PPIC PUBLIC POLICY INSTITUTE OF CALIFORNIA

25 YEARS

SEPTEMBER 2019

Julian Betts, Laura Hill, Karen Bachofer, Joseph Hayes, Andrew Lee, and Andrew Zau

Supported with funding by the Institute of Education Sciences, US Department of Education, and the William T. Grant Foundation

English Learner Trajectories and Reclassification

SUMMARY

CONTENTS

Technical appendices to this report are available on the PPIC website.

More than 40 percent of students in California's public schools speak a language other than English at home. In the 2016–17 school year, 21 percent of all students, or more than 1.3 million, were English Learners (ELs). When former English Learners are included, the population of "ever ELs" expands to 38 percent of all K–12 students in the state (CDE Dataquest 2017). A key issue for California's K–12 schools is when to reclassify English Learner (EL) students as English Proficient. If they are reclassified too soon, they may have difficulty handling core academic classes. But if they wait too long, they may be deprived of subject matter they are capable of handling. ELs have historically performed far below proficient English speakers. However, former ELs are some of the strongest performers, sometimes scoring better than native English speakers on standardized tests.

Across California's public schools, the decision of when to reclassify ELs is guided by district-specific policies. California Education Code 313 (f) establishes four criteria for reclassification procedures, which provide overarching direction on reclassification. However, the criteria laid out in these policies have not been uniform across districts, nor static over time. That is now changing. Under the federal Every Student Succeeds Act (ESSA), the state's approximately 1,000 school districts must standardize reclassification policies. Thus, it is important to understand how reclassification policies have worked in the past if we are to put in place effective policies that reclassify ELs at the optimal time.

This report describes research carried out over two distinct reclassification eras between the 2002–03 and 2013–14 academic years in the Los Angeles Unified (LAUSD) and San Diego Unified School Districts (SDUSD). Using student-level data and the multiple reclassification policies, we are able to discern whether reclassification takes place at the optimal time for student success. We compare academic outcomes of students just below and just above the threshold for meeting reclassification requirements, and when we find that students' academic trajectories continue at the same pace, we determine that reclassification policies are set appropriately.

Overall, reclassification criteria appear mostly to have been appropriate, even though reclassification policies underwent changes during the period of our study and varied across the two districts. More specifically, we find

- In SDUSD, there is no strong evidence that EL students were reclassified either too soon or too late. In other words, students who just met reclassification criteria and those who narrowly missed reclassification later performed about the same.

- In LAUSD, in a few instances, some elementary school students may have been reclassified too soon, resulting in short-term negative outcomes. There is also some evidence of negative effects on on-time graduation for reclassified ELs at the high school level, but only in the first reclassification era. In all other cases, reclassified students performed neither substantially better nor worse than ELs whose language scores were just below the reclassification threshold. These findings provide useful guidance about what appropriate reclassification criteria might look like as standardized statewide criteria are set.
- In both districts, it does not appear that student characteristics such as length of time as an EL student, school demographic patterns, or neighborhood characteristics altered the impact of reclassification on students, with some minor exceptions. This suggests reclassification policies need not be adapted for different types of students or school contexts.

Because we find that students fared similarly under two different reclassification regimes in San Diego and that reclassified students in LAUSD and SDUSD fared well even though the districts' reclassification policies were not identical, this suggests that the context of EL instruction and conditions for reclassified students may be critical to understanding reclassification's impact. We did not study the details of the English Learner instruction and reclassified student monitoring. Once all reclassification criteria are standardized across the state's districts as recently required by ESSA, we could see differential impacts from reclassification by district. We recommend that state policymakers establish procedures for monitoring EL students to affirm that the new English language proficiency assessment and any new criteria based on the Smarter Balanced Assessment Consortium (SBAC) are not so rigorous as to prevent students who could succeed from being reclassified. And, as state law now requires, policies must be adopted to ensure all students, including ELs, receive core academic instruction.

Introduction

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 certain threshold on California's English proficiency test.[1] EL status is meant to be temporary—when students demonstrate sufficient English language proficiency, their official designation changes to Reclassified Fluent English Proficient.

English Learners should have English language development instruction, either in standalone English language development settings (designated) or as part of subject matter content (integrated). EL students receive core subject instruction, but the requirement that they receive English language development instruction may mean they cover less core content than other students. After they are reclassified, ELs no longer receive English language development instruction and take core subjects without additional support.[2]

To be reclassified, students must demonstrate English language proficiency on state assessments administered only to EL students. They must also demonstrate basic skills in English comparable to those of native English speakers. Individual school districts make reclassification decisions based on criteria recommended by state law and regulation and, in some cases, districts use additional criteria.

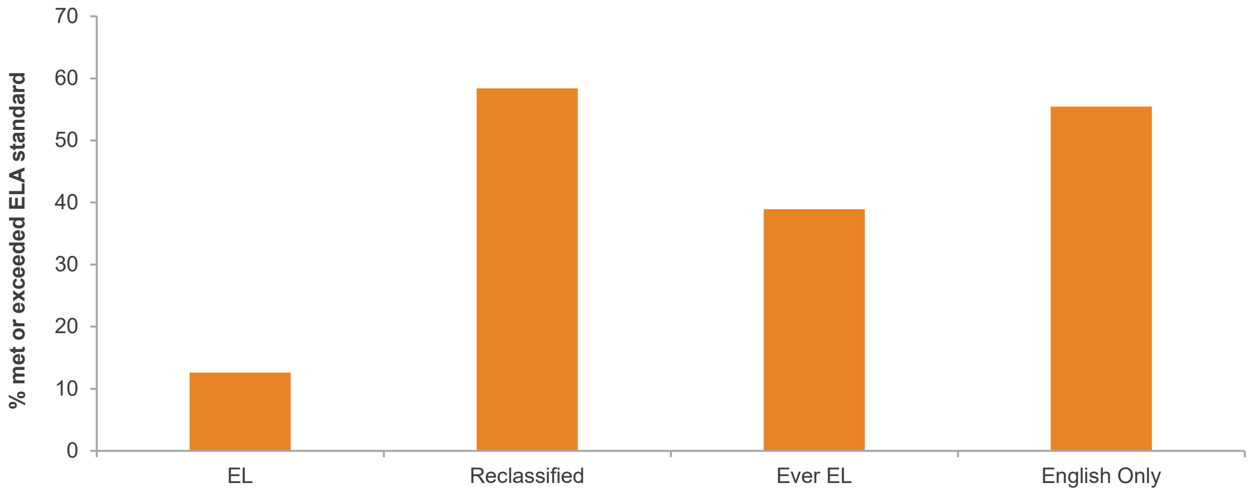
Student data have long shown that EL students do not do as well academically as their non-EL peers. However, reclassified students perform much better than current EL students on average, as measured by test scores, completion of college preparatory courses, and graduation rates. Moreover, they sometimes perform better than students who only speak English at home (Figure 1). Studies that just examine the association between being reclassified and student outcomes will not help us understand what causes the link between reclassification and positive student outcomes. Does becoming reclassified somehow improve student performance? Or do the reclassification criteria effectively distinguish between students who need additional English language support and those who do not? Without clarity about reclassification's causal effects, it is difficult to design reclassification policies that provide students English language support for the optimal length of time.

---

[1] California replaced the California English Language Development Test (CELDT) with the English Language Proficiency Assessments for California (ELPAC) in 2017–18. In January 2019, it became the required assessment to measure English proficiency for reclassification, with mandatory cutpoints.
[2] AB 2735, effective January 1, 2019, requires that EL students have access to core academic content.

FIGURE 1

Reclassified students meet English language standards at high rates compared with other student language groups



SOURCE: 2018 Smarter Balanced Assessments, CDE.

NOTE: All tested grades (3–8, 11) for all of California. The third bar combines current and reclassified students. The fourth bar, English Only students, are students who do not speak a language other than English at home.

This report uses student-level data from Los Angeles Unified (LAUSD) and San Diego Unified (SDUSD), California's two largest school districts, to evaluate reclassification's effects on academic performance of former ELs. In both California and the United States as a whole, Spanish is the primary language of most ELs.[3] Nonetheless, the two districts have very diverse student populations, reflecting the growing heterogeneity of US and California student populations. In SDUSD, 61 percent of students are low-income compared with 77 percent in LAUSD. The SDUSD student racial/ethnic distribution is closer to that of the overall state student population, while LAUSD has a higher share of Latino and a lower share of white students.

We use a form of statistical analysis known as regression discontinuity to examine academic outcomes for students just above and just below the thresholds used for EL reclassification (see technical appendices for details). This methodology allows us to compare two groups of students who appear very similar by most performance measures—specifically, students who qualified on all but one criterion and scored very close to the cutoff point on the final criterion, with the first group nearly passing and the second group nearly failing to meet the requirement. This approach allows us to estimate the causal effects of reclassification.

Using student-level data, we investigated whether reclassification criteria were set at the right degree of difficulty to ensure a smooth transition from EL to reclassified status. Making the switch at the right time is critical because ELs take courses that support non-native speakers, while reclassified students are placed in core content classes without additional support. If ELs change status too soon, they may have difficulty handling core classes. But if they wait too long, they may be deprived of subject matter they are capable of handling.[4] Our methodology allowed us not only to estimate how the duration of English language support drove student outcomes, but also determine which reclassification standards came closest to reclassifying students at the appropriate time.

[3] Over 80 percent of all ELs in California are Spanish speaking.

[4] Research from Estrada (2014) and Callahan (2005) suggests than many EL students lack access to all core academic content. A recent report from the Migration Policy Institute (Sugarman 2019) finds that four years might not be enough time for high school ELs to graduate. This is likely especially true in districts like LAUSD and SDUSD that, starting with the class of 2016, now require completing college preparatory "a–g" courses (with grades of at least "D") in order to graduate.

In addition, reclassification policy changes in LAUSD and SDUSD provide a unique opportunity to understand how reclassification in general and specific reclassification policies in particular can affect academic performance of former ELs. Because both districts changed reclassification policies, we were able to study four distinct reclassification regimes. We refer to the distinct regimes before and after these shifts as "eras." (See Table 1 for detailed characterizations of each era.)

This is an important time to consider how to measure when ELs are prepared to fully integrate into academic courses without English language support. Implementation of the Common Core State Standards and new English language development standards are underway, though unevenly, across the state (Gao, Lafortune, and Lee, forthcoming 2019; Warren and Murphy 2014; McLaughlin, Glaab, and Carrasco 2014). Also, the assessments that we used to determine how students are meeting the standards have changed. The Smarter Balanced Assessment Consortium (SBAC) test, the new statewide standards test given to all students, was administered for the first time in 2014–15. First-year results were particularly alarming for EL students (Hill and Ugo 2016), so determining appropriate cutpoints for this assessment as a reclassification tool will be important. The English Language Proficiency Assessments for California (ELPAC), the test of English proficiency for EL students that replaces the California English Language Development Test (CELDT), was fully implemented in the 2017–18 school year. This means that reclassification policies are necessarily in a time of transition.

In addition, the federal Every Student Succeeds Act (ESSA) puts reclassification policies into sharper focus. States are now expected to use standardized EL reclassification criteria in all school districts. Yet in each school year, starting in 2014–15, there has been a major shift in reclassification policy as old assessments ended and new ones were implemented. State guidance has allowed districts significant flexibility during this period. The trend continued in 2018–19 with the state issuing final guidance about reclassification cutpoints for the ELPAC.

This report starts with an overview of LAUSD and SDUSD reclassification policies. It then considers how to determine the causal effects of different reclassification policies on student outcomes using the regression discontinuity methodology. The report looks at demographic, linguistic, and other factors that may influence reclassification effects and reviews the student outcomes used in our analysis. We then present the results of our analysis of whether LAUSD and SDUSD EL students were reclassified at the optimal time during the four reclassification eras. Whether reclassification has different effects on long-term and late-arriving ELs, and how student, school, and neighborhood characteristics moderate reclassification are considered. The report ends with conclusions and recommendations.

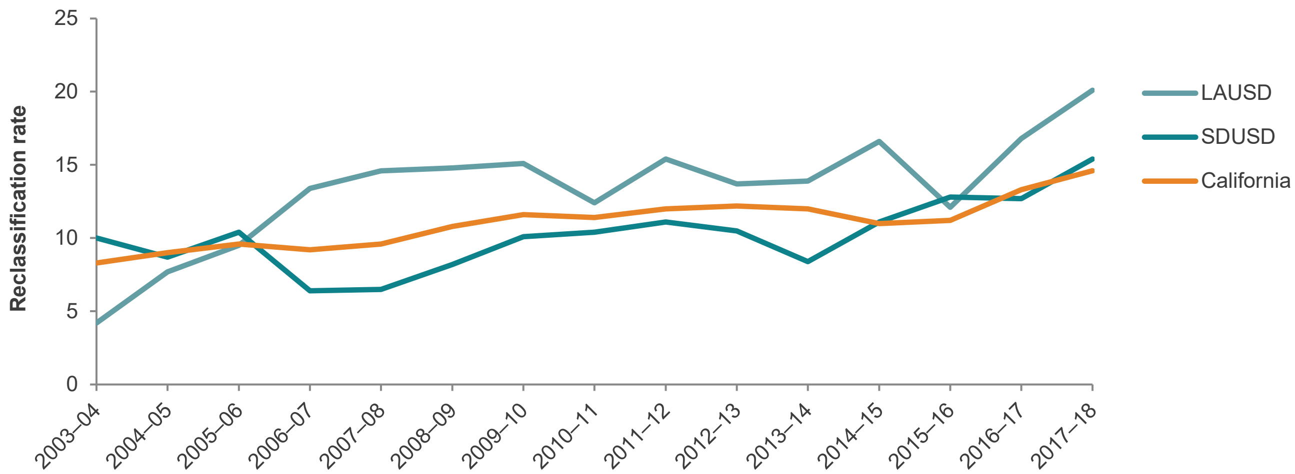
When Do English Learners No Longer Need Language Support?

There is a clear rationale for providing language support to English Learners. Standard classes presume students are able to understand and speak English, and that students are roughly at grade level in reading and writing. By contrast, EL programs are based on the idea that language acquisition is slower for ELs placed in traditional classes because facility with English is needed to understand teachers and curricular materials. Consequently, ELs should be put in special classes with additional language support. However, as proficiency with English grows, at some juncture ELs may perform as well as or better than native English speakers in traditional classes.

Figure 2 shows reclassification trends over the past 15 years in LAUSD, SDUSD, and the state. In all three, reclassification rates show a positive trend. One noticeable change occurred in LAUSD in 2006–07 when the district stopped using math grades as a reclassification requirement. Despite a move to a more rigorous CELDT that year, LAUSD reclassification rates rose 3.8 percentage points. Conversely, one of the few drops in the SDUSD reclassification rate occurred the same year, perhaps reflecting the renorming of the CELDT.

FIGURE 2

Reclassification rates are rising in California, LAUSD, and SDUSD



SOURCE: Authors' tabulations from CDE data, 2003–04 through 2017–18.

California's guidelines establish four criteria for reclassification (California Education Code, Section 313(f)). During our research, the reclassification guidelines included:

1. the California English Language Development Test (CELDT), with recommended overall and subtest scores;
2. the California Standards Test of English Language Arts (CST ELA), which tests basic English skills, with a recommended minimum score;
3. teacher evaluation;
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 state law, regulations, and guidance. Before January 2019, districts decided how to implement the four state reclassification criteria by setting local policy, but since January 2019, the English proficiency criterion has been standardized statewide.[5] The policies and practices for identifying EL students, assigning them to instructional programs, and reclassifying them are articulated in each school district's *Master Plan for English Learners*.[6]

[5] 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).

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

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 minimum scores on the CST ELA or the CELDT, the CST math test as another measure of basic skills, or course marks for teacher evaluation. The survey also suggested parental consultation was of limited importance.

SDUSD and LAUSD both have stricter standards than the minimum state guidelines, but their rigor differs (*Master Plan*, LAUSD 2012 and SDUSD 2009). These differences provide a unique opportunity to examine the relationship between reclassification criteria and reclassified student outcomes. Earlier research found that more-rigorous criteria in SDUSD and LAUSD were associated with roughly the same or better academic outcomes for reclassified students than those of native English speakers (Hill et al. 2014). During the period studied, LAUSD had more stringent criteria than SDUSD, which were associated with lower reclassification rates and stronger outcomes for reclassified students in LAUSD.[7]

LAUSD and SDUSD have had their English language instructional programs in place for over a decade and their reclassification policies remained the same from 2006 through the last years in which we study reclassification decisions—spring 2012 in LAUSD and spring 2014 in SDUSD. Different policies were in place in both districts from 2002–06. For example, SDUSD raised its basic skills reclassification criterion for students at all grade levels in 2005–06 and then lowered it to the earlier level, but only for secondary students, in 2006–07 and later years (Table 1.)

TABLE 1

Reclassification eras changed in both districts during study period

	Los Angeles		San Diego
2003–06	Basic ELA skills = 300+ CST ELA English proficiency = "Early Advanced" overall English proficiency = "Intermediate"+ subtests Teacher evaluation = 3+ or C or better in English *and* Math courses	**2002–05**	Basic ELA skills = 300+ CST ELA English proficiency = "Early Advanced" overall English proficiency = no more than one "Intermediate" on subtests
		2005–06	Increased CST ELA threshold to 333+
New CELDT era starting in 2006–07 school year			
2006–12	Dropped math course requirement	**2006–14**	333+ CST ELA for elementary, 300+ grades 6–12
2012–13	Allowed English course *or* Advanced ESL		
2013–14	Used 2012–13 CST *or* High school exit exam (grades 9–12) *or* Elementary reading assessment		

SOURCES: Los Angeles Unified School District. Reclassification of English Learners, Grades 2–12, Bulletin BUL-5619.0, October 17, 2011. San Diego Unified School District, Master Plan for English Learners, 2009

NOTES: CELDT proficiency cutoffs remained the same, but cut scores for proficiency levels were raised across all grades and subtests (CDE, 2007).

[7] Comparing reclassification stringency in the two districts was not easy. SDUSD had considerably higher standards than LAUSD on the CELDT test, but LAUSD required ELs to obtain sufficiently high grades in both math classes (up to 2005–06) and English classes to be reclassified. SDUSD did not set cutoffs in subject grades. The different results in the two districts could have derived from changes in their EL student populations not related to the reclassification criteria and from the effect of reclassifying students at different points in their academic progress.

Similarly, the LAUSD criterion for teacher evaluation of EL students specified minimum math marks from 2002–03 to 2005–06, but subsequently dropped the requirement. Both districts were subject to the state's rescaling of its English language proficiency test (the CELDT test) in 2006–07, which set a higher bar for reclassification by requiring higher scores to reach the Early Advanced level. Reclassification rates registered a small drop statewide in 2006–07, a larger drop in SDUSD, but in LAUSD the rate rose, probably due to the district's decision to stop using math grades as a reclassification criterion.

Effects of Reclassification on Student Outcomes

Ideally, students are reclassified when they no longer need EL support and are ready for an English-only instructional program. Because reclassification policies vary widely across the state, it is unlikely that all school districts reclassify students at the optimal time.

Previous research has found that reclassified ELs are among the best performing students according to a variety of academic measures (Hill et al. 2014; Saunders and Marcelletti 2013; Gándara and Rumberger 2006; EdSource 2008; Flores, Painter, and Pachon 2009). However, this does not indicate that reclassification itself improves outcomes. For example, when the new CELDT increased reclassification standards in 2006–07, the change may have artificially improved reclassified student outcomes by keeping lower-performing students—who in earlier years would have been reclassified—as English Learners in the new era. In contrast to these earlier studies, our method allows us to understand whether reclassified students do better as a result of leaving English Learner status or if reclassification criteria just screen for high-performing students.

This report looks at all LAUSD and SDUSD EL students in grades 3 through 12 between 2001–02 and 2015–16, and examines EL reclassification decisions in LAUSD between 2002–03 and 2011–12, and in SDUSD between 2003–04 and 2013–14. For both districts, data through 2015–16 are used to measure outcomes. In LAUSD, the mean number of years spent in US schools before reclassification was 6.1 during the study period. In SDUSD, the mean number of years since entering the district was 4.6. In both districts, almost exactly half of all students who were ever ELs were reclassified by grade 6. Because we control for student characteristics from the year before reclassification and measure outcomes one or more years after reclassification, we require that students be in the district three consecutive years.

Ours is one of only a few studies that has had access to the student-level data needed to establish causality. 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. 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. Robinson-Cimpian and Thompson (2016) investigated two LAUSD reclassification policy eras. This study looked at only one of five reclassification criteria, and therefore did not fully evaluate whether the overall policy in these periods was set appropriately.[8]

---

[8] As mentioned earlier, this paper is among the first to address whether reclassification causes better outcomes. We highlight here the similarities and differences between our report and the Robinson-Cimpian and Thompson paper (2016), because its focus is LAUSD and for similar school years. This important paper focuses on the CST ELA Regression Discontinuity experiment, but uses fewer years (2002–03 to 2009–10) than our current study. In addition, our paper is different in that we separate elementary and middle school students, running separate experiments for elementary, middle, and high school grades. The Robinson-Cimpian and Thompson

We are not asking whether EL students benefit from additional English language support. Rather, we are asking the more nuanced but critical policy question of whether ELs were reclassified at the optimal time. Our regression discontinuity design allows us to determine the causal effects of reclassification by using the threshold above which reclassification occurs and below which it does not. If we see students who met reclassification criteria in a given year subsequently perform better or worse than otherwise similar students who narrowly missed the cutoff, this is evidence that ELs were reclassified either too late or too early. If those students performed similarly, that indicates the reclassification criteria were appropriate.

In this way, our methodology approximates a true experiment with a treatment group and a control group. Reclassification's effects can be observed by examining outcomes on either side of the EL reclassification cutoff. The differences in underlying achievement between students just above and just below the reclassification cutoff should be small, so any differences in outcomes between those above and those below are likely to be caused by reclassification.

Because the two districts used various test scores in their reclassification decisions, we can test whether the cutpoints for these separate tests were appropriate. The components of our analysis include the CST cutoffs, the CELDT Overall Proficiency Level (OPL) in LAUSD only, and the CELDT subtest for reading. We treat each component separately.[9]

Figure 3 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 too late and reclassification criteria were too rigorous. If we find no statistically significant difference in test score gain post reclassification, then reclassification criteria were optimally set for student progress.

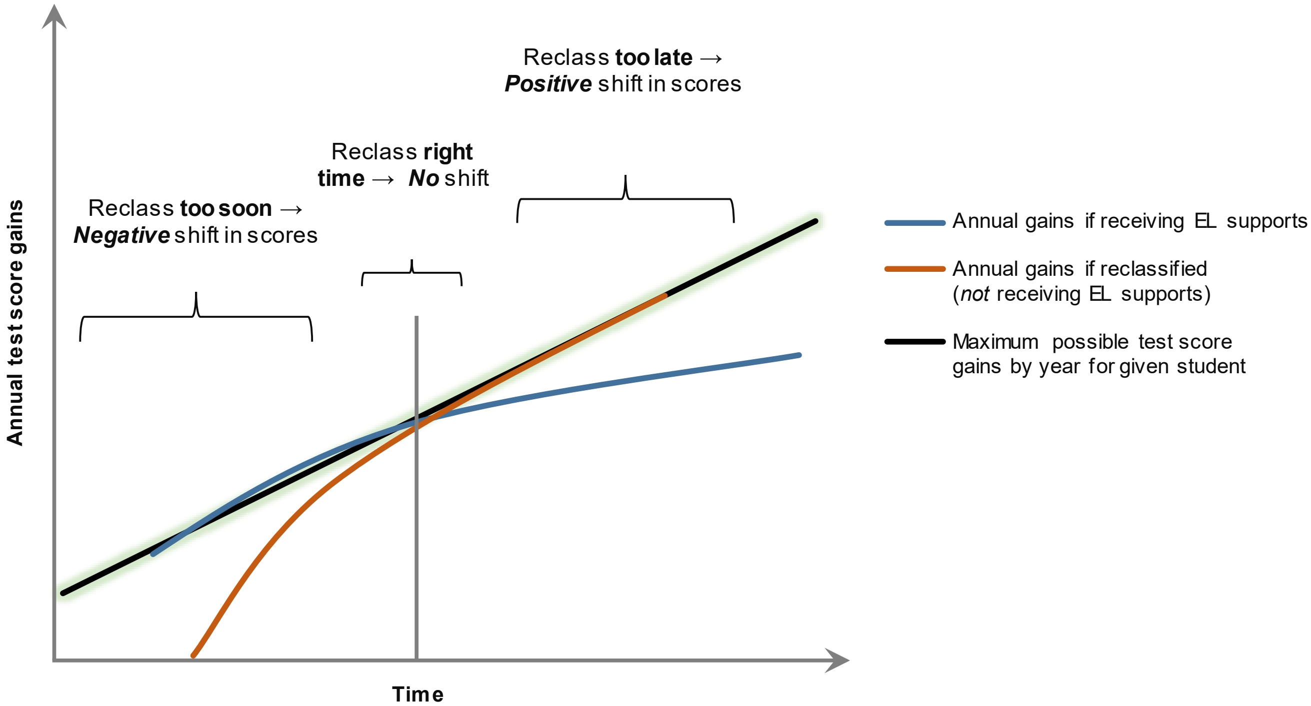
---

paper focuses on only Latino students, while our analyses include EL students of all ethnicities, while controlling for whether a student's home language is Spanish. Probably the most important difference is that we also account for the teacher recommendation component of LAUSD's reclassification policy (i.e., the requirement of math and ELA course marks through 2005–06 and ELA course marks from 2006–07 onwards). Finally, our report includes another important school district: SDUSD.

[9] Each of these components measures some aspect of English, so it might be expected that these test scores would be highly correlated and that few students would meet the requirements on one test but not another. This does not turn out to be an issue. The correlations between the CST and the CELDT OPL and subtests were positive, but relatively low for each grade span and era. In both districts, the highest correlation was for high school students between the CST and the CELDT reading subtest during the first reclassification era. More typically, correlations between the CST and components of the CELDT were substantially lower.

FIGURE 3

Test scores rise, fall, or stay the same depending on whether reclassification occurs at the right time



SOURCE: Authors' calculations.

NOTE: Hypothetical effects of EL reclassification.

What Factors Influence the Impact of Reclassification?

California's EL population is diverse, varying in parental background and representing more than 60 different home languages (Hill 2018). The home languages spoken by students and their proficiency in those languages could influence reclassification effects and affect student performance (Friesen and Krauth 2011; Hill et al. 2014; Salazar and Hayes 2010). In addition, factors at the school and neighborhood level may attenuate or amplify reclassification effects.[10] For example, grade level at reclassification is correlated with student outcomes (Hill, Weston, and Hayes 2014; Hill et al. 2014).

Previous research has found that students reclassified under more rigorous criteria in elementary school grades usually experience slightly better outcomes than those reclassified under less rigorous criteria (Hill et al. 2014; Kim and Herman 2012). But those reclassified at older ages using more rigorous criteria may not do better (Hill, Weston, and Hayes 2014). Again, these are observational studies and say nothing about causation.

Having more ELs of the same home language in a school could potentially make learning English more difficult because students have fewer opportunities to practice English. Alternatively, having a critical mass of ELs speaking the same language could mean that a district or school would be more likely to invest in teacher training

[10] Friesen and Krauth (2011) find that being in a school with many other Chinese-speaking ELs is beneficial for Chinese-speaking ELs but that for Punjabi-speaking ELs, higher concentrations of Punjabi speakers is associated with lower academic outcomes.

that specifically supports that language group.[11] Neighborhood concentration of non-English speakers could interact with school concentrations of ELs (Conger, Schwartz, and Steifel 2011).[12]

Furthermore, EL instructional programs vary across California districts (Parrish et al. 2006). Time to reclassification differs and can be affected by language spoken (Umansky and Reardon 2014). Betts, Zau, and Rice (2003) find that having teachers with a Crosscultural, Language, and Academic Development (CLAD) certificate or the Bilingual CLAD is not associated with better EL student outcomes. However, at the secondary level, teacher qualification was in some cases associated with gains in overall student learning.

Another variable that could influence the impact of reclassification is proper assignment of English language development (ELD) coursework. ELs should receive ELD instruction, but after reclassification they should no longer get it. A question for research is whether the percentage of a school's ELs not assigned ELD coursework influences the effects of reclassification.

Educational Outcomes That Reclassification May Affect

The outcomes we looked at include standardized test scores, which are available for students at various grade levels, and high school graduation.[13] Because the goal of EL programs is to improve student language ability so they can perform well in classes designed for native English speakers, some of our main outcomes involve English Language Arts (ELA) performance. Our chief ELA-skills measure is the grade-specific ELA CST.[14] From 2001–02 through 2012–13, California administered the CST to all students in grades 2 through 11, except some special education students and recently arrived immigrants.

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

Those EL students who narrowly miss reclassification are likely to be reclassified during the next one to three years, during which they continue to receive EL support.[15] The most obvious time frame for measuring ELA performance may be during the first full school year after reclassification.[16] But because reclassification might require short-run adjustment for a student but have no longer-run consequences, we also model test scores two years after the reclassification year.

[11] Using propensity score matching, Callahan et al. (2009) find that greater concentrations of EL students are associated with stronger academic achievement relative to lower concentrations. Bui (2013) finds that increasing shares of EL students are associated with increased reclassification, less grade retention, and increased math scores, but not reading and language scores.

[12] Peer effects have been shown to be a factor in student achievement, but measuring the effects is complicated by the potential for unobservable factors to determine both neighborhood and school concentrations (Hoxby 2000; Hanushek, Kain, and Rivkin 2002). The relationship between ELs and peer achievement is even more complicated.

[13] Future work will examine college attendance, persistence, and graduation.

[14] We convert the CST scores in years after the reclassification decision into Z-scores (standardized using the statewide mean and standard deviation of CST scores). This puts the scores onto a common scale across grades, which also makes coefficient sizes quite interpretable. (An increase of one point on this standardized version of the CST means that a student has moved up one standard deviation in the rankings of the scores of all students in the given grade statewide.)

[15] Recall that to be included in an experiment, students must have met all but one criterion for reclassification, and on the last criterion, they must be close to the cutpoint (either just above or just below).

[16] Students can be reclassified partway through a school year (most typically in spring in SDUSD and throughout the year in LAUSD), so by examining outcomes a year after reclassification, we can be fairly certain that a student received the reclassification "treatment" for the entire year. This also reduces concerns that student achievement growth could drop in the weeks immediately after reclassification simply because the student's daily routine has changed. Some students might experience a mid-year change in classes as a psychological shock requiring time to adjust.

On-time graduation can be measured based on the expected year of graduation according to the grade in which students are enrolled when their records are included in our statistical analysis for that year. For those above grade 9, projected graduation date is based on the year they first enrolled in grade 9.[17]

Main Findings from the Effect of Being Reclassified

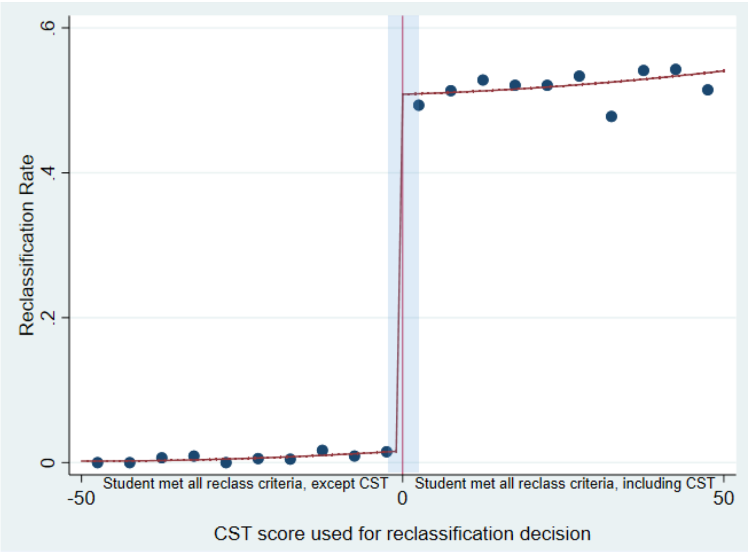
Our regression discontinuity approach is very powerful but only if certain conditions apply. In order for it to approximate a true experiment, we have to determine whether district policies are implemented in such ways as to simulate a treatment and a control group. To do this, we examine students who have met all reclassification criteria except one. For example, if the one remaining criterion is the CST ELA test, we refer to scores on this test as the running variable and plot the probability of being reclassified by the value of the score on it. If districts are adhering to their stated reclassification policy, we should see a marked change in a student's probability of being reclassified based on the value of this running variable.

Figure 4 shows one example of how dramatically that probability changes based on a test score falling on either side of the reclassification cutpoint in LAUSD, grades 3–5, during the first era, 2003–06. The key to understanding this figure is that on the horizontal axis, a test score of zero shows the cutpoint for reclassification. Students below the cutpoint are rarely reclassified, but the reclassification probability of students above it jumps 49 percentage points. (The dots represent students grouped by CST test scores. For example, the dot to the right of the cutpoint represents all students who scored between 0 and 5 points above the CST cutpoint. The line represents the best fit to the underlying student data). The figure demonstrates that the LAUSD reclassification criteria applied to the CST requirement reasonably simulate an experiment with a treatment. The same is true for both school districts and both reclassification eras under consideration (See Technical Appendix B, Tables 1A and 1B). For example, in SDUSD for era 1, the probability of being reclassified for grade 3–5 elementary students who meet the CST requirement is 59 percentage points higher than for students who do not. Also included in Technical Appendix B are the results of further statistical tests. The technical appendices explain what these statistical tests are and why they tell us we have a valid experiment, focusing on the CST cutoff. These tests confirm that students just above or just below the cutpoint are otherwise similar, and that the test scores used in the reclassification decision are not being manipulated.

[17] We cannot distinguish between LAUSD dropouts, who do not graduate on time, and students who leave for another school district. We treat all of these students as not having graduated on time. To determine whether this is a problem, we repeated all the SDUSD analyses of graduation on time. In the main results, we drop those who leave the district and instead treat these students as not having graduated on time. This did not change the SDUSD results much (as shown in the technical appendices).

FIGURE 4

Probability of EL reclassification increases dramatically when ELs meet CST reclassification requirement, controlling for meeting all other reclassification criteria in 2003–06 (LAUSD, grades 3–5, era 1)



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, meaning that students are grouped by test scores into groups 5 points wide. The range of years listed here and in later figures and tables refers to spring of the given school year. For example, the graph for LAUSD for "2007–2012" means that the school years 2006–2007 through 2011–2012 are used.

Results

Have the districts set appropriate cutoffs on the CST?

The key question we investigate is whether the two districts have historically chosen suitable cutoffs on the CST test. We are fortunate to observe two different "eras" in each district, during which each district changed its reclassification criteria—either moving the cutoff point for a particular test or including additional criteria, such as course marks. In both districts, the first era includes years when the older CELDT test was used, while the second includes years when the newer, more rigorous CELDT was in place, allowing us to identify which set of standards comes closest to reclassifying students at the appropriate time.

Our approach allows us to estimate the causal impact of being at or above the CST cutoff point at the elementary, middle, and high school levels.[18] As Figure 4 shows, not all students above the CST cutoff actually get reclassified. To arrive at a causal estimate of reclassification itself, a few additional assumptions are required.[19]

In LAUSD in a few instances it appears that students were reclassified too soon. High school students reclassified in 2003–06 had a five percentage point lower chance of graduating on time than those students near the reclassification cutoffs who remained ELs (see Technical Appendix B, Table 5A).[20] Figure 5 shows the effects of

---

[18] Our definition of just above and just below is a bandwidth of 50 CST points, which is approximately the number of points need to move from one level of proficiency to another.

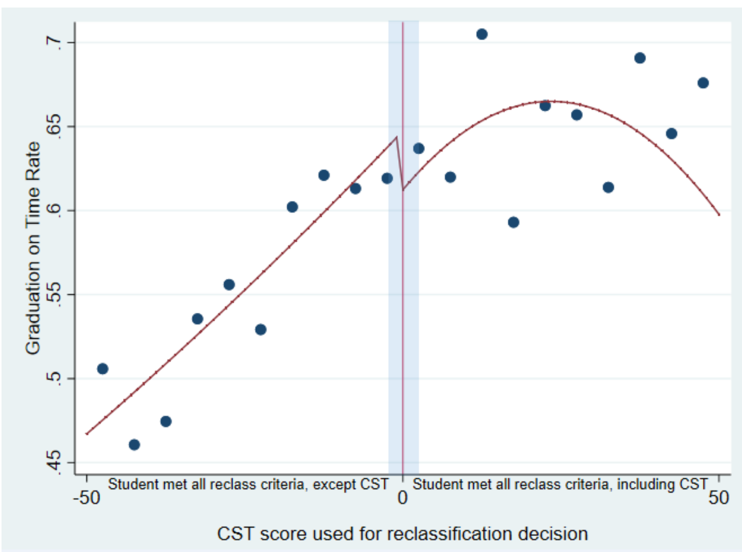
[19] In Technical Appendix B and the many tables there, we refer to the estimates of the impacts of being above the cutoff and of actually being reclassified as the intent-to-treat estimate and the estimate of the impact of treatment on the treated, respectively.

[20] The causal impact of being at or above the cutoff is also negative but about half as large. This effect is larger in the models where we estimate the impact of actually being reclassified than the impact of being at or above the cutoff. We estimated models that allowed the outcomes to depend on the CST score in a simple linear way, but we also estimated models that allowed for a non-linear relationship: quadratic models (also known as second order polynomial models), third, and fourth order polynomial models. Our estimates on being above the reclassification threshold were usually, but not always, consistent. See technical appendix Tables 5A and 5B for the full results of the experiments where we examined the impact of being just above and just below the CST cutpoint. The most common pattern where the four

on-time graduation of being at or above the cutoff. Overall, there is a positive link between CST scores and the on-time graduation rate, which makes sense. But to identify the impact of reclassification our interest lies in whether there is a distinct break in the on-time graduation rate at the cutoff. The shaded region of the graph demonstrates the effect of being reclassified for high school students. There we see that those students just above the CST cutpoint have lower rates of graduating on time than similar students who were just below the cutpoint.[21] Students in the shaded area are nearly identical in characteristics we can observe (our tests to ensure our experiments are valid concluded that they are essentially the same demographically) and all included students have met other reclassification requirements, with the exception of the CST cutpoint. The only differences among students just above and just below the CST cutpoint is that those just above it received a test score high enough to be reclassified. And those just below the cutpoint missed a reclassifying score by one or just a few points. As CST scores increase (x-axis), chances of graduating on time increase, up until the cutpoint. Then just after the cutpoint, chances of high school graduation decrease. Since students just above and below the CST cutpoint are similar in all other ways we can observe, we can attribute the dip in graduation chances to being reclassified. The results suggest that there were longer-term consequences of being reclassified early in LAUSD. Given that many of the high school students who narrowly missed the cutoff were likely reclassified a year or two later, this estimated drop in on-time graduation rates is quite significant from a policy viewpoint. Robinson-Cimpian and Thompson (2016) had similar results for graduation.

FIGURE 5

High school reclassification caused lower on-time graduation rates in 2003–06 (LAUSD, grades 9–12, era 1)



SOURCE: Authors' estimates, see Table 5A in the technical appendix.

NOTE: CST is the running variable and has been rescaled to zero. The lines in the graphs show the regression fit using a quadratic polynomial estimated independently on the two sides of the cutoff. Each dot represents a bin size of 5. The discontinuity was -0.0529 and was significant. Sample size was 22,979. In LAUSD, on-time graduation is calculated for students who persist to 12th grade and does not include dropouts.

---

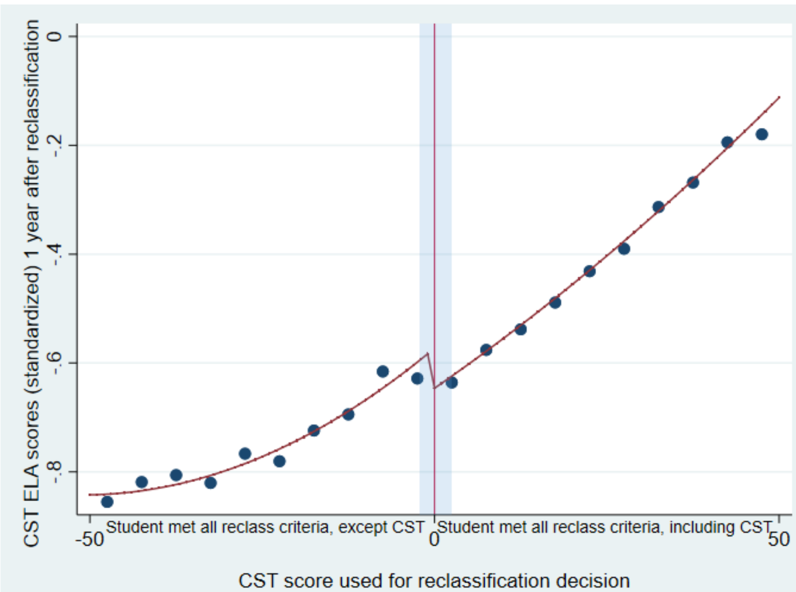
models differed in significance was that the simple linear model sometimes disagreed with the models with higher order polynomials. We interpret this as most likely being due to some non-linearity in the relation between the running variable and the given outcome, which is not allowed for in the linear model. In the main text we focus on the results that allow for a second order polynomial model. Visual inspection of results suggested this was the most sensible approach given that the relationship between the running variable and various dependent variables was not linear, and because the models with third and fourth order polynomials often showed signs of overfitting, with occasional but dramatic non-linearities in the fitted line that are classic signs of overfitting.

[21] It is important to note that in LAUSD, if a student exits the school district, they are absent from our sample, whereas in SDUSD, the district tracks students that are dropouts. Therefore, LAUSD's on-time graduation outcome in this study is calculated among students who persist to 12th grade. On-time graduation in SDUSD includes dropouts in the denominator.

After the CELDT was rescaled and the math course grade requirement was dropped in LAUSD in 2006–07, the impacts of reclassification were slightly different. Elementary school students reclassified in the 2007–12 school years had lower CST ELA and math scores a year after reclassification.[22] This is seen in Figure 6 by examining the shaded area. Figure 6 illustrates the slightly lower CST ELA scores one year after reclassification of students just above the CST cutpoint relative to students just below it. However, two years after reclassification, reclassified students were similar to those who just missed being reclassified two years earlier—there was no statistically significant difference in CST scores two years after reclassification for students just above and just below the cutpoint (see Table 5A in Technical Appendix B). So while there is a short-term negative consequence to being reclassified for elementary school students in the 2007–12 era, there were not longer-term consequences for these students.

FIGURE 6

Elementary ELs reclassified in 2007–12 have lower CST ELA scores one year after (LAUSD, grades 3–5, era 2)



SOURCE: Authors' estimates. See Table 5A in Technical Appendix B.

NOTE: CST is the running variable and has been rescaled to zero. CST one year after reclassification is the outcome variable. The lines in the graphs show the regression fit using a quadratic polynomial estimated independently on the two sides of the cutoff. Each dot represents a bin size of 5. The discontinuity was -0.0962 and was significant. Sample size was 48,361.

In SDUSD, in most cases those reclassified after meeting the CST cutoff performed the same as those just below it. Importantly, this suggests SDUSD set its reclassification criteria appropriately.

There were two exceptions. Figure 7 shows that in the original CELDT era in 2003–04 and 2004–05, SDUSD students meeting the CST cutoff in high school had lower CST ELA scores a year later.[23] Figure 8 shows that in the rescaled CELDT era from 2006–07 through 2013–14, SDUSD middle school students meeting the CST cutoff performed better on the CST ELA test two years after the reclassification decision than those just below the cutoff. While both discontinuities are statistically significant, the drop in test scores for those above the cutpoint in Figure 7 is more apparent than is the rise in test scores in Figure 8.

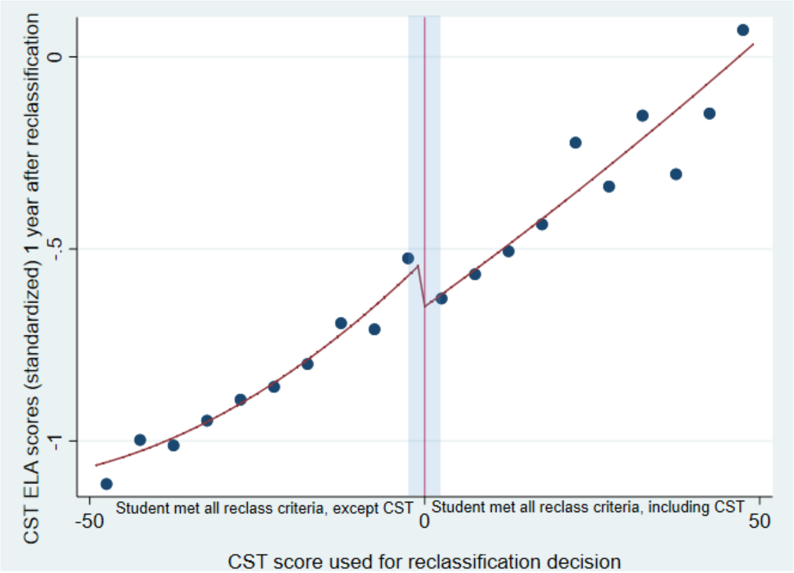
---

[22] Math scores were significant in all but the quadratic model.

[23] However, two years later, CST ELA scores were higher for high schoolers (in all but the quadratic model).

FIGURE 7

High school ELs reclassified in 2004–05 have lower CST ELA scores one year after (SDUSD, Grades 9–12, era 1)

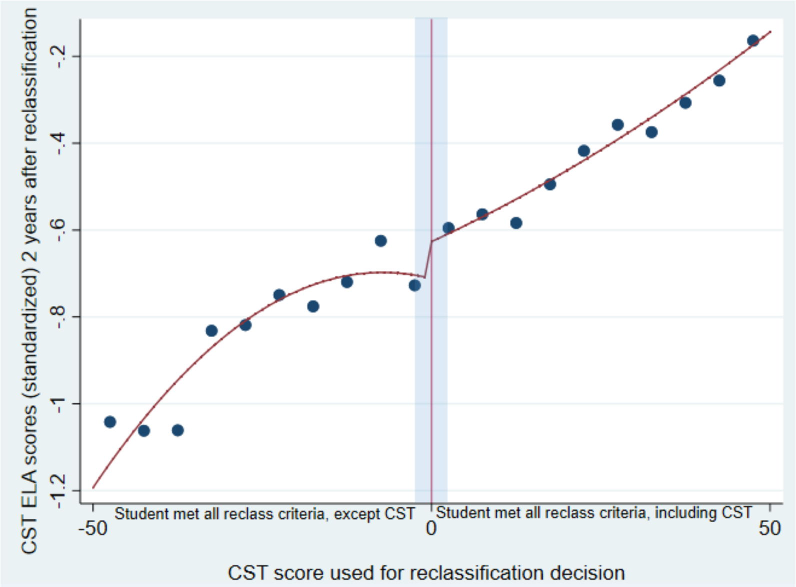


SOURCE: Authors' estimates. See Table 5B in the technical appendix.

NOTE: CST is the running variable and has been rescaled to zero. The lines in the graphs show the regression fit using a quadratic polynomial estimated independently on the two sides of the cutoff. Each dot represents a bin size of 5. The discontinuity was -0.138 in the model with other controls and was significant. Sample size was 1,513.

FIGURE 8

Middle school ELs reclassified in 2007–14 have higher CST ELA scores two years after (SDUSD, Grades 6–8, era 2)



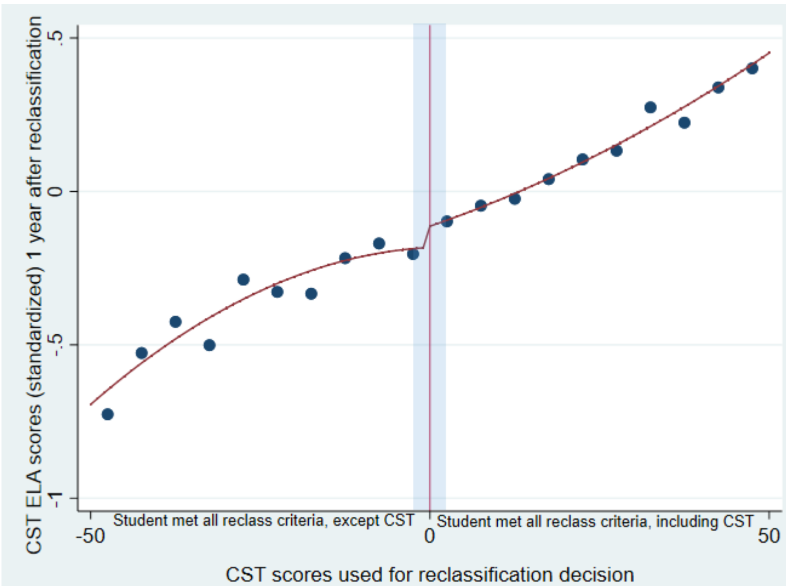
SOURCE: Authors' estimates. See Table 5B in the technical appendix.

NOTE: CST is the running variable and has been rescaled to zero. The lines in the graphs show the regression fit using a quadratic polynomial estimated independently on the two sides of the cutoff. Each dot represents a bin size of 5. The discontinuity was 0.114 in the model that included other controls and was significant. Sample size was 5,155.

In almost all other cases, SDUSD students at or above the cutoff performed about as well as those below the cutoff, which implies that EL supports were withdrawn at about the right time.[24] For example, Figure 9 shows CST ELA scores one year after reclassification for the newer CELDT era for grades three to five from 2007–14. At the reclassification cutoff, the change in the mean CST score in the following year is almost imperceptible.

FIGURE 9

Elementary ELs reclassified in 2007–14 have shown no change in CST ELA scores 1 year after (SDUSD, Grades 3–5, era 2)



SOURCE: Authors' estimates. See Table 5B in the technical appendix.

NOTE: CST is the running variable and has been rescaled to zero. The lines in the graphs show the regression fit using a quadratic polynomial estimated independently on the two sides of the cutoff. Each dot represents a bin size of 5. The discontinuity was -0.009 in a model that added background controls and was not significant. Sample size was 7,926. (The discontinuity is positive in the above graph at +0.067, and the difference in sign is because the graph shows the simple model that controls only for the running variable.)

We set out in this section to answer whether the two districts in the past have set an appropriate cutoff for reclassification on the CST test. LAUSD and SDUSD had the same CST cutpoint during each reclassification era for secondary schools. In contrast, the two districts had the same CST cutpoint in era 1 but SDUSD had a more rigorous cutpoint in era 2. However, the two districts had different CELDT subtest requirements and LAUSD required course grades. Despite these differences in reclassification requirements and the increase in the demands of the CELDT in 2006–07, it appears that each district had set an appropriate CST cutpoint for a smooth academic transition to reclassified status in most cases. Exceptions include on-time graduation in LAUSD in the old reclassification era, and short-term consequences (i.e., effects observed one but not observed two years after reclassification) in both districts. How could the reclassification criteria differ, but the outcomes be largely the

[24] Although SDUSD is the second largest district in California, it is much smaller than LAUSD. One might wonder whether the finding that there were no statistically significant changes in outcomes at the CST score cutoff merely reflects imprecise estimates. However, the power analysis we conducted before beginning the research, $\alpha = 0.05$, and power set to 0.80, suggested that the Minimum Detectable Effect Size for the 2004–2005 analysis of CST scores one year later was on the order of 0.07 to 0.1, depending on the grade spans studied. Those results indicated we could detect quite small impacts of reclassification. The actual standard errors we obtained were in line with these projections. For the second CELDT era, with four times as many years of data available, the standard errors are even smaller. The one place where statistical power becomes an issue in the SDUSD data is for the various potential experiments with CELDT test scores as the running variable, where we found only one potential experiment, based on CELDT reading scores, with enough observations just above and below the required scores to achieve a reasonable level of power. We discuss those results in a later section.

same? One possibility is that the cutpoints are appropriate given the conditions at the district that we have not considered, such as the academic environment or demographics of EL and reclassified students.

Have the districts set an appropriate cutoff on the CELDT?

In LAUSD, we were able to conduct similar experiments in which we considered the impact of the requirement that students score at an overall Early Advanced performance level on the CELDT.[25] We performed similar validity checks to our CST experiments and concluded that, in most cases, we could confidently measure an effect of being just above the cutpoint.[26] In these experiments, we considered students who met the CST requirement, earned a course mark of "C" or better in English and math in 2003–06, and met all CELDT subtest requirements by scoring at least Intermediate on reading, writing, listening, and speaking. Correlations among the CELDT subtests and the overall CELDT score are high, but 27 percent of students meeting all CELDT subtest criteria did not meet the overall CELDT cutpoint by having an overall score of Early Advanced. We then compared students just above and just below the CELDT Early Advanced cutpoint.[27] It appears that the district reclassified students according to policy at very high levels in the second reclassification era. Those just meeting the CELDT overall requirement had between 61 and 85 percentage point gains in the reclassification rate in the new era. Adherence to policy was weaker in the previous era, with reclassification rate jumps between 33 and 47 percentage points.

In the earlier era, there were no significant effects of being reclassified at any grade level on CST ELA scores one or two years after reclassification, on CST math scores one or two years after reclassification, or on on-time graduation (see Technical Appendix B, Table 6). This implies that the CELDT Early Advanced overall performance requirement was set approximately correctly for smooth transition from EL to reclassified status given the other reclassification requirements.

Once the CELDT became more challenging in 2006–07 and LAUSD dropped the math grade requirements, reclassified middle school students who scored just above the CELDT OPL cutpoint had slightly lower CST ELA scores than similar students who scored just below the cutpoint in the first, but not later years.

In SDUSD, there were too few students just above or below the CELDT OPL cutpoint for the experiment to be valid. With only a handful of students who met all other criteria but were just below the OPL cutpoint, there were not enough observations to draw meaningful conclusions.

Both districts had enough students above and below the CELDT reading subtest cutpoint to permit examination of the effect of reclassification in at least some grade levels.[28] For the CELDT reading subtest, as with the CELDT OPL, LAUSD reclassification policy seems to have been implemented more regularly in the new era. In this era, the probability of reclassification jumps at or above the CELDT reading cutpoint 76 to 86 percentage points compared with 31 to 46 in the old era.

While the CELDT OPL cutpoint is Early Advanced in LAUSD, scores on any subtests need only be Intermediate for an EL student to be eligible for reclassification. We had enough students on either side of the cutpoint to estimate the effect of being just above the cutpoint for elementary students in the first reclassification era and elementary, middle, and high school students in the second era. (See Technical Appendix B, Table 7A.) We found no impact on our academic outcomes of interest and conclude that, given the other reclassification requirements,

---

[25] Cutpoints for the Early Advanced level are pegged to scaled scores. Scaled scores required for Early Advanced vary by grade level.

[26] However, we do not have confidence in the results for middle school reclassification in the first reclassification era when it comes to our estimates of the impact on math scores. The size of the discontinuity at the CELTD OPL cutpoint for the math outcomes was just 0.147.

[27] Validity checks are available from the authors on request.

[28] Validity checks are available from the authors on request.

CELDT reading cutpoints for elementary school students were set at the proper level in both reclassification eras. For middle and high school students, CELDT reading cutpoints were also properly set in the new era.

In SDUSD, we had enough observations to evaluate the CELDT reading cutoff in the 2007–2014 period. Overall, we found no cases with a statistically significant jump or drop in outcomes for students at or above the CELDT reading cutpoint, suggesting that the Early Advanced criterion was appropriate (See Technical Appendix B, Table 7B).[29]

Overall, then, did the two districts set appropriate cutpoints on the CELDT? The answer appears to be yes. Despite SDUSD and LAUSD requiring different CELDT subtest cutpoints and the increase in rigor of the CELDT experienced by both districts, we did not see differences in the impact of being reclassified across the two districts or over time. EL students in both districts had relatively smooth academic transitions to reclassified status. We do note, however, relatively large increases in LAUSD's fidelity to the CELDT reclassification policy in the new era. Recall that in the earlier reclassification era, the jump in reclassification rate was between 33 to 47 percentage points for students who met all criteria including the CELDT as compared to those who met all criteria but the CELDT. In the newer era, the jump ranged between 61 to 85 percentage points. Further, here we do not control for differences in the academic programs or all demographic characteristics of ELs across the districts.

Does Reclassification Have Different Effects on Long-Term and Late-Arriving English Learners?

At the secondary school level, ELs are heterogeneous in the length of time they were in the district before reclassification. Long-Term English Learners (LTELs) are those who have already completed at least five years of schooling and have not been reclassified.[30] Late-Arriving English Learners are those arriving for the first time at grade 6 or higher while scoring at the bottom proficiency level on the CELDT (Beginning level) when they first take the test. Hill et al. (2019) studies these two populations in LAUSD and SDUSD, and finds that late-arriving ELs have much lower reading achievement than LTELs, but on average make faster gains. In both districts, LTELs greatly outnumber late-arriving ELs. A third, and very small, group consists of ELs who arrive before grade 6 but have not yet spent five years as ELs in a school district.

In LAUSD, we examined reclassification effects among LTELs in grades 6–8 and 9–12. On the CST, we found that a negative effect of reclassification on high school graduation was only weakly evident in one of the four models we estimated. Otherwise, LTELs just above the CST cutpoint had similar academic outcomes. Our estimates for LTEL academic outcomes after reclassification using the CELDT overall and reading scores also failed to find any statistically significant differences.

Because the numbers of late arrivers are relatively small, we combined grades 6–12 to estimate as many regression discontinuity experiments as possible for this group in LAUSD. In all instances for which we had

---

[29] As shown in Technical Appendix B, Table 7B, in no case did our favored specification with the outcome modeled as a quadratic function of the reading CELDT score become significant. For the outcome CST ELA scores a year later, in elementary school a negative significant coefficient emerged, but only in the third and fourth order polynomial models. Conversely, for graduation on time, in elementary schools a significant positive coefficient emerged, but only for the linear model.
[30] In SDUSD, this is defined as five years in the district and in LAUSD, it is defined as five years in the US.

enough students to perform the experiment, we found no statistically significant differences for students just above and below the cutpoint.[31]

In SDUSD, we lacked a sufficiently large sample of late-arriving ELs to repeat the analyses for this group. But we did have enough LTELs to repeat all of the experiments for this subgroup in grades 6–8 and 9–12. The results were extraordinarily similar to the main results described above. In no case was the impact of being above the cutoff or of being reclassified statistically significant. Furthermore, in the few cases where there was a significant impact for all ELs, the direction of the estimated effect was the same. LTELs constituted well over half the samples used in our earlier analyses, which may explain why findings were similar when we used all ELs in middle and high schools.

We conclude that separate reclassification policies for late-arriving and long-term ELs do not appear to be necessary, though we caution that, even in LAUSD, our sample of late-arriving ELs is small.

Do Moderating Factors Impact the Effect of Reclassification?

The appropriate timing of reclassification may vary according to student, school, or neighborhood characteristics. If so, the optimal reclassification strategy could vary depending on student subgroups or school contexts. We examine five types of potential factors that might affect our CST ELA test score experiment one year later:

1. language spoken at home;
2. student demographics, including concentrations of students by language and diversity of languages spoken at a school;
3. neighborhood characteristics, such as language, poverty, and nativity in the school catchment area, based on 2006–16 zip code data estimates from the American Community Survey;
4. average teacher qualifications at a school;[32]
5. fidelity to English Language Development course placement as outlined in each district's *Master Plan for English Learners*.

In both districts, we estimated separate regression models to determine if any of these moderating factors appears to have an impact on reclassification.

[31] In the CST experiment, we were able to estimate CST ELA scores one year post reclassification in the new reclassification era and high school graduation in both eras. For the CEDLT OPL experiment, we were able to estimate CST ELA scores and gradation in the second reclassification era. For the CELDT Reading experiment, we were only able to estimate high school graduation in the second reclassification era.

[32] We measure average teacher tenure and percentage of teachers with at least a master's degree at each school among all teachers, using publicly available data from the California Department of Education.

LAUSD Results

In LAUSD, all the following variables appeared to have a negative impact on CST ELA scores for students just above the CST reclassification cutpoint relative to those just below, when estimated in separate regression models.[33] (See Technical Appendix B, Table 8.)

- being a Spanish speaker;
- school language characteristics, including overall EL percentage, the percentage of ELs speaking the same language as the student, the percentage of the overall student body speaking the same language as the student, and a measure of language homogeneity among EL students;
- neighborhood language characteristics, including the percentages of foreign-born and Spanish-speaking people.

However, these findings must be considered inconclusive because the measures of language are closely correlated. We then estimated a model with just one variable from each of the five types of moderating factors listed above, but included these potentially moderating factors simultaneously. We found little evidence of a systematic effect of moderators on CST ELA test scores across reclassification eras or grade levels. (See Technical Appendix B, Table 9.)

For example, in the first reclassification era in LAUSD, we test the five moderating variables in three models (elementary, middle, and high school), and in only two of the 15 instances were the moderating variables statistically significant for students just above the reclassification threshold. We do not see enough evidence to draw conclusions about the effects of these variables in the first reclassification era. In the more recent reclassification era, more of these variables seem to affect reclassification outcomes for students reclassified in middle and high school. However, these results are either not robust to variations in student grade level or reclassification era, or they are largely offset by reclassification's positive overall effect. Overall, these results do not appear to provide a compelling reason to modify reclassification policy according to school context or linguistic subgroup.

SDUSD Results

In San Diego, moderating variables that measured students, schools, and neighborhood characteristics did not, for the most part, appear to have an independent effect on ELA standardized test scores one year after reclassification, an important outcome.

There were a few exceptions in the CST ELA experiment. For both elementary and middle schools in the new CELDT era, the effect of meeting the CST reclassification cutoff was smaller if teachers had more experience. One interpretation is that, in that era, it might have been worthwhile postponing reclassification for students in schools with highly experienced teachers. There were two other moderating variables that were significant, but this pattern appeared in only one grade span and one era, which raises concerns that the results could have been accidental.[34]

[33] Technical Appendix A explains how we test for evidence of impact of potentially moderating variables and reclassification outcomes.

[34] First, in the newer CELDT era for middle schools, the proportion of teachers holding a master's degree had negative and significant interactions with the indicator for meeting the CST criterion. Second, in the older CELDT era there was a negative and significant interaction only between the indicator for whether the student's home language was Spanish and meeting the cutoff in high school. This was also found in the LAUSD results, but in that case, the negative interaction was in the new CELDT era.

In the new CELDT era, the impact of meeting the CST cutoff was smaller for high school students as the share of foreign-born rose in the school zip code. Notably, this pattern was not statistically significant when we used the share of foreign-born in the zip codes where students lived.

We tentatively conclude there is little evidence that student, school, or neighborhood characteristics had a separate effect on academic outcomes as a result of reclassification. This probably is good news for educators and policymakers who must standardize EL reclassification criteria across the state.[35] However, there may be other characteristics of the EL instructional environment or the reclassified student environment that we do not measure that could be relevant and could also contribute to an explanation about why reclassification criteria can differ across LAUSD and SDUSD but still be appropriate for students in their districts.

Conclusion and Recommendations

Overall, in the two school districts and two reclassification eras we studied, reclassification criteria appear to largely have been appropriate. Below, we discuss why this might be the case despite the fact that reclassification policies differed over time and across the districts. However, we find that some LAUSD elementary school students may have been reclassified too soon, resulting in short-term negative outcomes. Perhaps the English Language development instruction given to students just below the reclassification cutpoint provided a special protective effect that ELs barely above the cutpoint did not receive. Within two years, the negative effects had mostly dissipated under the newer CELDT. Thus, it is possible that the negative outcomes could be eliminated or mitigated by augmented monitoring or transition services of limited duration.[36] At the high school level in LAUSD, we found negative impacts on on-time graduation for reclassified ELs, but only in the first reclassification era.

In SDUSD, students just above and below the CST and the CELDT reading cutpoints performed about equally on the full set of outcomes we considered, suggesting that, in both CELDT eras, SDUSD set reclassification criteria appropriately.

We cannot evaluate what would have happened had either district used significantly more or less stringent reclassification criteria. Our analysis does not tell us whether those with far lower or higher achievement would have been affected similarly by reclassification. In addition, although we can estimate the causal effect of reclassification, in most cases we cannot compare one reclassification era with another in a district because multiple policies may have changed between the first and second eras. SDUSD middle schools and high schools are an exception because in those cases the only change between eras was the introduction of the more difficult CELDT.

Our research raises several questions:

Why was the impact not greater when reclassification policy changed in the two districts?

It is possible that these two districts modified their English Language Development programs in reaction to the new reclassification policy, leaving reclassification outcomes largely unchanged. We do not assess the content of instruction or any potential adaptations.

LAUSD had much more compliance with its reclassification policies in the new era, especially in CELDT. In the old era, staff and teachers may only have reclassified students whom they thought could manage

[35] To put into context these four exceptions to the rule that no variables moderate the impact of being reclassified in the CST experiment, we performed 42 different tests. So even if none of the moderators influenced the impact of reclassification, we would expect two tests to find significant effects. We are close to this rate of significant findings.

[36] ESSA already requires that reclassified students be monitored for four years.

without English Language support. Alternatively, they may not have believed criteria were rigorous enough and were applying stricter standards on their own.

In SDUSD, the move to a more-rigorous CELDT may have had little impact because the district already had stricter requirements on the test than LAUSD. (Only one subtest score of Intermediate is allowed in San Diego). Thus, moving to more difficult CELDT requirements may not have changed the profile of students who were reclassified in San Diego.

Another explanation is simply that there is a range of cutpoints that could be appropriate, rather than a single cutpoint. This idea emerges in the San Diego results, where the reclassification criteria clearly became more stringent in the new CELDT era, but in both eras, ELs appear to have been reclassified appropriately.

How will standardization affect reclassification policy?

Standardization of reclassification policies is required by the federal Every Student Succeeds Act (ESSA). 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. Once reclassification policy is standardized for all four criteria across school districts, we may learn more about the effects of different district instructional models.

Nevertheless, standardization will probably not align school district reclassification rates, even if instructional models are similar. Other factors, such as the percentage of recently arrived immigrant students, may play a role. Furthermore, standardization will affect districts differently. For example, if the new state standard is more rigorous than a district's old policy, reclassification rates will fall.

The California Department of Education is working with WestEd to develop an observational protocol to standardize teacher evaluations of student English language proficiency. Whether the protocol is implemented in a standard way will depend on funding and the effectiveness of accompanying professional development.

How can reclassification standards be set to ensure all students get core academic instruction?

Slightly more than half of the state's K–12 English-only speaking students, but only 12 percent of ELs, meet or exceed ELA standards on the new statewide SBAC test. There is a risk that access to core academic instruction will be restricted if new state policy requires rigorous SBAC cutpoints. November 2018 results indicate that EL students scoring high enough for reclassification on ELPAC on average have scores below those of their English-only speaking peers in grades 5–8 and grade 11 on the SBAC ELA assessment (Linquanti et al. 2018). Further research could provide evidence on the ability of students meeting the ELPAC reclassification standard, but not scoring at the "met standard" level on the SBAC, to make the transition to reclassified status.

State policy has also changed to ensure K–12 English Learners have access to college preparatory curriculum. Two new bills aim to help English Learner students gain more-robust access to the core academic curriculum in the K–12 setting. The first, AB 2735, requires school districts to allow ELs full access to a school's standard instructional programs, including a–g courses and AP classes.[37] Another bill, AB 2121, mandates a fifth year of high school for newcomers and students in the Migrant Education Program. California's higher education institutions may provide a model by reducing student time in developmental courses, while allowing students into college-level courses more quickly.

Recommendations

As we move into a new era of standardization, district and state policymakers have an opportunity to ensure that EL students are well served by school ELD programs and reclassification policies. To be sure, reclassification rates will not be uniform across districts—they will vary depending on the English proficiency of each district's EL population and the availability and efficacy of EL instruction.

[37] It does grant districts an exception for students who have just arrived in the US.

The reclassification criteria used in the past by LAUSD, and San Diego in particular, seem to have been set appropriately. Our research shows that reclassified students performed neither substantially better nor worse than ELs whose language scores were just below the cutpoint. This provides important hints about what appropriate reclassification criteria might look like using the new assessments: the SBAC ELA test and ELPAC, the new state test of EL English proficiency.

Currently, EL students must meet multiple criteria to be reclassified. The state should consider the advisability of using fewer criteria. Some research suggests it may be beneficial to use just one criterion to assess whether EL students are ready for reclassification (Umansky et al. 2015 and Hill et al. 2014). If just one criterion is used, ELPAC may be ideal because it is aligned to the state's new ELA standards. Another option is to allow any of a few criteria to be sufficient for reclassification. This is the direction California's community colleges and the California State University system are moving in assessing student readiness for college-level courses.

Policymakers at each level should treat ESSA's requirement to monitor students for four years after reclassification seriously to ensure that any new reclassification policies are effective. In particular, state policymakers should monitor EL students to make sure the new English language proficiency assessment and any new criteria based on SBAC are not so rigorous that they prevent students who could succeed from being reclassified. And the rigor of new reclassification criteria must not make it harder for students to access core content.

As California overhauls its reclassification criteria, it is imperative that we understand the effects of the policies of the past in order to create a roadmap for the future.

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.

Betts, J., A. Zau, and L. Rice. 2003. *Determinants of Student Achievement: New Evidence from San Diego*. Public Policy Institute of California.

Callahan, Rebecca. 2005. "Tracking and High School English Learners: Limiting Opportunity to Learn." *American Educational Research Journal* 42 (2): 305-328.

Callahan, R., L. Wilkinson, C. Muller, and M. Frisco. 2009. "ESL Placement and Schools: Effects on Immigrant Achievement." *Educational Policy* 23 (2): 355–384.

California Department of Education. 2007. "Technical Report for the California English Language Development Test (CELDT): 2006-07 Edition (Form F)."

California Department of Education. 2016–17. DataQuest.

California Department of Education. 2019. "Updated Reclassification Guidance for 2018-2019."

Conger, D., A.E. Schwartz, and L. Steifel. 2011. "The Effects of Immigrant Communities on Foreign-Born Student Achievement." *International Migration Review* 45 (3): 675–701.

EdSource. 2008. "English Learners in California: What the Numbers Say."

Estrada, Peggy. 2014. "English Learner *Curricular Streams* in Four Middle Schools: Triage in the Trenches." *The Urban Review* 46 (4): 535–573.

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

Friesen, J., and B. Krauth. 2011. "Ethnic Enclaves in the Classroom." *Labour Economics* 18 (5): 656–663.

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

Hanushek, E.A., J.F. Kain, and S.G. Rivkin. 2002. "New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement." NBER Working Paper 8741, National Bureau of Economic Research.

Hill, L. 2018. *K-12 Reforms and California's English Learner Achievement Gap*. Public Policy Institute of California.

Hill, L., J. Betts, M. Hopkins, M. Lavadenz, K. Bachofer, J. Hayes, A. Lee, M. Murillo, T. Vahdani, and A. Zau. 2019. *Academic 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*. Public Policy Institute of California.

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

Hoxby, C. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." NBER Working Paper 7867, National Bureau of Economic Research.

Imbens, G.W., and T. Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142 (2): 615–635.

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, Robert, Min Huang, and Eric Crane. 2018. Attachment 3. Report on Supplemental Empirical Analyses of the English Language Proficiency Assessment for California. State Board of Education.

Los Angeles Unified School District. 2012. *English Learner Master Plan*.

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. 2014. “Implementing Common Core State Standards in California: A Report from the Field.” Stanford University and PACE.

Parrish, T., A. Merickel, M. Perez, R. Linquanti, M. Socias, 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*. American Institutes for Research and WestEd.

Pope, N.G. 2016. "The Marginal Effect of K-12 English Language Development Programs: Evidence from 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.” *Educational 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.

Salazar, J.J., and K. Hayes. 2010. “The Role of Home Language Proficiency and Its Impact on English Learners' Academic Achievement.” Paper presented at the annual meeting of the American Educational Research Association.

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.

Sugarman, Julie. 2019. *The Unintended Consequences for English Learners of Using the Four-Year Graduation Rate for School Accountability*. Migration Policy Institute.

Umansky, I., and S. Reardon. 2014. “Reclassification Patterns among Latino English Learner Students in Bilingual, Dual Immersion, and English Immersion Classrooms.” *American Educational Research Journal* 51 (5): 879–912.

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

US Department of Education. 2014. *What Works Clearinghouse: Procedures and Standards Handbook Version 3.0.* What Works Clearinghouse, Institute of Education Sciences.

Warren, P. and P. Murphy. 2014. *Implementing the Common Core State Standards in California*. Public Policy Institute of California.

ABOUT THE AUTHORS

Julian Betts is a professor and former chair of the Department of Economics at the University of California, San Diego, and is executive director of the department's San Diego Education Research Alliance (SanDERA). He is also an adjunct fellow at PPIC and research associate at the National Bureau of Economic Research. He has performed student-level research using San Diego Unified School District data over the past 19 years. His research on determinants of student achievement and evaluation of interventions has examined differences in outcomes between EL and non-EL students. He has extensive experience with experimental and quasi-experimental methods. He holds a PhD in economics from Queen's University, Kingston, Ontario, Canada.

Laura Hill is a policy director and senior fellow at the Public Policy Institute of California. Her areas of expertise are K–12 education and immigration. She is currently researching English Learners in California schools, special education, and ESL students in community colleges. Her recent publications examine student achievement on the state's K–12 assessments and the link between language reclassification policies and student success. Prior to joining PPIC, she was a research associate at the SPHERE Institute and a National Institute of Aging postdoctoral fellow. She holds a PhD in demography and an MA in economics from the University of California, Berkeley.

Karen Bachofer is director of SanDERA. Previously she was executive director of San Diego Unified School District's Research and Evaluation Division, where her responsibilities included direct oversight of national, state, and district assessment, and accountability processes and reporting. She also oversaw internal, external, and commissioned evaluation activities, including Title III reporting and production of the district's annual English Learner Progress Report. Her SanDERA research has included studies of the impact of EL reclassification policies and student learning trajectories. She holds a PhD in education from the Claremont Graduate School and San Diego State University.

Joseph Hayes is a research associate at the Public Policy Institute of California, where he studies population change, educational policy, and corrections issues. Recent projects have focused on estimates of the undocumented immigration population, English Learner reclassification policies in California public schools, and the changing composition of the state's prison and parole populations. He holds an MS in agricultural economics from the University of Wisconsin, Madison.

Andrew Lee is a research associate at the Public Policy Institute of California, focusing on K–12 education, specifically as it relates to English Learners in California schools. Prior to joining PPIC, he contributed to research on health policy and agricultural development policy at Acumen LLC and the UN Food and Agriculture Organization, respectively. He holds an MA in development practice from the University of California, Berkeley, and a BA in economics and English from the University of Michigan, Ann Arbor.

Andrew C. Zau is senior statistician at SanDERA. His research focuses on the effects of policy on student achievement in K–12 education. He has 19 years of experience analyzing SDUSD student-level longitudinal data and has played key roles in evaluations of the impact of reading reforms, diagnostic math testing, and school choice in San Diego, using experimental and quasi-experimental approaches. He holds an MPH in epidemiology from San Diego State University and a BS in bioengineering from the University of California, San Diego.

ACKNOWLEDGMENTS

Niu Gao, Olga Rodriguez, Veronica Aguila, Richard Murnane, and Lynette Ubois provided very helpful reviews, and Sam Zuckerman and Rebecca Morgan provided much-appreciated editorial support. We thank the many staff in central offices and school sites in both districts for sharing their insights. In particular, we thank Sandra Cephas at SDUSD and Kathy Hayes, Lydia Acosta-Stephens, and Hilda Maldonado at LAUSD for many helpful conversations. The research reported here was supported by the Institute of Education Sciences, US Department of Education, through R305A170288 to PPIC. The opinions expressed are those of the authors and do not represent views of the Institute or the US Department of Education. IES provided the main financial support for this report; funding from the W.T. Grant Foundation supported the regression discontinuity analyses of long-term and late-arriving ELs.

PPIC | PUBLIC POLICY INSTITUTE OF CALIFORNIA

25 YEARS

The Public Policy Institute of California is dedicated to informing and improving public policy in California through independent, objective, nonpartisan research.

Public Policy Institute of California
500 Washington Street, Suite 600
San Francisco, CA 94111
T: 415.291.4400
F: 415.291.4401
PPIC.ORG

PPIC Sacramento Center
Senator Office Building
1121 L Street, Suite 801
Sacramento, CA 95814
T: 916.440.1120
F: 916.440.1121

PPIC PUBLIC POLICY INSTITUTE OF CALIFORNIA

25 YEARS

English Learner Trajectories and Reclassification

Technical Appendices

CONTENTS

Julian Betts, Laura Hill, Karen Bachofer, Joseph Hayes, Andrew Lee, and Andrew Zau

The research reported here was supported by the Institute of Education Sciences, US Department of Education, through R305A170288 to PPIC. The opinions expressed are those of the authors and do not represent views of the Institute or the US Department of Education.

Appendix A. Methodology

We use a regression discontinuity design (RD) that exploits the rules both districts have established for reclassification of English Learners (ELs), as presented in Table 1. The assignment variables related to reclassification are California Standards Test (CST) scores and the overall and subtest scores on the California English Language Development Test (CELDT) and (in LAUSD) course marks. Subject to a number of conditions, an RD study will produce a consistent estimate of the causal effect.

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 for reclassification remain in EL support (described in the main text), and serve as 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 the act of being reclassified and having English language development classes removed.

Conceptually, there are (at least) two ways to handle RD models with more than one forcing variable. 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.

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 Y_{ist} denote the test score or other outcome of student i in school s. Year t is the year in which we observe a reclassification decision, but the outcome observed will be in a future school year. Let $Z_{i,t-1}$ be a vector of baseline characteristics representing a vector of background variables measured the year before the reclassification decision. Note that the vector Z is measured in period t-1 because, in the year of reclassification, reclassified students receive treatment for part of the school year before being reclassified. These will include time invariant demographic variables such as indicators for race/ethnicity, language spoken at home, and gender. But $Z_{i,t-1}$ will also include 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 (and on course grades in LAUSD). The key regressor is a 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 local 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:

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

Here, the key coefficient is δ. If it 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.) In later experiments, we use the CELDT overall performance level and the CELDT reading subtest as the running variable.

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 weighted 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.[1]

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

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

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 an outcome such as ELA achievement the year after the reclassification decision, but replace actual reclassification with predicted reclassification $\hat{R}_{it}$. We estimate local 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:

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

Here, the key coefficient is π. If it is not significantly different from zero, then we retain the null hypothesis of a zero causal impact of reclassification on test scores. We cluster the standard errors at the student level.

Establishing the Validity of the Regression Discontinuity Approach

We showed in Figure 4 that in the case of the CST requirement, indeed, there is reason to expect the reclassification criteria are implemented by school districts in such a way as to simulate an experiment with a treatment and control group. We include all students who have met all reclassification criteria except (perhaps) the CST. For example, in the case of LAUSD in 2003-05, we include ELs who met the CELDT OPL cutoff, scored at least intermediate on all CELDT subtests, and earned at least course marks of "C" in ELA and math. Then we plot the probability of being reclassified by CST score. (The CST has been rescaled to zero to signify the cutpoint score of 300); Appendix B, Table 1A and 1B demonstrate the difference in reclassification probabilities for the CST requirement across our grade levels and during the two reclassification eras. They range from 0.49 to 0.92 in LAUSD and 0.39 to 0.78 in SDUSD.

In LAUSD, after 2006-07, there was much more fidelity to the district's reclassification policy than in the earlier years of that decade. Our earlier research examining how EL students were assigned to English Language Development courses found that LAUSD has been moving in the direction of greater automation which coincided with increases in correct course placement. We know similar automation was undertaken for reclassification decisions, which may explain why the later years see sharper discontinuities (e.g., the probability of reclassification increased to 0.8 or higher for those just above the cutpoint in era 2).

[1] 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 Control Trials.

We performed similar analyses for the various cutoffs on the CELDT test (on the Overall Proficiency Level and the subtests, the most daunting of which appears to be the reading subtest). Our tests to ensure that the CELDT overall proficiency level and CELDT reading subtests also demonstrated discontinuities and are available from the authors upon request. However, we found adequate sample sizes to conduct these CELDT experiments for only some grades and years.

Tests for Manipulation of the Running Variable(s)

To have institutional integrity, the running variables such as the CST ELA score cannot 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 ELA and math tests are statewide assessments and are scored outside of the given school district, which makes it unlikely that the scores could be manipulated locally. The same is true of the CELDT.[2]

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. We expect to see no discontinuity in the frequency of students above versus below the cutpoint. Appendix B, Tables 2A and 2B show results for this test for the sample used to model the impact of reclassification on CST ELA scores the year after reclassification. They also show the same for the sample used to model the impact of graduation on time.[3] The coefficient on ABOVE is not significantly different from zero in either subsample, or for any of the three grade spans. Appendix B, Figures 1A and 1B give one example from each district, showing the distribution of students versus the running variable for elementary grades in the more recent ("new") eras in either district.

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. It is not appropriate to use the year of the reclassification decision as the source for baseline data because reclassification can occur part way through the year, and thus all outcomes in that year are potentially endogenous should that student be transitioned out of English language development courses during that year.

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. Appendix B, Tables 3A and 3B show the coefficient on ABOVE in models where the dependent variable is one of a number of student background characteristics. For SDUSD, the discontinuity variable is almost always statistically insignificant, with three exceptions (out of 53 tests). All three were from the second CELDT era. For elementary school students, those at or above the CST cutpoint were significantly more likely to be observed in an earlier year and an earlier grade, but the differences were small (0.3 year, 0.1 grade level difference). For high school students, those at or above the CST cutpoint exhibited a drop in CELDT overall proficiency level of 7 points. (All students in the CST experiment had to meet the CELDT Early Advanced score corresponding to Early Advanced. In grades 9-12, this minimum score ranged from 579 to 591, so the discontinuity represents roughly a 1 percent drop from the minimum needed to meet the CELDT requirements.)

---

[2] The annual CELDT, which is what we use, is scored off site. A student's answers to the speaking subtest are recorded by the examiner, but are scored off-site by the test vendor. (CDE 2007)

3 This second set of checks using graduation on time is really a test not just for manipulation of the CST test scores, but of differential attrition in the 1 to nine years between the reclassification decision and when students are expected to be in grade 12.

For LAUSD, slightly more of these tests produce significant discontinuities than one would expect by chance. In LAUSD, there are no discontinuities observed in the numbers of students who are female or Spanish speakers just above and below the CST cutoff. However, there are a number of negative discontinuities observed just above and below the CST cutoff for the various CELDT subtests (especially the reading subtests), as well as with the overall score on the CELDT. Academic year and grade level are often statistically significant as well, as they sometimes are in SDUSD.

Overall, it appears that both districts have adhered to their stated reclassification policies quite closely, and that the fuzzy RD design is appropriate. However, in our analyses of impacts on outcomes, we control for the baseline characteristics to increase precision, and this seems warranted given the occasional difference in these variables on either side of the CST cutoff.

Another potential issue is whether the bandwidth influences the findings. Wider bandwidths will generate greater precision but potentially at the cost of greater bias. In results available from the authors we reproduced the results for the CST experiment that follows (which used a bandwidth of +-50 scale score points), using bandwidths of +-100 and +-30 points. Results are broadly similar, although in a few cases with the smallest bandwidth results become statistically insignificant, even though the sign of the estimated impact does not change. We believe that this merely represents a loss of precision as we trim the sample. We performed similar bandwidth tests for the experiments based on the CELDT and again results were similar.

As mentioned in the main text, the LAUSD data do not distinguish between dropouts and students who leave for another school district. In the models of graduation on time, only students who appear in at least one 12th grade year are included in our analysis. In the SDUSD data, by contrast, we can distinguish between those who leave the district and those who drop out. We repeated all the SDUSD analyses of graduation on time, where in the main results we drop those who leave the district, and instead treated these students as not having graduated on time.

The results do not change markedly. Table 15 in Appendix B shows the original regression discontinuity results for era 2, for the CST experiment, separately for elementary, middle and high school. Adding those who left the district to enroll elsewhere into the data, and treating them as not having graduated on time, increases the sample size in the models meaningfully. Nonetheless the results are very similar. The biggest difference was for middle schools where the ITT estimate dropped by 0.036 and the TOT effect dropped by 0.053. In no case was there a change in statistical significance. For elementary and high school the results were the same to within 0.003 in all cases. The similarity in results may be because the key coefficient in RD analysis is the indicator for being at or above the CST cutpoint. There is no reason to think that those moving away were more likely to be just below versus just above the CST cutpoint, and so the estimate of the discontinuity should not be much affected.

Moderator Analysis

Perhaps even more policy-relevant than our main question of “Are ELs being reclassified at the right time?” is the question “Does the appropriate time of reclassification vary by characteristics of the student, school or neighborhood?” We test whether theoretically possible moderating factors are related in practice to heterogeneity in the effects of being reclassified. (Although studies have examined heterogeneity in the academic progress of ELs, we are asking a distinct and more novel question: Do appropriate reclassification criteria vary by subgroup?) Our research tests for heterogeneity along five dimensions:

- language spoken at home

- demographics of the student body, including concentrations of students by language, diversity of languages spoken at the school
- neighborhood characteristics (e.g. language, poverty, and nativity in the school catchment area) based on 2006-2016 zip code data estimates from the American Community Survey
- average teacher qualifications at a given school
- fidelity to the English Language Development courses specified in each district's Master Plan for English Learners.

We model home language with a single dummy variable for whether the home language is Spanish, "SPANISH," given the aforementioned strong majorities of ELs for whom Spanish is the home language. We modify equation (1) by adding a control for SPANISH, as well as an interaction of SPANISH with the indicator for being above the cutpoint:

$$Y_{is,t+1} = \alpha + \theta SPANISH_i + \beta CST_{i,t-1} + \delta ABOVE_{it} + \gamma ABOVE_{it} \cdot CST_{i,t-1} + \pi SPANISH_i \cdot ABOVE_{it} + Z_{i,t-1}{}'\Delta + \mu_{it} \quad (3)$$

The key coefficient here is π. If we retain the null that this coefficient is zero, then the causal impact of being above the cutpoint is the same between ELs who speak Spanish at home and ELs with other home languages.

The remaining moderators are tested using models similar to (3). We test for differential effects of reclassification related to concentrations of one's own home language among students at the school. For example, for ELs who speak Spanish, the appropriate measure will be the percentage of students at the school with Spanish as a home language. We consider it a test of enclave effects: If a large fraction of *all* students speak a given language, it may reduce students' incentive to master English.

Instead of all students, we can measure the percentage of students at the school who are currently ELs and have the same home language. This is a test of the idea that if one's own language group comprises a major share of the ELs in the school, teachers may focus their energies on helping those in one's language group.

Diversity of languages spoken captures a separate concept. If a large number of languages are spoken among ELs, it suggests that teachers may have relatively more difficulty teaching the EL students. We will use the Herfindahl concentration index to measure this language diversity. It is defined in terms of the squared shares of each language group j among ELs in a school:

$$H = \sum_{j=1}^{n} Share_i^2 \quad (4)$$

The value of our measure, known as a Herfindahl index, can be interpreted as the probability that any two ELs picked randomly at a given school would speak the same home language.

We measure average teacher qualifications at the school by determining the proportion of all teachers holding a Master's degree or higher, and by mean years of teaching experience. The skill level of teachers could influence the causal impact of reclassification on the students' performance. For instance, if more highly educated and experienced teachers are better at differentiating instruction, it could ease EL students' transition into non-EL classrooms. Of course, these skill measures could also relate to the quality of instruction for ELs as well as reclassified students.

We run separate models for elementary, middle and high school. Finally, in painstaking detail, we have determined if English Learner students are assigned to appropriate English Language Development (ELD) coursework, determining if they are placed correctly, in classes that are more rigorous than recommended by district policy, less rigorous than recommended, or in no ELD course. We focus on the share of ELs at a school who are not in any ELD course given evidence from these two districts that ELs tend to progress more slowly in schools where high shares of ELs do not take ELD courses (Hill, Betts, et al. 2019).

Appendix B: Supporting Tables and Figures

As noted in the main text, we adopted the following labeling of years in tables and figures and repeat that approach here: “2004-2005” means that the analysis was performed for the school years 2003-04 through 2004-05.

Tables 1A and 1B below establish that there is a significant jump in the share of students reclassified at the CST cutoff. (For all regression tables in the appendix, coefficients shaded green are positive and significant while those shaded red are negative and significant.)

TABLE 1A

LAUSD 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		
		g3-5	g6-8	g9-12	g3-5	g6-8	g9-12
CST ELA 1 Year Later	Coef.	**0.492****	**0.660****	**0.551****	**0.769****	**0.905****	**0.825****
	S.E.	(0.009)	(0.008)	(0.014)	(0.011)	(0.005)	(0.009)
CST ELA 2 Years Later	Coef.	**0.513****	**0.663****	**0.522****	**0.770****	**0.901****	**0.804****
	S.E.	(0.010)	(0.008)	(0.020)	(0.013)	(0.006)	(0.014)
CST Math 1 Year Later	Coef.	**0.493****	**0.612****		**0.768****	**0.859****	
	S.E.	(0.009)	(0.014)		(0.011)	(0.011)	
CST Math 2 Years Later	Coef.	**0.512****			**0.770****		
	S.E.	(0.010)			(0.013)		
Graduation on Time	Coef.	**0.516****	**0.668****	**0.582****	**0.891****	**0.921****	**0.869****
	S.E.	(0.012)	(0.010)	(0.012)	(0.010)	(0.006)	(0.007)

SOURCES: Authors’ calculations.

NOTES: “CELDT Era 1” (2003-2006) is before the CELDT was rescaled and “CELDT Era 2” (2007-2012) is after the rescaling. Tables 1A and 1B show coefficients and standard errors for the “ABOVE” dummy variable in a model of whether students are reclassified. The running variable was fit with a second-order polynomial. * and ** indicate significance at the 5% and 1% levels respectively. In LAUSD, on-time graduation is calculated for students who persist to 12th grade and does not include dropouts. For all regression tables in the appendix, coefficients shaded green are positive and significant while those shaded red are negative and significant.

TABLE 1B

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		
		g3-5	g6-8	g9-12	g3-5	g6-8	g9-12
CST ELA 1 Year Later	Coef.	**0.591****	**0.653****	**0.778****	**0.484****	**0.719****	**0.700****
	S.E.	(0.032)	(0.026)	(0.039)	(0.028)	(0.021)	(0.038)
CST ELA 2 Years Later	Coef.	**0.576****	**0.651****	**0.746****	**0.539****	**0.701****	**0.730****
	S.E.	(0.034)	(0.028)	(0.053)	(0.031)	(0.024)	(0.061)
CST Math 1 Year Later	Coef.	**0.592****	**0.668****		**0.486****	**0.638****	
	S.E.	(0.032)	(0.053)		(0.028)	(0.040)	
CST Math 2 Years Later	Coef.	**0.578****			**0.548****		
	S.E.	(0.034)			(0.031)		
Graduation on Time	Coef.	**0.592****	**0.645****	**0.763****	**0.394****	**0.729****	**0.494****
	S.E.	(0.036)	(0.028)	(0.039)	(0.061)	(0.026)	(0.031)

SOURCES: Authors' calculations.

NOTE: "CELDT Era 1" sample uses data from 2004-2005, meaning school years 2003-04 and 2004-05; "CELDT Era 2" uses data from 2007-2014 Tables 1A and 1B show coefficients and standard errors for the "ABOVE" dummy variable in a model of whether students are reclassified. The running variable was fit with a second-order polynomial. * and ** indicate significance at the 5% and 1% levels respectively. Coefficients shaded green are positive and significant while those shaded red are negative and significant.

Recall that to be included in this experiment, a student must have met all but the CST requirement. For LAUSD, we find between 20 and 30 percent of all EL students are included in each CST experiment. As another point of context, for students who are not reclassified, we examine the percentage of students who meet none of the criteria for reclassification. In LAUSD, that ranges from 16 percent to 30 percent, depending on the reclassification era and the grade span.

Tables 2A and 2B show the coefficient indicating whether there is a discontinuity at the cutoff in the number of students. There are no discontinuities in the SDUSD data, as desired. For LAUSD, only for graduation on time in the old reclassification era and for middle school students is there a significant discontinuity (positive). LAUSD data cannot distinguish between students who drop out versus students who transfer to other districts. Therefore, LAUSD's on time graduation calculation is based on students who persist to 12th grade, and students who drop out before 12th grade are excluded from our sample.

TABLE 2A

LAUSD McCrary test for discontinuity in number of students in the sample just above CST reclassification cutpoint

Outcome		CELDT Era 1			CELDT Era 2		
		g3-5	g6-8	g9-12	g3-5	g6-8	g9-12
CST ELA 1 Year Later (low attrition)	Coef.	87.70	769.7	196.1	215.5	-350.3	259.0
	S.E.	(904.8)	(394.1)	(215.6)	(861.4)	(307.5)	(275.1)
Graduation on Time (high attrition)	Coef.	27.21	560.2	208.2	36.77	-235.9	193.6
	S.E.	(555.3)	(264.8)	(232.5)	(267.2)	(218.4)	(359.8)

SOURCES: Authors' calculations.

NOTES: * and ** indicate significance at the 5% and 1% levels respectively. See Tables 1A and 1B for the years used to model outcomes during CELDT eras 1 and 2 in the two districts. Coefficients shaded green are positive and significant while those shaded red are negative and significant.

TABLE 2B

SDUSD McCrary test for discontinuity in number of students in the sample just above CST reclassification cutpoint

Outcome		CELDT Era 1			CELDT Era 2		
		g3-5	g6-8	g9-12	g3-5	g6-8	g9-12
CST ELA 1 Year Later (low attrition)	Coef.	31.67	70.14	2.715	-91.85	-145.4	7.058
	S.E.	(50.84)	(51.22)	(24.95)	(114.6)	(92.61)	(33.29)
Graduation on Time (high attrition)	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)

SOURCES: Authors' calculations.

NOTES: * and ** indicate significance at the 5% and 1% levels respectively. See Tables 1A and 1B for the years used to model outcomes during CELDT eras 1 and 2 in the two districts. Coefficients shaded green are positive and significant while those shaded red are negative and significant.

Tables 3A and 3B test for discontinuities in background variables of students. In most cases no significant discontinuity arises, but there are some exceptions. For this reason, all outcomes regression models discussed in the main text and shown in the tables below condition on these baseline variables. The figures however do not control for background variables.

TABLE 3A

LAUSD Checks for discontinuities in background variables above the CST reclassification cutpoint

Baseline Variables		CELDT Era 1			CELDT Era 2		
		g3-5	g6-8	g9-12	g3-5	g6-8	g9-12
Spanish Home-Language	Coef.	-0.002	-0.001	-0.007	0.007	0.001	0.013
	S.E.	(0.007)	(0.007)	(0.012)	(0.006)	(0.006)	(0.009)
CELDT Listening Subscore	Coef.	1.841	0.779	2.684	-1.341	**-9.243****	**-4.878***
	S.E.	(1.536)	(1.390)	(1.749)	(1.537)	(1.578)	(2.161)
CELDT Reading Subscore	Coef.	0.877	**-2.943****	0.655	**-2.087***	**-5.766****	**-3.121***
	S.E.	(0.856)	(0.658)	(1.007)	(0.917)	(0.817)	(1.259)
CELDT Writing Subscore	Coef.	-0.430	**-2.589****	0.517	-1.079	**-4.405****	-2.130
	S.E.	(0.915)	(0.778)	(1.172)	(0.951)	(0.809)	(1.387)
CELDT Speaking Subscore*	Coef.				0.390	-2.741	1.215
	S.E.				(2.065)	(1.500)	(2.378)
CELDT Overall Score	Coef.	1.017	-0.989	1.619	-1.028	**-5.535****	**-2.250***
	S.E.	(0.833)	(0.758)	(0.992)	(0.730)	(0.658)	(0.989)
Math_Basic	Coef.	-0.017			-0.007		
	S.E.	(0.014)			(0.014)		
Math_Proficient	Coef.	**-0.0307***			-0.023		
	S.E.	(0.016)			(0.016)		
Female	Coef.	0.028	-0.000	-0.011	-0.018	-0.002	0.007
	S.E.	(0.017)	(0.014)	(0.020)	(0.016)	(0.013)	(0.018)
Academic Year	Coef.	0.0425	**0.0955****	**0.124****	-0.046	**0.298****	**0.185****
	S.E.	(0.028)	(0.030)	(0.040)	(0.056)	(0.042)	(0.059)
Grade Level	Coef.	**0.140****	-0.020	**0.075****	-0.035	**-0.187****	0.007
	S.E.	(0.022)	(0.023)	(0.020)	(0.021)	(0.021)	(0.019)

SOURCES: Authors' calculations.

NOTES: *Listening and Speaking are combined tests in CELDT Era 1; * and ** indicate significance at the 5% and 1% levels respectively. "CELDT Era 1" is 2003-2006; "CELDT Era 2" is 2007-2012. Coefficients shaded green are positive and significant while those shaded red are negative and significant.

.

TABLE 3B

SDUSD Checks for discontinuities in background variables above the CST reclassification cutpoint

Baseline Variables		CELDT Era 1			CELDT Era 2		
		g3-5	g6-8	g9-12	g3-5	g6-8	g9-12
Spanish Home-Language	Coef.	0.0205	0.0147	-0.0709	-0.0359	0.00223	0.0304
	S.E.	-0.0411	-0.0325	-0.0503	-0.0298	-0.0267	-0.0485
CELDT Listening Subscore	Coef.	-8.689	-0.44	4.997	0.709	5.395	-6.343
	S.E.	-5.792	-3.517	-4.369	-2.991	-4.203	-6.322
CELDT Reading Subscore	Coef.	-1.891	2.193	3.367	1.491	-2.171	-6.248
	S.E.	-2.974	-1.839	-3.18	-1.866	-2.075	-4.025
CELDT Writing Subscore	Coef.	0.019	-1.486	0.279	-0.129	1.103	-6.362
	S.E.	-2.739	-2.165	-3.711	-1.746	-2.204	-4.305
CELDT Speaking Subscore*	Coef.				5.243	-2.596	-7.866
	S.E.				-3.395	-4.035	-7.363
CELDT Overall Score	Coef.	-4.81	-0.0393	3.412	1.804	0.405	**-6.651***
	S.E.	-2.999	-1.881	-2.659	-1.346	-1.671	-2.89
Math_Basic	Coef.	0.0473					
	S.E.	-0.0554					
Math_Proficient	Coef.	-0.0106					
	S.E.	-0.0491					
Female	Coef.	0.0271	0.0546	0.0874	0.0625	-0.014	0.0109
	S.E.	-0.0575	-0.044	-0.0629	-0.0354	-0.0373	-0.0628
Academic Year	Coef.	-0.0763	-0.0028	0.114	**-0.285***	0.219	0.00482
	S.E.	-0.0556	-0.0431	-0.0591	-0.118	-0.122	-0.211
Grade Level	Coef.	-0.0474	0.069	0.0946	**-0.104***	0.0211	-0.111
	S.E.	-0.0652	-0.0669	-0.0636	-0.043	-0.0568	-0.0686

SOURCES: Authors' calculations.

NOTES: *Listening and Speaking are combined tests in CELDT Era 1, * and ** indicate significance at the 5% and 1% levels respectively. "CELDT Era 1" is 2004-2005; "CELDT Era 2 is 2007-2014". Coefficients shaded green are positive and significant while those shaded red are negative and significant.

Tables 4A and 4B provide summary statistics.

TABLE 4A

LAUSD Variable Summary Statistics for the Sample from the 2nd Order Polynomial, 50 Bandwidth, CST ELA experiment

Variable		Intent to Treat					
		CELDT Era 1			CELDT Era 2		
		g3-5	g6-8	g9-12	g3-5	g6-8	g9-12
Lag CST ELA*	# Obs	40,141	37,577	16,769	48,361	41,485	20,622
	Mean	318.634	303.874	294.734	320.831	301.872	295.819
	S.D.	21.120	24.028	23.140	20.052	23.848	23.594
CST ELA 1 year ahead	# Obs	40,141	37,577	16,769	48,361	41,485	20,622
	Mean	-0.199	-0.471	-0.579	-0.422	-0.766	-0.746
	S.D.	0.599	0.573	0.577	0.612	0.563	0.582
CST ELA 2 years ahead	# Obs	36,658	33,957	8,588	37,969	32,495	9,153
	Mean	-0.200	-0.445	-0.576	-0.440	-0.692	-0.733
	S.D.	0.617	0.589	0.612	0.619	0.589	0.601
CST math 1 year ahead	# Obs	40,094	13,743		48,296	13,005	
	Mean	-0.061	-0.334		-0.276	-0.516	
	S.D.	0.744	0.651		0.753	0.664	
CST math 2 years ahead	# Obs	36,510			37,919		
	Mean	-0.126			-0.312		
	S.D.	0.716			0.730		
Graduation on time	# Obs	22,851	24,136	12,957	12,799	22,864	14,944
	Mean	0.867	0.850	0.776	0.877	0.809	0.789
	S.D.	0.339	0.357	0.417	0.329	0.393	0.408
Reclassified	# Obs	40,141	37,577	16,769	48,361	41,485	20,622
	Mean	0.429	0.438	0.255	0.803	0.497	0.376
	S.D.	0.495	0.496	0.436	0.397	0.500	0.484
Spanish Home-Language	# Obs	40,141	37,577	16,769	48,361	41,485	20,622
	Mean	0.937	0.915	0.893	0.946	0.941	0.939
	S.D.	0.243	0.278	0.309	0.226	0.236	0.239
CELDT Listening Subscore	# Obs	40,141	37,577	16,769	48,361	41,485	20,622
	Mean	581.405	577.462	561.426	563.850	618.899	636.061
	S.D.	47.954	49.546	44.032	47.723	61.795	58.584
CELDT Reading Subscore	# Obs	40,141	37,577	16,769	48,361	41,485	20,622
	Mean	541.257	547.290	560.873	543.720	577.127	601.475
	S.D.	30.841	25.534	28.208	31.724	33.731	35.298
CELDT Writing Subscore	# Obs	40,141	37,577	16,769	48,361	41,485	20,622
	Mean	553.490	555.319	552.722	544.766	566.882	583.129
	S.D.	29.514	29.250	29.917	31.158	32.384	38.798
CELDT Speaking Subscore	# Obs				48361	41485	20622
	Mean				575.147	595.482	618.051
	S.D.				62.037	58.508	64.223

CELDT Overall Score	# Obs	40,141	37,577	16,769	48,361	41,485	20,622
	Mean	564.026	564.015	558.736	556.492	589.226	609.307
	S.D.	27.589	27.559	25.920	25.017	26.474	27.771
Math Basic (Proficiency)	# Obs	40,141			48,361		
	Mean	0.862			0.878		
	S.D.	0.345			0.328		
Math Proficient	# Obs	40,141			48,361		
	Mean	0.510			0.546		
	S.D.	0.500			0.498		
Female	# Obs	40,141	37,577	16,769	48,361	41,485	20,622
	Mean	0.544	0.557	0.519	0.515	0.494	0.421
	S.D.	0.498	0.497	0.500	0.500	0.500	0.494
Academic Year (Spring)	# Obs	40,141	37,577	16,769	48,361	41,485	20,622
	Mean	2005.074	2004.763	2004.817	2009.158	2009.063	2009.249
	S.D.	0.868	1.035	0.996	1.705	1.655	1.618
Grade Level	# Obs	40,141	37,577	16,769	48,361	41,485	20,622
	Mean	4.298	6.956	9.395	4.242	7.035	9.382
	S.D.	0.728	0.829	0.496	0.752	0.816	0.501

SOURCES: Authors' calculations.

NOTES: *'Lagcstela' values are raw CST-ELA scores from the previous academic year; the various "1/2 year(s) ahead" variable values are conversions of raw CST scores into z-scores of that grade/year score distribution.

TABLE 4B

SDUSD Variable Summary Statistics for the Sample from the 2nd Order Polynomial, 50 Bandwidth, CST ELA experiment

Variable		Intent to Treat					
		CELDT Era 1			CELDT Era 2		
		g3-5	g6-8	g9-12	g3-5	g6-8	g9-12
Lag CST ELA	# Obs	3,414	3,903	1,513	3,740	6,540	1,931
	Mean	318.336	303.414	294.571	328.275	315.021	308.220
	S.D.	21.238	23.062	22.033	17.579	22.311	23.636
CST ELA 1 year ahead	# Obs	3,414	3,903	1,513	3,740	6,540	1,931
	Mean	-0.141	-0.434	-0.651	-0.168	-0.530	-0.630
	S.D.	0.585	0.582	0.560	0.596	0.576	0.611
CST ELA 2 years ahead	# Obs	3,123	3,486	838	2,721	5,026	712
	Mean	-0.137	-0.443	-0.618	-0.138	-0.491	-0.557
	S.D.	0.611	0.600	0.598	0.606	0.601	0.622
CST math 1 year ahead	# Obs	3,412	1,059		3,735	2,052	
	Mean	-0.097	-0.219		-0.158	-0.311	
	S.D.	0.713	0.722		0.734	0.659	
CST math 2 years ahead	# Obs	3,114			2,663		
	Mean	-0.083			-0.089		
	S.D.	0.724			0.708		
Graduation on time	# Obs	2,595	3,309	1,384	893	4,118	1,750
	Mean	0.756	0.648	0.611	0.853	0.789	0.725
	S.D.	0.430	0.478	0.488	0.354	0.408	0.447
Reclassified	# Obs	3,414	3,903	1,513	3,740	6,540	1,931
	Mean	0.521	0.420	0.318	0.462	0.568	0.456
	S.D.	0.500	0.494	0.466	0.499	0.495	0.498
Spanish Home-Language	# Obs	3,414	3,903	1,513	3,740	6,540	1,931
	Mean	0.818	0.837	0.827	0.760	0.821	0.812
	S.D.	0.386	0.369	0.378	0.427	0.383	0.391
CELDT Listening Subscore	# Obs	3,414	3,903	1,513	3,740	6,540	1,931
	Mean	571.362	560.584	549.286	574.421	633.206	662.786
	S.D.	48.355	42.450	35.534	42.205	57.695	48.442
CELDT Reading Subscore	# Obs	3,414	3,903	1,513	3,740	6,540	1,931
	Mean	544.500	548.212	561.727	563.969	591.867	623.542
	S.D.	28.140	23.597	24.937	28.830	29.910	33.450
CELDT Writing Subscore	# Obs	3,414	3,903	1,513	3,740	6,540	1,931
	Mean	553.016	555.810	553.547	563.146	578.771	596.373
	S.D.	24.926	26.119	28.604	25.641	30.487	34.947
CELDT Speaking Subscore	# Obs				3,740	6,540	1,931
	Mean				548.212	585.831	607.986
	S.D.				48.957	53.702	56.872
CELDT Overall Score	# Obs	3,414	3,903	1,513	3,740	6,540	1,931
	Mean	559.705	555.922	553.083	562.059	597.038	622.295

	S.D.	25.977	23.272	21.370	19.652	23.974	22.816
Math Basic (Proficiency)	# Obs	3,414					
	Mean	0.745					
	S.D.	0.436					
Math Proficient	# Obs	3,414					
	Mean	0.384					
	S.D.	0.486					
Female	# Obs	3,414	3,903	1,513	3,740	6,540	1,931
	Mean	0.512	0.476	0.472	0.505	0.480	0.387
	S.D.	0.500	0.499	0.499	0.500	0.500	0.487
Academic Year (Spring)	# Obs	3,414	3,903	1,513	3,740	6,540	1,931
	Mean	2004.511	2004.531	2004.718	2009.717	2009.592	2009.606
	S.D.	0.500	0.499	0.450	1.685	1.615	1.651
Grade Level	# Obs	3,414	3,903	1,513	3,740	6,540	1,931
	Mean	4.452	7.086	9.354	4.400	6.985	9.484
	S.D.	0.644	0.788	0.503	0.620	0.791	0.517

SOURCES: Authors' calculations.
NOTES: *'Lagcstela' values are raw CST-ELA scores from the previous academic year; the various "1/2 year(s) ahead" variable values are conversions of raw CST scores into z-scores of that grade/year score distribution.

Table 5A

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

Outcome/Model			Intent to Treat				Treatment on Treated			
			4th	3rd	2nd	1st	4th	3rd	2nd	1st
CST ELA 1 Year Later	CELDT Era 1 ('03-'06)	Elm	-0.0388	-0.0332	-0.0286	**-0.0221***	-0.0855	-0.0726	-0.0605	**-0.0459***
			(0.0289)	(0.022)	(0.0162)	(0.0111)	(0.0638)	(0.0481)	(0.0343)	(0.0232)
		MS	0.0032	0.0068	0.0085	0.0167	0.0048	0.0103	0.0130	0.0251
			(0.0252)	(0.0185)	(0.0134)	(0.00902)	(0.0381)	(0.0283)	(0.0205)	(0.0135)
		HS	0.0099	0.0052	0.0033	0.0067	0.0185	0.0094	0.0062	0.0122
			(0.0331)	(0.026)	(0.0196)	(0.0136)	(0.0616)	(0.0471)	(0.0362)	(0.0248)
	CELDT Era 2 ('07-'12)	Elm	**-0.0825****	**-0.0818****	**-0.0742****	**-0.0389****	**-0.113****	**-0.110****	**-0.0962****	**-0.0498****
			(0.0292)	(0.0229)	(0.0172)	(0.0119)	(0.0406)	(0.031)	(0.0224)	(0.0153)
		MS	0.0023	-0.0150	0.0112	0.0066	0.0026	-0.0166	0.0124	0.0073
			(0.0204)	(0.0162)	(0.0124)	(0.00864)	(0.0227)	(0.0179)	(0.0137)	(0.00955)
		HS	0.0345	0.0344	0.0145	0.0199	0.0424	0.0421	0.0175	0.0239
			(0.0341)	(0.0261)	(0.0194)	(0.0134)	(0.0419)	(0.0319)	(0.0235)	(0.016)
CST ELA 2 Years Later	CELDT Era 1 ('03-'06)	Elm	-0.0164	-0.0221	**-0.0357***	-0.0160	-0.0345	-0.0462	**-0.0720***	-0.0320
			(0.0317)	(0.024)	(0.0177)	(0.0122)	(0.0668)	(0.0501)	(0.0357)	(0.0244)
		MS	0.0013	-0.0188	-0.0271	-0.0024	0.0020	-0.0286	-0.0412	-0.0037
			(0.0278)	(0.0206)	(0.0149)	(0.01)	(0.0422)	(0.0314)	(0.0227)	(0.015)
		HS	-0.0229	-0.0125	-0.0209	-0.0107	-0.0443	-0.0234	-0.0408	-0.0202
			(0.0486)	(0.0384)	(0.0291)	(0.0203)	(0.0942)	(0.0719)	(0.0567)	(0.0384)
	CELDT Era 2 ('07-'12)	Elm	-0.0120	-0.0046	0.0025	-0.0014	-0.0165	-0.0062	0.0032	-0.0017
			(0.0348)	(0.0271)	(0.0202)	(0.014)	(0.048)	(0.0362)	(0.0262)	(0.0177)
		MS	-0.0349	-0.0216	0.0013	-0.0081	-0.0391	-0.0240	0.0014	-0.0090
			(0.0252)	(0.02)	(0.0153)	(0.0106)	(0.0282)	(0.0222)	(0.017)	(0.0118)
		HS	0.0638	0.0289	-0.0245	-0.0112	0.0803	0.0365	-0.0305	-0.0139
			(0.0524)	(0.0403)	(0.0303)	(0.0211)	(0.0657)	(0.0508)	(0.0377)	(0.026)
CST Math 1 Year Later	CELDT Era 1 ('03-'06)	Elm	-0.0023	-0.0002	0.0105	**-0.0271***	-0.0052	-0.0004	0.0222	**-0.0563***
			(0.0349)	(0.0262)	(0.0194)	(0.0134)	(0.0768)	(0.0572)	(0.0409)	(0.0278)
		MS	0.0883	0.0341	0.0243	0.0218	0.1430	0.0566	0.0399	0.0346

			(0.0526)	(0.0400)	(0.0298)	(0.0206)	(0.085)	(0.0664)	(0.049)	(0.0327)
		HS								
	CELDT Era 2 ('07-'12)	Elm	**-0.0790***	**-0.0658***	-0.0247	**-0.0580****	**-0.109***	**-0.0883***	-0.0320	**-0.0744****
			(0.0343)	(0.0269)	(0.0203)	(0.014)	(0.0475)	(0.0362)	(0.0264)	(0.0179)
		MS	-0.0057	-0.0110	0.0077	0.0066	-0.0065	-0.0125	0.0089	0.0077
			(0.0500)	(0.0387)	(0.0291)	(0.0204)	(0.0574)	(0.0442)	(0.0338)	(0.0238)
		HS								
CST Math 2 Years Later	CELDT Era 1 ('03-'06)	Elm	0.0339	0.0251	0.0237	-0.0064	0.0713	0.0523	0.0478	-0.0129
			(0.0364)	(0.0273)	(0.0201)	(0.0138)	(0.0766)	(0.0569)	(0.0404)	(0.0277)
		MS								
		HS								
	CELDT Era 2 ('07-'12)	Elm	-0.0260	-0.0127	-0.0027	-0.0278	-0.0359	-0.0170	-0.0036	-0.0353
			(0.0396)	(0.0305)	(0.0227)	(0.0155)	(0.0547)	(0.0407)	(0.0293)	(0.0196)
		MS								
		HS								
Graduation on Time	CELDT Era 1 ('03-'06)	Elm	-0.0530	-0.0405	**-0.0315***	-0.0192	-0.1070	-0.0805	**-0.0625***	-0.0384
			(0.0273)	(0.0209)	(0.0157)	(0.011)	(0.0551)	(0.0415)	(0.0312)	(0.0221)
		MS	0.0155	0.0087	0.0148	**0.0186***	0.0235	0.0131	0.0224	**0.0273***
			(0.0242)	(0.018)	(0.013)	(0.00877)	(0.0367)	(0.0271)	(0.0196)	(0.0129)
		HS	**-0.0628****	**-0.0365***	**-0.0319***	-0.0049	**-0.110****	**-0.0619***	**-0.0549***	-0.0085
			(0.0235)	(0.0182)	(0.0136)	(0.0094)	(0.0413)	(0.0309)	(0.0235)	(0.0162)
	CELDT Era 2 ('07-'12)	Elm	-0.0230	-0.0177	-0.0044	0.0045	-0.0255	-0.0196	-0.0049	0.0051
			(0.0385)	(0.0297)	(0.0219)	(0.0151)	(0.0427)	(0.0328)	(0.0243)	(0.0171)
		MS	-0.0068	-0.0021	-0.0035	-0.0110	-0.0075	-0.0023	-0.0038	-0.0119
			(0.022)	(0.0177)	(0.0136)	(0.00932)	(0.0241)	(0.0193)	(0.0148)	(0.0101)

		HS	-0.0275	-0.0053	0.0090	0.0029	-0.0317	-0.0062	0.0103	0.0033
			(0.0208)	(0.0161)	(0.0119)	(0.0082)	(0.0241)	(0.0186)	(0.0137)	(0.0093)

SOURCES: Authors' estimations.

NOTES: Model of CST Math 1 Year Later for middle school students is only for students whose reclassification decision was in grade 6. Coefficients shaded green are positive and significant while those shaded red are negative and significant.

TABLE 5B

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

Outcome/Model			Intent to Treat				Treatment on Treated			
			4th	3rd	2nd	1st	4th	3rd	2nd	1st
CST ELA 1 Year Later	CELDT Era 1 ('04-'05)	Elm	-0.0332	-0.0250	-0.0034	-0.0304	-0.0569	-0.0434	-0.00571	-0.0510
			(0.0993)	(0.0725)	(0.0526)	(0.0358)	(0.170)	(0.125)	(0.0879)	(0.0599)
		MS	0.0383	-0.00883	-0.033	-0.0026	0.0621	-0.0139	-0.0507	-0.0038
			(0.0766)	(0.0567)	(0.0418)	(0.0287)	(0.124)	(0.0893)	(0.0638)	(0.0427)
		HS	-0.0624	-0.171*	-0.138*	-0.0794	-0.0837	-0.216*	-0.177*	-0.102
			(0.110)	(0.0869)	(0.0668)	(0.0458)	(0.146)	(0.108)	(0.0849)	(0.0587)
	CELDT Era 2 ('07-'14)	Elm	0.0415	-0.000182	-0.0087	-0.0113	0.0896	-0.00040	-0.0174	-0.0195
			(0.0648)	(0.0491)	(0.0362)	(0.0244)	(0.140)	(0.108)	(0.0727)	(0.0422)
		MS	-0.0674	0.0183	-0.0185	-0.0320	-0.0978	0.0255	-0.0263	-0.0484
			(0.0670)	(0.0536)	(0.0400)	(0.0271)	(0.0969)	(0.0743)	(0.0566)	(0.0408)
		HS	-0.101	0.0800	-0.0557	-0.0481	-0.158	0.123	-0.0794	-0.0691
			(0.122)	(0.0896)	(0.0663)	(0.0468)	(0.189)	(0.137)	(0.0945)	(0.0670)
CST ELA 2 Years Later	CELDT Era 1 ('04-'05)	Elm	-0.120	-0.0947	-0.124*	-0.0744	-0.219	-0.171	-0.214*	-0.127
			(0.104)	(0.0786)	(0.0582)	(0.0404)	(0.191)	(0.144)	(0.102)	(0.0690)
		MS	0.0729	0.0157	-0.0649	-0.0320	0.121	0.0252	-0.0995	-0.0482
			(0.0901)	(0.0662)	(0.0484)	(0.0324)	(0.149)	(0.106)	(0.0743)	(0.0488)
		HS	0.327*	0.265*	0.129	0.134*	0.464*	0.346*	0.171	0.178*
			(0.146)	(0.114)	(0.0877)	(0.0632)	(0.218)	(0.154)	(0.117)	(0.0843)
	CELDT Era 2 ('07-'14)	Elm	-0.0102	-0.0206	-0.0453	-0.0432	-0.0186	-0.0393	-0.0827	-0.0711
			(0.0780)	(0.0590)	(0.0435)	(0.0290)	(0.142)	(0.112)	(0.0793)	(0.0477)
		MS	0.273**	0.170**	0.114*	0.00717	0.409**	0.238**	0.167*	0.0113
			(0.0760)	(0.0606)	(0.0466)	(0.0326)	(0.115)	(0.0851)	(0.0684)	(0.0513)
		HS	-0.227	-0.0434	-0.0515	-0.0988	-0.355	-0.0659	-0.0710	-0.143
			(0.198)	(0.160)	(0.118)	(0.0791)	(0.305)	(0.239)	(0.161)	(0.113)
CST Math 1 Year Later	CELDT Era 1 ('04-'05)	Elm	0.0181	0.0848	0.0302	-0.0306	0.0310	0.147	0.0505	-0.0512
			(0.123)	(0.0889)	(0.0644)	(0.0444)	(0.209)	(0.153)	(0.107)	(0.0743)

		MS	0.143	0.0116	-0.0267	0.0418	0.211	0.0164	-0.0396	0.0598
			(0.176)	(0.141)	(0.110)	(0.0784)	(0.254)	(0.197)	(0.163)	(0.112)
		HS								
	CELDT Era 2 ('07-'14)	Elm	-0.0689	-0.119	-0.0712	-0.0368	-0.149	-0.262	-0.142	-0.0637
			(0.0814)	(0.0636)	(0.0474)	(0.0321)	(0.176)	(0.141)	(0.0948)	(0.0556)
		MS	-0.397*	-0.265	-0.0709	-0.0857	-0.501*	-0.346	-0.106	-0.151
			(0.189)	(0.145)	(0.106)	(0.0676)	(0.241)	(0.189)	(0.158)	(0.119)
		HS								
CST Math 2 Years Later	CELDT Era 1 ('04-'05)	Elm	-0.0902	-0.0162	-0.0651	-0.0327	-0.165	-0.0293	-0.112	-0.0558
			(0.123)	(0.0900)	(0.0660)	(0.0463)	(0.226)	(0.163)	(0.114)	(0.0789)
		MS								
		HS								
	CELDT Era 2 ('07-'14)	Elm	-0.0892	-0.0956	-0.0460	-0.0489	-0.162	-0.180	-0.0828	-0.0802
			(0.107)	(0.0796)	(0.0570)	(0.0380)	(0.194)	(0.150)	(0.102)	(0.0621)
		MS								
		HS								
Graduation on Time	CELDT Era 1 ('04-'05)	Elm	-0.00236	0.0940	0.0666	0.0499	- 0.00409	0.167	0.113	0.0841
			(0.106)	(0.0790)	(0.0582)	(0.0396)	(0.184)	(0.143)	(0.0991)	(0.0669)
		MS	-0.0964	0.0318	-0.0398	-0.0132	-0.163	0.0513	-0.0615	-0.0201
			(0.0850)	(0.0623)	(0.0454)	(0.0302)	(0.144)	(0.100)	(0.0701)	(0.0459)
		HS	0.145	0.0709	0.0657	0.0426	0.196	0.0908	0.0858	0.0553
			(0.102)	(0.0799)	(0.0605)	(0.0412)	(0.137)	(0.102)	(0.0787)	(0.0533)
	CELDT Era 2 ('07-'14)	Elm	0.142	0.103	0.0289	0.0153	0.291	0.242	0.0721	0.0335
			(0.112)	(0.0795)	(0.0550)	(0.0353)	(0.238)	(0.193)	(0.137)	(0.0768)

		MS	0.00220	0.0225	-0.00984	-0.00789	0.00318	0.0312	-0.0138	-0.0121
			(0.0659)	(0.0522)	(0.0393)	(0.0268)	(0.0953)	(0.0718)	(0.0552)	(0.0408)
		HS	0.0476	0.0267	0.00961	-0.00643	0.108	0.0558	0.0189	-0.0119
			(0.0543)	(0.0416)	(0.0309)	(0.0207)	(0.123)	(0.0865)	(0.0607)	(0.0384)

SOURCES: Authors' estimations

NOTES: Model of CST Math 1 Year Later for middle school students is only for students whose reclassification decision was in grade 6. Coefficients shaded green are positive and significant while those shaded red are negative and significant.

Tables 6, 7A and 7B show similar results for the various experiments using the CELDT overall proficiency level and the CELDT Reading Level as RD experiments.

TABLE 6

LAUSD Robustness of the Estimates to the Order of Polynomials in the Running Variable CELDT OPL, whole level as bandwidth

Outcome/Model			Intent to Treat				Treatment on Treated			
			4th	3rd	2nd	1st	4th	3rd	2nd	1st
CST ELA 1 Year Later	CELDT Era 1 ('03-'06)	Elm	0.0008	0.0083	-0.0048	0.0027	0.0027	0.0250	-0.0146	0.0076
			(0.0278)	(0.0218)	(0.0165)	(0.0114)	(0.0879)	(0.0652)	(0.0499)	(0.0327)
		MS	-0.0222	-0.0132	0.0020	-0.0067	-0.0440	-0.0274	0.0043	-0.0138
			(0.0467)	(0.0366)	(0.0282)	(0.0193)	(0.0924)	(0.0760)	(0.0588)	(0.0398)
		HS	0.0428	-0.0013	0.0440	-0.0140	0.0857	-0.0027	0.0926	-0.0313
			(0.1050)	(0.0799)	(0.0580)	(0.0396)	(0.2090)	(0.1650)	(0.1220)	(0.0883)
	CELDT Era 2 ('07-'12)	Elm	-0.0151	-0.0208	-0.0201	0.0112	-0.0182	-0.0249	-0.0238	0.0131
			(0.0191)	(0.0152)	(0.0116)	(0.0081)	(0.0230)	(0.0181)	(0.0137)	(0.0095)
		MS	**-0.0817***	-0.0430	**-0.0410***	**-0.0382****	**-0.128***	-0.0686	**-0.0659***	**-0.0620****
			(0.0320)	(0.0256)	(0.0195)	(0.0136)	(0.0510)	(0.0411)	(0.0316)	(0.0221)
		HS	-0.0038	0.0064	-0.0151	-0.0099	-0.0048	0.0084	-0.0204	-0.0134
			(0.0525)	(0.0423)	(0.0324)	(0.0225)	(0.0668)	(0.0553)	(0.0436)	(0.0305)
CST ELA 2 Years Later	CELDT Era 1 ('03-'06)	Elm	-0.0266	-0.0066	-0.0052	0.0077	-0.0770	-0.0186	-0.0151	0.0214
			(0.0322)	(0.0251)	(0.0189)	(0.0128)	(0.0935)	(0.0712)	(0.0543)	(0.0356)
		MS	-0.0127	-0.0200	0.0284	0.0052	-0.0250	-0.0407	0.0586	0.0107
			(0.0479)	(0.0378)	(0.0291)	(0.0205)	(0.0941)	(0.0769)	(0.0597)	(0.0418)
		HS	0.0649	0.0842	0.0782	0.0672	0.1350	0.1770	0.1670	0.1610
			(0.1230)	(0.0988)	(0.0752)	(0.0527)	(0.2560)	(0.2090)	(0.1610)	(0.1260)
	CELDT Era 2 ('07-'12)	Elm	-0.0130	-0.0294	-0.0236	0.0012	-0.0157	-0.0352	-0.0279	0.0014
			(0.0229)	(0.0182)	(0.0139)	(0.0097)	(0.0276)	(0.0218)	(0.0165)	(0.0114)
		MS	-0.0384	-0.0392	-0.0359	**-0.0375***	-0.0612	-0.0650	-0.0603	**-0.0624***
			(0.0367)	(0.0297)	(0.0230)	(0.0162)	(0.0588)	(0.0495)	(0.0388)	(0.0272)
		HS	-0.1090	-0.0051	0.0005	0.0007	-0.1440	-0.0065	0.0007	0.0010

			(0.0857)	(0.0693)	(0.0527)	(0.0370)	(0.1140)	(0.0879)	(0.0712)	(0.0503)
CST Math 1 Year Later	CELDT Era 1 ('03-'06)	Elm	-0.0236	-0.0278	-0.0124	0.0059	-0.0707	-0.0811	-0.0367	0.0167
			(0.0378)	(0.0297)	(0.0222)	(0.0151)	(0.1130)	(0.0870)	(0.0657)	(0.0431)
		MS	0.0437	0.0150	0.0079	0.0083	0.0860	0.0315	0.0178	0.0183
			(0.0681)	(0.0539)	(0.0409)	(0.0283)	(0.1340)	(0.1130)	(0.0915)	(0.0625)
		HS								
	CELDT Era 2 ('07-'12)	Elm	0.0306	0.0220	0.0057	0.0281**	0.0367	0.0263	0.0068	0.0330**
			(0.0252)	(0.0201)	(0.0154)	(0.0107)	(0.0303)	(0.0240)	(0.0182)	(0.0125)
		MS	-0.0999	-0.0422	-0.0246	-0.0483*	-0.1950	-0.0819	-0.0484	-0.0959*
			(0.0596)	(0.0466)	(0.0353)	(0.0241)	(0.1190)	(0.0910)	(0.0696)	(0.0481)
		HS								
CST Math 2 Years Later	CELDT Era 1 ('03-'06)	Elm	-0.0174	-0.0175	-0.0040	0.0037	-0.0506	-0.0497	-0.0116	0.0103
			(0.0396)	(0.0309)	(0.0230)	(0.0158)	(0.1150)	(0.0876)	(0.0662)	(0.0440)
		MS								
		HS								
	CELDT Era 2 ('07-'12)	Elm	0.0146	0.0076	0.0123	0.0306*	0.0176	0.0090	0.0146	0.0357*
			(0.0295)	(0.0233)	(0.0177)	(0.0123)	(0.0356)	(0.0279)	(0.0209)	(0.0144)
		MS								
		HS								
Graduation on Time	CELDT Era 1 ('03-'06)	Elm	(0.0176)	(0.0024)	(0.0210)	(0.0199)	(0.0459)	(0.0063)	(0.0558)	(0.0516)
			-0.0247	-0.0193	-0.0147	-0.0103	-0.0646	-0.0502	-0.0390	-0.0267
		MS	(0.0270)	(0.0407)	(0.0245)	(0.0099)	(0.0526)	(0.0806)	(0.0507)	(0.0206)
			-0.0355	-0.0281	-0.0219	-0.0157	-0.0692	-0.0561	-0.0455	-0.0326
		HS	(0.1020)	(0.0414)	(0.0463)	(0.0473)	(0.2080)	(0.0862)	(0.0966)	(0.0966)
			-0.0667	-0.0514	-0.0374	-0.0260	-0.1360	-0.1060	-0.0777	-0.0531

	CELDT Era 2 ('07-'12)	Elm	0.0158	0.0165	0.0109	0.0176	0.0212	0.0223	0.0143	0.0226
			(0.0272)	(0.0218)	(0.0170)	(0.0122)	(0.0366)	(0.0295)	(0.0223)	(0.0156)
		MS	0.0188	0.0207	0.0075	-0.0044	0.0289	0.0326	0.0118	-0.0071
			(0.0326)	(0.0263)	(0.0202)	(0.0143)	(0.0501)	(0.0414)	(0.0321)	(0.0229)
		HS	0.0077	0.0209	0.0143	0.0114	0.0095	0.0263	0.0182	0.0144
			(0.0318)	(0.0253)	(0.0193)	(0.0137)	(0.0391)	(0.0318)	(0.0246)	(0.0172)

SOURCES: Authors' estimations

NOTES: Model of CST Math 1 Year Later for middle school students is only for students whose reclassification decision was in grade 6. Coefficients shaded green are positive and significant while those shaded red are negative and significant.

TABLE 7A

LAUSD Robustness of the Estimates to the Order of Polynomials in the Running Variable CELDT Reading Subtest, whole level as bandwidth

Outcome/Model			Intent to Treat				Treatment on Treated			
			4th	3rd	2nd	1st	4th	3rd	2nd	1st
CST ELA 1 Year Later	CELDT Era 1 ('03-'06)	Elm	-0.0687	-0.0518	0.00455	0.00367	-0.242	-0.148	0.0150	0.0116
			(0.0599)	(0.0402)	(0.0280)	(0.0188)	(0.213)	(0.115)	(0.0923)	(0.0598)
		MS	-0.111	-0.141	-0.133	-0.106	-0.146	-0.214	-0.270	-0.224
			(0.127)	(0.0972)	(0.0752)	(0.0557)	(0.171)	(0.153)	(0.162)	(0.123)
		HS	-0.927	**-1.209****	-0.560	-0.514	-2.224	**-1.727***	-1.124	-1.034
			(0.522)	(0.434)	(0.380)	(0.266)	(1.699)	(0.749)	(0.811)	(0.556)
	CELDT Era 2 ('07-'12)	Elm	0.0445	-0.0230	0.0214	-0.0344	0.0511	-0.0266	0.0249	-0.0393
			(0.0471)	(0.0380)	(0.0289)	(0.0202)	(0.0541)	(0.0441)	(0.0337)	(0.0231)
		MS	-0.0578	-0.0546	-0.0453	0.00203	-0.0777	-0.0696	-0.0593	0.00282
			(0.0957)	(0.0695)	(0.0511)	(0.0365)	(0.128)	(0.0882)	(0.0670)	(0.0508)
		HS	0.0328	-0.0267	0.00479	0.00952	0.0465	-0.0369	0.00620	0.0124
			(0.0867)	(0.0713)	(0.0559)	(0.0384)	(0.123)	(0.0985)	(0.0721)	(0.0500)
CST ELA 2 Years Later	CELDT Era 1 ('03-'06)	Elm	-0.00239	0.0350	0.0481	0.0104	-0.00861	0.100	0.156	0.0323
			(0.0663)	(0.0444)	(0.0306)	(0.0205)	(0.238)	(0.127)	(0.0992)	(0.0639)
		MS	-0.0540	-0.0699	-0.0726	-0.116	-0.0675	-0.101	-0.139	-0.258
			(0.148)	(0.109)	(0.0827)	(0.0608)	(0.186)	(0.160)	(0.162)	(0.143)
		HS	-0.239	-0.363	-0.0927	0.0690	-0.408	-0.704	-0.128	0.138
			(0.536)	(0.428)	(0.341)	(0.255)	(0.900)	(0.829)	(0.461)	(0.505)
	CELDT Era 2 ('07-'12)	Elm	**0.114***	0.0567	0.0503	0.00197	**0.129***	0.0655	0.0579	0.00226
			(0.0577)	(0.0470)	(0.0359)	(0.0251)	(0.0655)	(0.0542)	(0.0413)	(0.0287)
		MS	-0.138	-0.106	-0.113	-0.0598	-0.212	-0.145	-0.155	-0.0873
			(0.121)	(0.0864)	(0.0678)	(0.0486)	(0.185)	(0.118)	(0.0933)	(0.0711)
		HS	-0.00216	-0.00693	0.0346	0.0330	-0.00405	-0.0131	0.0515	0.0498
			(0.127)	(0.107)	(0.0846)	(0.0579)	(0.237)	(0.203)	(0.125)	(0.0870)
CST Math 1 Year Later		Elm	-0.0704	**-0.111***	-0.0316	0.0109	-0.249	**-0.317***	-0.104	0.0346
			(0.0808)	(0.0546)	(0.0377)	(0.0252)	(0.287)	(0.157)	(0.124)	(0.0798)

	CELDT Era 1 ('03-'06)	MS	0.144	-0.0940	0.0511	-0.133	0.198	-0.138	0.104	-0.291
			(0.173)	(0.131)	(0.0981)	(0.0739)	(0.238)	(0.194)	(0.198)	(0.172)
		HS								
	CELDT Era 2 ('07-'12)	Elm	**0.125***	-0.000379	0.0159	-0.00401	**0.143***	-0.000437	0.0185	-0.00457
			(0.0630)	(0.0515)	(0.0395)	(0.0281)	(0.0722)	(0.0593)	(0.0459)	(0.0320)
		MS	0.160	-0.0543	-0.0445	-0.0417	0.199	-0.0745	-0.0647	-0.0690
			(0.195)	(0.135)	(0.0936)	(0.0640)	(0.251)	(0.185)	(0.136)	(0.106)
		HS								
CST Math 2 Years Later	CELDT Era 1 ('03-'06)	Elm	-0.0493	-0.0633	-0.00559	0.00206	-0.178	-0.183	-0.0182	0.00645
			(0.0826)	(0.0564)	(0.0391)	(0.0260)	(0.300)	(0.163)	(0.127)	(0.0815)
		MS								
		HS								
	CELDT Era 2 ('07-'12)	Elm	0.0438	-0.0117	-0.0300	0.00447	0.0497	-0.0135	-0.0345	0.00511
			(0.0758)	(0.0611)	(0.0466)	(0.0330)	(0.0860)	(0.0706)	(0.0537)	(0.0378)
		MS								
		HS								
Graduation on Time	CELDT Era 1 ('03-'06)	Elm	-0.0470	-0.0340	-0.00120	-0.0195	-0.183	-0.106	-0.00405	-0.0585
			(0.0555)	(0.0369)	(0.0261)	(0.0175)	(0.218)	(0.115)	(0.0882)	(0.0528)
		MS	0.0445	**-0.173***	-0.0241	-0.0789	0.0523	-0.254	-0.0461	-0.176
			(0.119)	(0.0881)	(0.0635)	(0.0484)	(0.139)	(0.140)	(0.123)	(0.114)
		HS	-0.231	-0.473	**-0.428***	-0.168	-0.410	-1.015	**-0.652***	-0.362
			(0.347)	(0.267)	(0.198)	(0.144)	(0.612)	(0.589)	(0.296)	(0.309)
	CELDT Era 2 ('07-'12)	Elm	-0.0408	0.0657	-0.0673	-0.0429	0.557	0.0789	-0.0869	-0.0515

		(0.127)	(0.0929)	(0.0631)	(0.0419)	(0.621)	(0.112)	(0.0817)	(0.0504)
	MS	-0.00850	0.0483	**0.135***	0.0484	-0.0113	0.0606	**0.177***	0.0701
		(0.111)	(0.0827)	(0.0612)	(0.0440)	(0.147)	(0.104)	(0.0811)	(0.0639)
	HS	-0.0559	-0.00553	-0.000462	0.0248	-0.0699	-0.00672	-0.000555	0.0304
		(0.0546)	(0.0435)	(0.0332)	(0.0239)	(0.0687)	(0.0529)	(0.0399)	(0.0292)

SOURCES: Authors' estimates.

NOTES: Model of CST Math 1 Year Later for middle schoolers is only for grade 6. Coefficients shaded green are positive and significant while those shaded red are negative and significant.

TABLE 7B

SDUSD Robustness of the Estimates to the Order of Polynomials in the Running Variable CELDT Reading Subtest, whole level as bandwidth

Outcome/Model			Intent to Treat				Treatment on Treated			
			4th	3rd	2nd	1st	4th	3rd	2nd	1st
CST 1 Year Later	CELDT Era 2 ('07-'14)	Elm	**-0.0896***	**-0.0799****	-0.0352	0.00310	**-0.271***	**-0.232***	-0.109	0.00896
			(0.0378)	(0.0302)	(0.0231)	(0.0162)	(0.119)	(0.0900)	(0.0719)	(0.0468)
		MS	0.0441	-0.0115	-0.0154	-0.00876	0.197	-0.0490	-0.0568	-0.0323
			(0.0615)	(0.0472)	(0.0349)	(0.0233)	(0.277)	(0.201)	(0.129)	(0.0859)
		HS	-0.0655	0.0977	0.0429	0.0774	-0.386	0.319	0.147	0.304
			(0.146)	(0.0984)	(0.0679)	(0.0451)	(0.877)	(0.343)	(0.235)	(0.185)
CST 2 Years Later	CELDT Era 2 ('07-'14)	Elm	-0.00597	-0.0137	0.0417	**0.0502***	-0.0185	-0.0413	0.132	**0.146***
			(0.0451)	(0.0365)	(0.0281)	(0.0195)	(0.140)	(0.110)	(0.0896)	(0.0573)
		MS	-0.0499	-0.0842	-0.0198	-0.00237	-0.182	-0.306	-0.0673	-0.00811
			(0.0709)	(0.0550)	(0.0410)	(0.0284)	(0.264)	(0.211)	(0.140)	(0.0971)
		HS	-0.247	-0.302	-0.206	-0.0913	-0.602	-0.745	-0.526	-0.436
			(0.244)	(0.170)	(0.118)	(0.0801)	(0.591)	(0.447)	(0.306)	(0.392)
CST Math 1 Year Later	CELDT Era 2 ('07-'14)	Elm	-0.0602	-0.0575	-0.0202	0.00479	-0.182	-0.166	-0.0625	0.0138
			(0.0546)	(0.0431)	(0.0321)	(0.0227)	(0.166)	(0.126)	(0.0992)	(0.0654)
		MS	0.123	**0.214***	0.0822	0.0639	0.532	0.895	0.338	0.245
			(0.132)	(0.0970)	(0.0697)	(0.0453)	(0.594)	(0.473)	(0.292)	(0.175)
		HS								

CST Math 2 Years Later	CELDT Era 2 ('07-'14)	Elm	-0.0447	-0.0613	-0.00948	0.0175	-0.139	-0.183	-0.0295	0.0505
			(0.0654)	(0.0518)	(0.0393)	(0.0275)	(0.204)	(0.156)	(0.122)	(0.0796)
		MS								
		HS								
Graduation on Time	CELDT Era 2 ('07-'14)	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	0.0292	-0.0374	-0.0167	0.0127	0.118	-0.144	-0.0646	0.0475
			(0.0642)	(0.0493)	(0.0356)	(0.0239)	(0.259)	(0.193)	(0.137)	(0.0898)
		HS	0.00402	-0.0224	-0.00220	-0.0217	0.0166	-0.0843	-0.00836	-0.0867
			(0.106)	(0.0644)	(0.0421)	(0.0267)	(0.435)	(0.244)	(0.160)	(0.108)

SOURCES: Authors' estimations

NOTES: Model of CST Math 1 Year Later for middle schoolers is only for grade 6. Coefficients shaded green are positive and significant while those shaded red are negative and significant.

Tables 8A and 8B below shows the relevant coefficients from a series of models of outcomes on the indicator for being above the CST ELA cutpoint, the listed moderator, and the interaction between the above indicator and the moderator. Thus each trio of coefficients in a column refers to a different model than the other trios of coefficients in the same column.

For LAUSD, there is a pattern where the indicator for being at or above the CST cutpoint is smaller (or more negative) when the student's home language is Spanish, when there is greater language homogeneity among ELs at the school, or when Spanish is more predominant in the neighborhood of the school or in the neighborhoods from which a school's students come.

TABLE 8A

Estimates of the effect of moderator variables on CST ELA one year after reclassification, CST as running variable, each trio of coefficients is a uniquely estimated model, LAUSD, 2003-12

	2003-06			2007-12		
	Elementary	Middle School	High School	Elementary	Middle School	High School
above CST cutpoint	-0.0986**	0.0650**	0.0488	-0.138**	0.120**	0.137**
	(0.0377)	(0.0233)	(0.0314)	(0.0450)	(0.0289)	(0.0408)
Spanish * above	0.0734*	-0.163**	-0.205**	0.0664	-0.114**	-0.131**
	(0.0355)	(0.0169)	(0.0195)	(0.0428)	(0.0272)	(0.0375)
Spanish Home-Language	-0.280**	-0.0611**	-0.0508	-0.327**	-0.177**	-0.163**
	(0.0339)	(0.0204)	(0.0270)	(0.0416)	(0.0241)	(0.0298)
above CST cutpoint	0.00691	0.0315	-0.0180	-0.113**	0.0396*	0.0431
	(0.0307)	(0.0209)	(0.0303)	(0.0306)	(0.0190)	(0.0282)
% EL * above	-0.0606	-0.0547	0.0621	0.0791	-0.0894*	-0.101
	(0.0442)	(0.0380)	(0.0674)	(0.0511)	(0.0454)	(0.0708)
% EL	-0.0144	0.0366	-0.132**	-0.0621	-0.00563	-0.00448
	(0.0408)	(0.0299)	(0.0433)	(0.0483)	(0.0370)	(0.0494)
above CST cutpoint	-0.0449	0.0503*	0.0484	-0.109**	0.0945**	0.115**
	(0.0299)	(0.0197)	(0.0278)	(0.0321)	(0.0214)	(0.0317)
same language as all students*above	0.000222	-0.000619**	-0.000716*	0.000490	-0.00118**	-0.00147**
	(0.000337)	(0.000213)	(0.000303)	(0.000378)	(0.000246)	(0.000359)
same language as all student	-0.000923**	0.000262	-0.000434	-0.00143**	-0.000178	0.000989**
	(0.000328)	(0.000195)	(0.000254)	(0.000365)	(0.000212)	(0.000288)

above CST cutpoint	-0.0995**	0.0677**	0.0650*	-0.136**	0.128**	0.152**
	(0.0381)	(0.0236)	(0.0317)	(0.0431)	(0.0280)	(0.0401)
same language as EL students*above	0.000769*	-0.000673**	-0.000743*	0.000661	-0.00128**	-0.00155**
	(0.000370)	(0.000218)	(0.000294)	(0.000420)	(0.000274)	(0.000390)
same language as EL students	-0.00108**	-0.000369	-0.00177**	-0.00239**	-0.00107**	0.00148**
	(0.000414)	(0.000340)	(0.000484)	(0.000446)	(0.000346)	(0.000564)
above CST cutpoint				-0.146**	0.119**	0.228**
				(0.0556)	(0.0364)	(0.0574)
language homogeneity of ELs * above				0.0802	-0.119**	-0.242**
				(0.0565)	(0.0377)	(0.0604)
language homogeneity of ELs				-0.258**	-0.0616*	0.197**
				(0.0542)	(0.0306)	(0.0404)
above CST cutpoint	-0.0151	0.00972	-0.0894	-0.103**	0.0307	-0.0134
	(0.0257)	(0.0247)	(0.0478)	(0.0295)	(0.0243)	(0.0401)
% teachers have MA * above	-0.0521	-0.00686	0.296*	0.0849	-0.0543	0.0692
	(0.0778)	(0.0780)	(0.139)	(0.0730)	(0.0582)	(0.0869)
% teachers have MA	0.0589	0.259**	0.0531	-0.0784	0.0383	-0.176**
	(0.0714)	(0.0608)	(0.0879)	(0.0689)	(0.0489)	(0.0654)
above CST cutpoint	-0.0626	-0.0706*	-0.0494	0.00375	0.0243	-0.00692
	(0.0332)	(0.0291)	(0.0490)	(0.0353)	(0.0237)	(0.0360)
median years teaching experience * above	0.00290	0.00814**	0.00468	-0.00547*	-0.00116	0.00177
	(0.00249)	(0.00265)	(0.00400)	(0.00214)	(0.00177)	(0.00255)
median years teaching experience	-0.00161	-0.00456*	-0.00185	0.00636**	-0.00108	0.00138
	(0.00227)	(0.00207)	(0.00262)	(0.00209)	(0.00157)	(0.00185)
above CST cutpoint	-0.0293	0.00551	-0.00635	-0.0740**	-0.0318	0.0247
	(0.0164)	(0.0193)	(0.0470)	(0.0175)	(0.0180)	(0.0345)
% no ELD at school * above	0.0495	0.00856	0.0218	-0.00586	0.103**	-0.0275
	(0.142)	(0.0382)	(0.0957)	(0.145)	(0.0313)	(0.0757)
% no ELD at school	0.178	0.0138	-0.0535	0.199	-0.0738**	-0.0887
	(0.134)	(0.0291)	(0.0612)	(0.135)	(0.0243)	(0.0497)
above CST cutpoint	-0.00918	0.0806*	0.0306	-0.0672*	0.0564**	0.0843**
	(0.0393)	(0.0325)	(0.0477)	(0.0280)	(0.0176)	(0.0266)

% home language Spanish school zip neighborhood * above	-0.0636	-0.0663	-0.0657	-0.00589	-0.0730**	-0.120**
	(0.0475)	(0.0360)	(0.0538)	(0.0327)	(0.0205)	(0.0310)
% home language Spanish school zip neighborhood	0.0318	-0.0667*	-0.0306	-0.120**	-0.0530**	0.115**
	(0.0430)	(0.0286)	(0.0342)	(0.0308)	(0.0159)	(0.0211)
above CST cutpoint	-0.00506	0.102**	0.0608	-0.0620*	0.0703**	0.111**
	(0.0436)	(0.0385)	(0.0577)	(0.0311)	(0.0212)	(0.0321)
% home language Spanish home zip neighborhood * above	-0.0704	-0.0979*	-0.110	-0.0144	-0.0929**	-0.156**
	(0.0563)	(0.0489)	(0.0724)	(0.0386)	(0.0274)	(0.0408)
% home language Spanish home zip neighborhood	0.0373	-0.0710	0.0154	-0.131**	-0.0827**	0.158**
	(0.0510)	(0.0382)	(0.0470)	(0.0364)	(0.0211)	(0.0277)
above CST cutpoint	-0.00881	0.130**	0.0560	-0.0765	0.0505	0.0283
	(0.0588)	(0.0477)	(0.0915)	(0.0423)	(0.0259)	(0.0489)
% foreign-born in school zip neighborhood * above	-0.0914	-0.211*	-0.154	0.00913	-0.0913	-0.0341
	(0.118)	(0.0950)	(0.199)	(0.0849)	(0.0528)	(0.105)
% foreign-born in school zip neighborhood	-0.0434	-0.102	0.100	-0.0677	0.0372	0.115
	(0.107)	(0.0705)	(0.118)	(0.0801)	(0.0411)	(0.0726)
above CST cutpoint	0.00458	0.121	-0.0347	-0.0877	0.0542	0.0260
	(0.0750)	(0.0676)	(0.104)	(0.0522)	(0.0358)	(0.0582)
% foreign-born in home zip neighborhood * above	-0.124	-0.187	0.0602	0.0347	-0.0988	-0.0277
	(0.160)	(0.145)	(0.223)	(0.110)	(0.0771)	(0.125)
% foreign-born in home zip neighborhood	-0.0191	-0.121	0.171	-0.0597	0.0329	0.0778
	(0.144)	(0.106)	(0.137)	(0.104)	(0.0598)	(0.0878)
above CST cutpoint	-0.0339	0.0970*	0.0240	-0.0952**	0.0408*	0.0671*
	(0.0457)	(0.0385)	(0.0583)	(0.0323)	(0.0208)	(0.0321)
% in poverty in school zip neighborhood * above	-0.0319	-0.121	-0.0694	0.0443	-0.0593	-0.109*
	(0.0729)	(0.0617)	(0.0963)	(0.0509)	(0.0339)	(0.0520)
% in poverty in school zip neighborhood	-0.0744	-0.0844	-0.0241	-0.161**	-0.0539*	0.0623
	(0.0654)	(0.0485)	(0.0593)	(0.0480)	(0.0265)	(0.0351)

above CST cutpoint	-0.0483	0.137**	0.0377	-0.0905*	0.0462	0.105**
	(0.0531)	(0.0467)	(0.0736)	(0.0372)	(0.0255)	(0.0398)
% in poverty in home zip neighborhood*above	-0.00380	-0.199*	-0.0942	0.0357	-0.0691	-0.182**
	(0.0915)	(0.0814)	(0.130)	(0.0621)	(0.0442)	(0.0693)
% in poverty in home zip neighborhood	-0.140	-0.0758	-0.00197	-0.182**	-0.105**	0.131**
	(0.0821)	(0.0631)	(0.0796)	(0.0587)	(0.0343)	(0.0470)

SOURCES: Authors' estimates.

NOTES: Each set of rows represents six different regressions. Coefficients and standard errors are reported for being just above the CST cutpoint, the moderator, and moderator * above. Other control variables used in the regressions reported in Table 5A are included for each regression. Regressions are Intent to treat, second order polynomial. Coefficients shaded green are positive and significant while those shaded red are negative and significant.

TABLE 8B

Estimates of the effect of moderator variables on CST ELA one year after reclassification, CST as running variable, each trio of coefficients is a uniquely estimated model, SDUSD, 2004-14

	2004-05			2007-14		
	Elementary	Middle School	High School	Elementary	Middle School	High School
above CST cutpoint	0.0344	-0.0496	-0.101	-0.0103	-0.00848	-0.0953
	(0.0698)	(0.0558)	(0.0942)	(0.0441)	(0.0519)	(0.0895)
Spanish * above	-0.0443	0.0197	-0.131**	0.00212	-0.165**	-0.259**
	(0.0569)	(0.0439)	(0.0489)	(0.0329)	(0.0394)	(0.0623)
Spanish Home-Language	-0.0763	-0.137**	-0.0472	-0.136**	-0.0120	0.0492
	(0.0516)	(0.0340)	(0.0789)	(0.0302)	(0.0427)	(0.0722)
above CST cutpoint	-0.000210	-0.0178	-0.205**	-0.0467	-0.0345	-0.0323
	(0.0717)	(0.0504)	(0.0766)	(0.0467)	(0.0464)	(0.0777)
% EL * above	-0.000112	-0.000571	0.00283	0.000878	0.000597	-0.00114
	(0.00122)	(0.000984)	(0.00185)	(0.000720)	(0.000915)	(0.00222)
% EL	-0.00133	0.000397	0.000290	-0.00151*	-0.000425	0.00126
	(0.00110)	(0.000744)	(0.00114)	(0.000654)	(0.000824)	(0.00189)
above CST cutpoint	-0.0159	-0.0525	-0.159	-0.0398	-0.0318	-0.0607
	(0.0627)	(0.0493)	(0.0831)	(0.0414)	(0.0468)	(0.0807)
same language as all students*above	0.000253	0.000418	0.000529	0.000695	0.000324	0.000126
	(0.000734)	(0.000647)	(0.00129)	(0.000471)	(0.000565)	(0.00116)
same language as all student	-0.000179	0.000765	0.000584	-0.00147**	-0.000856	-5.63e-05
	(0.000699)	(0.000554)	(0.000892)	(0.000451)	(0.000527)	(0.00106)
above CST cutpoint	0.0171	-0.0566	-0.111	-0.0199	-0.0214	-0.165
	(0.0720)	(0.0577)	(0.0993)	(0.0460)	(0.0545)	(0.0918)
same language as EL students*above	-0.000335	0.000231	-0.000495	0.000222	5.91e-05	0.00150
	(0.000687)	(0.000529)	(0.00102)	(0.000424)	(0.000518)	(0.000874)
same language as EL students	0.000728	0.000678	0.00213	-0.000986*	-0.000746	-0.000938
	(0.000751)	(0.000773)	(0.00111)	(0.000474)	(0.000614)	(0.00102)
above CST cutpoint	-0.0536	-0.0732	-0.136	-0.0579	-0.0249	-0.195
	(0.0864)	(0.0798)	(0.127)	(0.0527)	(0.0761)	(0.116)
language homogeneity of ELs * above	0.0607	0.0433	0.00684	0.0667	-0.0207	0.197

	(0.0938)	(0.0879)	(0.152)	(0.0583)	(0.0901)	(0.129)
language homogeneity of ELs	-0.100	-0.0820	0.152	-0.129*	-0.107	-0.214
	(0.0859)	(0.0756)	(0.0885)	(0.0545)	(0.0817)	(0.111)
above CST cutpoint	0.0494	0.0219	-0.107	0.0375	0.132*	-0.309*
	(0.0830)	(0.0771)	(0.141)	(0.0765)	(0.0650)	(0.155)
% teachers have MA * above	-0.128	-0.150	-0.0327	-0.0830	-0.259**	0.419
	(0.147)	(0.142)	(0.248)	(0.110)	(0.0979)	(0.240)
% teachers have MA	0.105	0.0582	-0.255	0.177	0.126	-0.458*
	(0.131)	(0.113)	(0.156)	(0.0986)	(0.0892)	(0.196)
above CST cutpoint	0.0487	0.00380	-0.0595	0.0904	0.133*	-0.139
	(0.0844)	(0.0874)	(0.138)	(0.0614)	(0.0608)	(0.119)
median years teaching experience * above	-0.00515	-0.00429	-0.00540	-0.00721*	-0.0119**	0.00667
	(0.00610)	(0.00671)	(0.00959)	(0.00352)	(0.00396)	(0.00754)
median years teaching experience	0.00201	-0.000415	-0.0106	0.00708*	0.0106**	-0.0104
	(0.00543)	(0.00540)	(0.00579)	(0.00316)	(0.00354)	(0.00649)
above CST cutpoint					0.0183	-0.0384
					(0.0524)	(0.0724)
% no ELD at school * above					-0.000461	-0.00103
					(0.000684)	(0.00138)
% no ELD at school					0.000333	0.00133
					(0.000628)	(0.00122)
above CST cutpoint				-0.0119	-0.0399	-0.0110
				(0.0527)	(0.0461)	(0.0740)
% home language Spanish school zip neighborhood * above				0.0767	0.111	-0.201
				(0.0869)	(0.0785)	(0.163)
% home language Spanish school zip neighborhood				-0.176*	-0.149*	0.0156
				(0.0790)	(0.0706)	(0.135)
above CST cutpoint				-0.0198	-0.0782	-0.0103
				(0.0573)	(0.0548)	(0.0965)
% home language Spanish home zip neighborhood * above				0.0965	0.209	-0.115
				(0.105)	(0.109)	(0.211)

% home language Spanish home zip neighborhood	-0.197*	-0.269**	-0.00853
	(0.0955)	(0.0975)	(0.180)
above CST cutpoint	0.0366	-0.0229	0.124
	(0.0705)	(0.0570)	(0.0955)
% foreign-born in school zip neighborhood * above	-0.0774	0.0431	-0.746*
	(0.178)	(0.153)	(0.298)
% foreign-born in school zip neighborhood	-0.100	-0.107	0.355
	(0.162)	(0.138)	(0.253)
above CST cutpoint	0.0392	-0.0722	0.205
	(0.0859)	(0.0833)	(0.154)
% foreign-born in home zip neighborhood * above	-0.0815	0.207	-0.871
	(0.233)	(0.248)	(0.482)
% foreign-born in home zip neighborhood	-0.0858	-0.322	0.159
	(0.213)	(0.221)	(0.416)
above CST cutpoint	-0.0250	-0.0618	0.0383
	(0.0615)	(0.0532)	(0.0915)
% in poverty in school zip neighborhood * above	0.0906	0.140	-0.259
	(0.102)	(0.0935)	(0.183)
% in poverty in school zip neighborhood	-0.203*	-0.164	0.0400
	(0.0924)	(0.0841)	(0.154)
above CST cutpoint	-0.0510	-0.0983	-0.0112
	(0.0669)	(0.0667)	(0.119)
% in poverty in home zip neighborhood*above	0.150	0.215	-0.0899
	(0.119)	(0.126)	(0.237)
% in poverty in home zip neighborhood	-0.252*	-0.291*	-0.0335
	(0.108)	(0.113)	(0.203)

SOURCES: Authors' estimates.
NOTES: Each set of rows represents six different regressions. Coefficients and standard errors are reported for being just above the CST cutpoint, the moderator, and moderator * above. Other control variables used in the regressions reported in Table 5A are included for each regression. Regressions are Intent to treat, second order polynomial. Coefficients shaded green are positive and significant while those shaded red are negative and significant.

TABLE 9

Estimates of the effect of moderator variables on CST ELA one year after reclassification, CST as running variable, moderators estimated simultaneously, LAUSD, 2003-12

	CELDT Era 1 ('03-'06)			CELDT Era 2 ('07-12)		
	Elementary	Middle	High School	Elementary	Middle	High School
above	-0.0262	0.106	0.232	-0.180**	0.128**	0.175*
	(0.0789)	(0.0861)	(0.151)	(0.0587)	(0.0465)	(0.0690)
Spanish * above	0.0206	-0.0421	-0.0840	0.0683	-0.107**	-0.181**
	(0.0554)	(0.0422)	(0.0563)	(0.0436)	(0.0276)	(0.0515)
Spanish Home-Language	-0.254**	-0.158**	-0.170**	-0.299**	-0.159**	-0.112**
	(0.0521)	(0.0356)	(0.0399)	(0.0423)	(0.0244)	(0.0383)
% no ELD at school	-0.114	-0.0815	0.0610	0.273*	-0.0766**	0.0184
	(0.227)	(0.0626)	(0.109)	(0.139)	(0.0267)	(0.0785)
% no ELD at school * above	0.600*	0.0152	-0.131	0.0294	0.107**	-0.285**
	(0.252)	(0.0834)	(0.179)	(0.150)	(0.0348)	(0.0620)
% EL	0.0477	-0.0648	-0.244**	0.0206	0.0596	0.0843
	(0.0718)	(0.0755)	(0.0916)	(0.0534)	(0.0464)	(0.0868)
% EL * above	-0.0446	0.0539	0.205	0.105	0.0678	-0.0847
	(0.0794)	(0.0994)	(0.150)	(0.0565)	(0.0585)	(0.0670)
% teachers have MA	-0.00738	0.288*	0.337*	-0.111	-0.0492	0.0114
	(0.103)	(0.121)	(0.161)	(0.0703)	(0.0510)	(0.0896)
% teachers have MA*above	0.0275	0.0383	-0.225	0.0868	-0.0631	0.233**
	(0.115)	(0.156)	(0.265)	(0.0745)	(0.0600)	(0.0346)
% home language Spanish home zip neighborhood	0.0250	-0.0312	0.127	-0.128**	-0.119**	-0.151**
	(0.0558)	(0.0498)	(0.0650)	(0.0396)	(0.0249)	(0.0494)
% home language Spanish home zip neighborhood * above	-0.0549	-0.105	-0.169	-0.0573	-0.0927**	0.00701**
	(0.0624)	(0.0657)	(0.101)	(0.0420)	(0.0321)	(0.00145)
Constant	-2.283**	-3.048**	-2.918**	128.2**	128.7**	70.87**
	(0.131)	(0.147)	(0.253)	(3.007)	(4.141)	(6.218)

Observations	15,127	11,482	5,186	48,325	40,320	20,416
R-squared	0.343	0.277	0.239	0.287	0.246	0.168

Robust standard errors in parentheses

** p<0.01, * p<0.05

SOURCES: Authors' estimates

NOTES: Each column is a separate regression. Other control variables used in the regressions reported in Table 5A are included for each regression. Regressions are Intent to treat, second order polynomial. Coefficients shaded green are positive and significant while those shaded red are negative and significant.

Table 10A

Years in US Schools by Grade, Current/Reclassified English-Learners (LAUSD)

Student Grade	0-3 Years in US Schools, ELs		4-5 Years in US Schools, ELs		6+ Years in US Schools, ELs		Reclassified		Total
	#	%	#	%	#	%	#	%	
PRE-K	16,491	99.79%	9	0.05%	3	0.02%	22	0.13%	16,525
K	174,689	99.54%	141	0.08%	63	0.04%	598	0.34%	175,491
g1	258,803	98.54%	2,829	1.08%	157	0.06%	847	0.32%	262,636
g2	204,012	75.89%	47,908	17.82%	572	0.21%	16345	6.08%	268,837
g3	20,205	7.34%	211,816	76.98%	5,809	2.11%	37,325	13.57%	275,155
g4	14,182	5.27%	142,783	53.02%	51,747	19.21%	60,600	22.50%	269,312
g5	12,811	4.85%	13,703	5.18%	144,107	54.52%	93,688	35.45%	264,309
g6	12,524	4.91%	8,398	3.29%	110,483	43.32%	123,639	48.48%	255,044
g7	12,544	4.77%	7,533	2.86%	87,693	33.34%	155,253	59.03%	263,023
g8	12,741	4.77%	7,138	2.67%	76,025	28.44%	171,433	64.13%	267,337
g9	23,277	7.47%	8,709	2.80%	88,177	28.31%	191,352	61.43%	311,515
g10	16,676	6.61%	7,162	2.84%	57,316	22.71%	171,185	67.84%	252,339
g11	10,730	5.07%	6,953	3.29%	39,347	18.60%	154,514	73.04%	211,544
g12	3,350	2.00%	8,042	4.79%	28,651	17.08%	127,737	76.13%	167,780

SOURCE: Authors' calculations.

Table 10B

Years Since First Enrolling in SDUSD by Grade, Current/Reclassified English-Learners (SDUSD)

Student Grade	0-3 Years in US Schools, ELs		4-5 Years in US Schools, ELs		6+ Years in US Schools, ELs		Reclassified		Total
	#	%	#	%	#	%	#	%	
K	30,594	94.58	1,429	4.42	197	0.61	127	0.39	32,347
g1	27,923	83.28	5,173	15.43	301	0.90	134	0.40	33,531
g2	24,726	76.08	6,489	19.97	1,138	3.50	147	0.45	32,500
g3	4,443	14.02	19,188	60.57	2,855	9.01	5,194	16.40	31,680
g4	3,626	11.82	14,205	46.32	4,599	15.00	8,239	26.86	30,669
g5	2,942	9.76	1,545	5.12	12,302	40.79	13,369	44.33	30,158
g6	2,521	8.60	1,322	4.51	10,608	36.21	14,847	50.68	29,298
g7	2,309	8.05	1,147	4.00	9,081	31.65	16,158	56.31	28,695
g8	2,152	7.71	1,010	3.62	7,960	28.53	16,777	60.13	27,899
g9	3,138	9.81	1,149	3.59	10,487	32.78	17,215	53.82	31,989
g10	2,641	9.31	1,145	4.03	8,366	29.48	16,227	57.18	28,379
g11	1,815	7.96	910	3.99	5,825	25.53	14,265	62.52	22,815
g12	860	3.79	1,043	4.59	6,776	29.84	14,026	61.77	22,705

SOURCE: Authors' calculations

Notes: The results use students who are currently or were formerly ELs in the school years 2007-08 through 2013-14. Pre-K observations were excluded in calculating years in district.

Table 10B differs in a subtle way from Table 10A, because the year of first enrollment in a US school was not available for SDUSD. Instead, we measure in Table 10B the number of years elapsed since first enrollment in SDUSD.

The next two tables show the characteristics of these two related measures (years in US schools for LAUSD, and years since first enrolled in SDUSD for SDUSD), for the sample of ELs who appear in any of our RD experiments.

TABLE 11A

Distribution of years in US schools upon reclassification, students included in at least one RD Experiment, LAUSD

Percentiles			
1%	2	Obs	180,392
5%	4	Sum of Wgt.	180,392
10%	4		
25%	4	Mean	6.0858
50% (Median)	6	Std. Dev.	2.2471
75%	7		
90%	9	Variance	5.0492
95%	11	Skewness	1.1143
99%	13	Kurtosis	4.4758

SOURCE: Authors' calculations

NOTES: The sample is based on an RD sample that uses a CST bandwidth of 50.

TABLE 11B

Distribution of years since first enrolled in SDUSD upon reclassification, students included in at least one RD Experiment, SDUSD

Percentiles			
1%	1	Obs	48,617
5%	1	Sum of Wgt.	48.617
10%	2		
25%	3	Mean	4.5564
50% (Median)	4	Std. Dev.	2.3623
75%	6		
90%	8	Variance	5.5806
95%	9	Skewness	0.6925
99%	11	Kurtosis	3.3088

SOURCE: Authors' calculations

NOTES: The sample is based on an RD sample that uses a CST bandwidth of 50.

The main report discusses the correlation between the CST and various CELDT test scores, and states that the scores are not strongly correlated, meaning that it is appropriate to consider an RD experiment for one test for the subsample of students who meet the reclassification criteria on the other test. Table 12A examines this question in a slightly different way. It reports the correlation between dummy variables for whether a student was included in a given experiment and another RD experiment. In LAUSD these indicators are positively related and strongly so in the case of the CELDT OPL and CELDT reading experiments. The table shows an average for the two CELDT eras. In SDUSD the analysis is simpler because it was only in era two, and between the CST and CELDT reading experiments, that we conducted two RD experiments together. In that case, the correlation between dummy variables indicating inclusion in the given experiment was 0.45.

Table 12A

Correlation Between Inclusion in the Three Experiments (LAUSD)

Experiment	lagcstela	celdtOPL	celdtREAD
lagcstela			
celdtOPL	0.7448		
celdtREAD	0.7974	0.9113	

SOURCE: Authors' calculations combining both CELDT eras.

Tables 13A and 13B show for LAUSD and SDUSD the percentage of all ELs who did and did not meet the overall (OPL) CELDT requirement for reclassification, which in both districts was at the Early Advanced level, versus the percentage who met and did not meet the districts' subtest requirements. The most dramatic difference between the districts is that in SDUSD virtually no students who failed to meet the OPL requirement met the subtest requirement, and this reflects the very high requirement in SDUSD, where students must score Early Advanced on all but one subtest, with one allowed at the Intermediate level.

Table 13A

Rate of Students Meeting CELDT Subtest Requirements but NOT Overall Requirement in LAUSD (All EL Students)

	Subtest Not Met		Met Subtest	
Overall Not Met	749,783	77.36%	219,412	22.64%
Met Overall	67,913	10.27%	593,343	89.73%

Table 13B

Rate of Students Meeting CELDT Subtest Requirements but NOT Overall Requirement in SDUSD (All EL Students)

	Subtest Not Met		Met Subtest	
Overall Not Met	208,187	96.8%	6,789	3.2%
Met Overall	38,742	37.8%	63,833	62.2%

Tables 14A and B show the same cross-tabulation, but this time limiting the sample to those ELs who in a given year were included in one of the RD experiments. In LAUSD the percentage of those not meeting the overall CELDT proficiency level who met the CELDT subtest requirements was quite high relative to the same cell in Table 13A, which is for the entire EL population. In Table 14A, each included EL was included in at least one of the RD experiments, so they must have also met the CST and CELDT OPL requirements. The table for SDUSD has no entries in the "Overall Not Met" row because all three of the RD experiments we used (CST in eras 1 and 2 and CELDT Reading in era 2) required students to have met the other reclassification requirements, including the CELDT Overall Proficiency Level requirement.

Additional insights emerge by comparing Tables 13 and 14, because the earlier tables include all ELs and the later tables include only EL students used in an experiment. A comparison of the numbers shows that in LAUSD about one quarter of all EL observations were included in the RD experiments, compared to about one tenth in SDUSD. The smaller proportion of ELs used in San Diego mostly reflects the much higher CELDT subtest requirements in SDUSD compared to LAUSD. Recall that in SDUSD, only one CELDT subtest may be at the Intermediate level, and all other subtests must be at the Early Advanced level. In LAUSD, all subtests may be at the Intermediate level.

Table 14A

Rate of LAUSD Students Meeting CELDT Subtest Requirements but NOT Overall Requirement (Students In At Least One Experiment)

	Subtest Not Met		Met Subtest	
Overall Not Met	12,437	24.22%	38,910	75.78%
Met Overall	7,701	2.47%	304,125	97.53%

SOURCE: Authors' calculations.

Table 14B

Rate of SDUSD Students Meeting CELDT Subtest Requirements but NOT Overall Requirement (Students In At Least One Experiment)

	Subtest Not Met		Met Subtest	
Overall Not Met	0	N/A	0	N/A
Met Overall	3,742	12.0%	27,420	88.0%

SOURCE: Authors' calculations.

NOTES: N/A – Not Applicable.

Finally, Table 15 shows what happens to the ITT and TOT estimates of the impact of reclassification on graduating on time in San Diego when students who moved to other districts before graduation are treated as not graduating on time. This exercise is useful because in LAUSD the data did not distinguish between dropouts and those leaving for other districts, meaning than in the LAUSD models of graduating on time leavers were treated as not graduating on time. For further discussion of this table see Appendix A.

TABLE 15

A Comparison of the SDUSD CST Regression Discontinuity Results for Era 2 (2007-2014) for the Outcome Graduation on Time, Treating District Leavers as Not Having Graduated on Time Versus the Original Results Which Exclude District Leavers Who Did Not Drop Out

	Treating Leavers as Non-Graduates		Original Results	
Grade Span	Intent to Treat	Treatment on Treated	Intent to Treat	Treatment on Treated
3-5	0.0309 (0.0571) N=2957	0.069 (0.127)	0.0289 (0.0550) N=1814	0.0721 (0.137)
6-8	-0.0464 (0.0372) N=6216	-0.0670 (0.0539)	-0.00984 (0.0393) N=4271	-0.0138 (0.0552)
9-12	0.00842 (0.0304) N=6376	0.0165 (0.0594)	0.00961 (0.0309) N=5715	0.0189 (0.0607)

SOURCE: Authors' calculations.

NOTES: The results use bandwidth of 50, with a second order polynomial.

PPIC PUBLIC POLICY INSTITUTE OF CALIFORNIA

25 YEARS

The Public Policy Institute of California is dedicated to informing and improving public policy in California through independent, objective, nonpartisan research.

Public Policy Institute of California
500 Washington Street, Suite 600
San Francisco, CA 94111
T: 415.291.4400
F: 415.291.4401
PPIC.ORG

PPIC Sacramento Center
Senator Office Building
1121 L Street, Suite 801
Sacramento, CA 95814
T: 916.440.1120
F: 916.440.1121