101)
		HS	0.0126	-0.0301	-0.0152	-0.000887	0.0174	-0.0392	-0.0205	-0.00120
			(0.0189)	(0.0157)	(0.0130)	(0.00955)	(0.0262)	(0.0203)	(0.0173)	(0.0129)
	CELDT Era 2 ('07-'12)	Elm	-0.0398	-0.0247	0.00332	0.00867	-0.0716	-0.0499	0.00665	0.0158
			(0.0238)	(0.0169)	(0.0118)	(0.00769)	(0.0443)	(0.0349)	(0.0235)	(0.0140)
		MS	-0.0166	-0.0120	-0.00370	-0.00419	-0.0246	-0.0168	-0.00524	-0.00638
			(0.0176)	(0.0128)	(0.00968)	(0.00691)	(0.0263)	(0.0179)	(0.0137)	(0.0105)

		HS	0.0167	0.00902	0.0116	-0.00825	0.0351	0.0180	0.0226	-0.0150
			(0.0133)	(0.0116)	(0.00994)	(0.00671)	(0.0282)	(0.0231)	(0.0193)	(0.0122)

Table 7
Robustness of the Estimates to the Order of Polynomials in the Running Variable CELDT, Bandwidth=50. CELDT reading.

		Grade	Intent to Treat				Treatment on Treated			
Outcome	Era	Span	4th	3rd	2nd	1st	4th	3rd	2nd	1st
Graduation on time	CELDT Era 2 ('07-'12)	Elm	0.116	0.0404	0.0557	0.0606*	0.264	0.106	0.187	0.192*
			(0.0810)	(0.0562)	(0.0404)	(0.0248)	(0.199)	(0.150)	(0.140)	(0.0811)
		MS/HS	-0.0106	-0.0295	-0.0129	-0.00310	-0.0547	-0.135	-0.0550	-0.0121
			(0.0437)	(0.0349)	(0.0259)	(0.0178)	(0.226)	(0.162)	(0.110)	(0.0691)
Postsecondary 2 or 4yr school, 1st year	CELDT Era 2 ('07-'12)	Elm	0.0750	0.0428	0.100**	0.0600*	0.311	0.162	0.362**	0.198*
			(0.0633)	(0.0478)	(0.0354)	(0.0242)	(0.272)	(0.183)	(0.134)	(0.0804)
		MS/HS	0.0587	0.0397	0.0291	0.0107	0.258	0.168	0.113	0.0398
			(0.0460)	(0.0367)	(0.0276)	(0.0187)	(0.208)	(0.158)	(0.108)	(0.0700)
Postsecondary 2 or 4yr school, 2nd year	CELDT Era 2 ('07-'12)	Elm	0.102	0.0568	0.107**	0.0614*	0.423	0.215	0.387**	0.202*
			(0.0641)	(0.0482)	(0.0356)	(0.0244)	(0.282)	(0.185)	(0.135)	(0.0809)
		MS/HS	0.0820	0.0461	0.0409	0.0147	0.360	0.195	0.159	0.0548
			(0.0457)	(0.0365)	(0.0273)	(0.0186)	(0.213)	(0.157)	(0.108)	(0.0694)
Postsecondary 2yr school, 1st year	CELDT Era 2 ('07-'12)	Elm	0.0159	-0.0112	0.0334	0.0159	0.0662	-0.0424	0.121	0.0523
			(0.0591)	(0.0456)	(0.0341)	(0.0234)	(0.245)	(0.172)	(0.124)	(0.0768)
		MS/HS	0.0787	0.0636	0.0545*	0.0260	0.345	0.269	0.212	0.0970
			(0.0456)	(0.0363)	(0.0272)	(0.0185)	(0.212)	(0.159)	(0.108)	(0.0693)
Postsecondary 2yr school, 2nd year	CELDT Era 2 ('07-'12)	Elm	0.0623	0.0344	0.0724*	0.0181	0.258	0.130	0.262*	0.0594
			(0.0611)	(0.0466)	(0.0346)	(0.0236)	(0.261)	(0.178)	(0.128)	(0.0777)
		MS/HS	0.108*	0.0665	0.0640*	0.0263	0.473*	0.281	0.249*	0.0982
			(0.0451)	(0.0360)	(0.0270)	(0.0183)	(0.219)	(0.158)	(0.109)	(0.0689)

		Grade	Intent to Treat				Treatment on Treated			
Outcome	Era	Span	4th	3rd	2nd	1st	4th	3rd	2nd	1st
Postsecondary 4yr school, 1st year	CELDT Era 2 ('07-'12)	Elm	0.0537	0.0333	0.0600*	0.0446*	0.223	0.126	0.217*	0.147*
			(0.0513)	(0.0393)	(0.0290)	(0.0201)	(0.219)	(0.150)	(0.107)	(0.0668)
		MS/HS	-0.0142	-0.0152	-0.0231	-0.0227*	-0.0622	-0.0642	-0.0897	-0.0846
			(0.0300)	(0.0235)	(0.0174)	(0.0115)	(0.132)	(0.0995)	(0.0679)	(0.0433)
Postsecondary 4yr school, 2nd year	CELDT Era 2 ('07-'12)	Elm	0.0698	0.0365	0.0506	0.0509**	0.290	0.138	0.183	0.167**
			(0.0501)	(0.0383)	(0.0281)	(0.0194)	(0.216)	(0.146)	(0.103)	(0.0646)
		MS/HS	-0.0220	-0.0195	-0.0240	-0.0184	-0.0966	-0.0825	-0.0932	-0.0689
			(0.0281)	(0.0218)	(0.0160)	(0.0105)	(0.125)	(0.0932)	(0.0630)	(0.0397)
AS in LE 4yrs	CELDT Era 2 ('07-'12)	Elm	0.0107	-0.00657	-0.00200	0.0130	0.0443	-0.0249	-0.00725	0.0429
			(0.0190)	(0.0149)	(0.0117)	(0.00834)	(0.0794)	(0.0564)	(0.0421)	(0.0275)
		MS/HS	-0.0198	-0.0227	-0.00839	0.000292	-0.0867	-0.0959	-0.0326	0.00109
			(0.0162)	(0.0130)	(0.00993)	(0.00690)	(0.0727)	(0.0564)	(0.0387)	(0.0258)
BS in LE 4yrs	CELDT Era 2 ('07-'12)	Elm	0.00825	0.0211*	0.0282**	0.0169*	0.0342	0.0798	0.102**	0.0555*
			(0.0133)	(0.0102)	(0.00871)	(0.00678)	(0.0558)	(0.0408)	(0.0335)	(0.0226)
		MS/HS	-0.00926	-0.0106	-0.0151*	-0.00802	-0.0407	-0.0450	-0.0587*	-0.0300
			(0.00987)	(0.00859)	(0.00653)	(0.00439)	(0.0440)	(0.0370)	(0.0261)	(0.0165)
AS in LE 5yrs	CELDT Era 2 ('07-'12)	Elm	0.0162	-0.00193	0.00416	0.0156	0.0674	-0.00732	0.0150	0.0512
			(0.0196)	(0.0154)	(0.0122)	(0.00870)	(0.0827)	(0.0584)	(0.0440)	(0.0287)
		MS/HS	-0.0226	-0.0190	-0.00841	0.00549	-0.0994	-0.0804	-0.0327	0.0205
			(0.0180)	(0.0145)	(0.0109)	(0.00765)	(0.0813)	(0.0622)	(0.0427)	(0.0286)
BS in LE 5yrs	CELDT Era 2 ('07-'12)	Elm	0.0272	0.0364**	0.0271**	0.0130	0.113	0.138*	0.0980*	0.0428
			(0.0160)	(0.0129)	(0.0104)	(0.00781)	(0.0718)	(0.0537)	(0.0392)	(0.0259)
		MS/HS	0.00622	0.00486	-0.00383	-0.000203	0.0273	0.0205	-0.0149	-0.000757
			(0.0158)	(0.0129)	(0.00991)	(0.00661)	(0.0696)	(0.0548)	(0.0385)	(0.0247)