Research Article

Educational Evaluation and Policy Analysis
June 2023, Vol. 45, No. 2, pp. 311–335
DOI: 10.3102/01623737221121786
Article reuse guidelines: sagepub.com/journals-permissions

Dual Identification? The Effects of English Learner (EL) Status on Subsequent Special Education (SPED) Placement in an Equity-Focused District

Mark Murphy
University of Hawai'i at Mānoa

Angela Johnson
NWEA

This study examines the effects of English Learner (EL) status on subsequent Special Education (SPED) placement. Through a research-practice partnership, we link student demographic data and initial English proficiency assessment data across seven cohorts of test takers and observe EL and SPED programmatic participation for these students over 7 years. Our regression discontinuity (RD) estimates at the English proficiency margin consistently differ substantively from positive associations generated through regression analyses. RD evidence indicates that EL status had no effect on SPED placement at the English proficiency threshold. Grade-by-grade and subgroup RD analyses at this margin suggest that ELs were modestly underidentified for SPED during Grade 5 and that ELs whose primary language was Spanish were underidentified for SPED.

Keywords: *English Learner, Special Education, educational policy, equity, evaluation, bilingual/bicultural, quasi-experimental analysis, regression discontinuity, descriptive analysis*

Introduction

The effectiveness of educational support services holds tremendous influence on the academic success of millions of public school students in the United States (Thurlow et al., 2006). When such services meet students' needs for linguistic or disability support, indicators of student success can improve (Berkeley et al., 2010; Morgan et al., 2010; Schwartz et al., 2021; Wang & Lam, 2017). Conversely, when needs are unmet, student success can falter (Morgan et al., 2010). Persistent achievement gaps between student populations who are eligible for services and those who are not suggest that there is substantial room for improvement nationwide (Albus & Thurlow, 2013; Pasternack, 2014; Schwartz et al., 2021). Furthermore, some students rely not solely on English Learner (EL) services or Special Education (SPED) services, but rather on supports for *both* a disability and developing English proficiency. We refer to these students as *dually identified* students (i.e., identified for both EL and SPED), and they are a central focus of this study (Carnock & Silva, 2019; Umansky et al., 2017).

Despite a substantial body of descriptive and qualitative research on the relation between EL and SPED placement, little is known about the causal link between the two. Prior work has highlighted disproportionate representation of EL students in SPED using descriptive regression and hazard analyses (Morgan et al., 2015, 2017, 2018; Umansky et al., 2017). This work has found grade-level heterogeneity with EL students' being underidentified for SPED in elementary grades, but overidentified for SPED in secondary grades (Umansky et al., 2017). However, a clear evaluation of how EL status *affects* subsequent SPED placement has not yet been conducted. Previous studies on this topic noted that to thoroughly examine disproportionate representation, student-level data are necessary (Morgan et al., 2017).

Through a research partnership with a large, urban school district in California, we employ student-level data across seven cohorts of kindergarten students to examine the relation between EL status and SPED placement. We principally leverage a regression discontinuity (RD) design to generate credibly causal estimates of the effects of EL status on subsequent SPED placement for students near the English proficiency margin. However, we also examine descriptive results using raw placement rates and regression analysis that allow us to understand the relation between EL status and SPED placement across the broader English language proficiency continuum. Results obtained through these different empirical approaches strengthen our understanding of how SPED placement is related to the range of possible levels of English language proficiency. For example, the probability of being in SPED is higher for students with lower levels of English proficiency. This paper makes at least two critical contributions.

First, our application of the RD research design to detailed, student-level data adds new evidence based on a rigorous quasi-experimental approach about the placement of EL students in SPED services. The RD design allows for the unbiased estimation of the Local Average Treatment Effect (LATE) by comparing individuals just above and below an arbitrary cut point or threshold that determines assignment to a particular intervention (Angrist & Pischke, 2009; Thistlewaite & Campbell, 1960). A core strength of this empirical approach is the modest number of assumptions that must hold in order for inferences to be considered valid (Angrist & Pischke, 2009). Furthermore, compared with other quasi-experimental research designs (e.g., propensity score matching; interrupted time series designs, etc.), the RD approach stands out because the identifying assumptions can be empirically tested (Calonico et al., 2014; McCrary, 2008).

Our main RD findings indicate that EL status largely led to proportionate representation of EL students in SPED at the English proficiency margin. These results, based on the pooled sample, suggest that the district identified a close-to-proportionate number of EL students for SPED placement at the English proficiency threshold during the years of our study.

However, a limitation of the RD approach is that our estimand (i.e., the LATE) is local and may not speak to patterns occurring across the full continuum of English language proficiency. Given this, we also explore descriptive results to provide a more complete picture of the complex relation between EL status and SPED placement. Our unconditional placement rate comparisons and ordinary least squares (OLS) regression analyses indicate that for the full analytical sample, EL students in this district are overrepresented in SPED by at least 2.3 percentage points; however, it is possible that the effect of EL status on SPED placement is heterogeneous at varying levels of English proficiency. Furthermore, across a wide range of bandwidths, we reject the null hypothesis that our RD estimate is statistically the same as OLS estimates. This highlights the novel information generated by the RD approach.

Second, we present highly policy-relevant evidence for practitioners that can facilitate the continuous improvement of SPED identification procedures for ELs. Initially, we explore how the *timing* of SPED placement may reflect disproportionate representation. Prior literature has found that disproportionate identification of students who entered kindergarten as ELs for SPED identification can vary by grade level (e.g., Umansky et al., 2017). We look specifically at RD results for SPED placement in each year following initial EL classification in kindergarten, from Grade 1 to Grade 6, and find evidence that slight underidentification of EL students for SPED placement occurred during Grade 5. These findings provide insight to district staff on the implementation of SPED placement procedures that may differ by grade and the implications of these practices. In addition, we explore effect heterogeneity by students' primary language to shed light on possible differences that occurred across language groups. To do this, we split our sample into three subgroups (students speaking Spanish as the primary language, students speaking Mandarin/Cantonese, and students speaking any other non-English language) and apply our RD approach to each subgroup separately. After doing this, we observe suggestive evidence of heterogeneity by primary language group. Spanish-speaking ELs were underidentified for SPED while Mandarin- or Cantonese-speaking ELs and ELs speaking other non-English languages were proportionately identified for SPED. Our findings help to illuminate potential differences in the way EL students were placed into SPED based on the primary language spoken. Such distilled information can aid district leaders to develop more targeted approaches for students in each of the three distinct subgroups.

In combination, these contributions build on previous literature and strengthen the available evidence pertaining to the intersection of EL status and SPED placement. While our main results do not raise clear concerns for inequitable SPED placement practices overall based on EL status in this district, subgroup analyses suggest areas for adjusting the SPED identification process for ELs.

Students in EL and SPED Services

English Learners

ELs are students between the ages of 3 and 21 years who need additional support to improve their English language listening, speaking, reading, and/or writing abilities to be able to succeed in academic courses where English is the language of instruction (U.S. Department of Education [DOE], 2016). Also referred to as students with limited English proficiency or as emergent bilingual/multilingual students, EL students have been a protected class of students since the 1974 Supreme Court decision in *Lau v. Nichols*. Title III of the Every Student Succeeds Act (2015) provides state educational agencies with substantial latitude in how EL students are to be identified, but most commonly (including in our partner district) the process involves a home language survey for newly enrolled students. Students whose families indicate that a language other than English is spoken at home are then given a formal assessment to determine if the student qualifies for EL classification. In California districts, including our partner district, the California English Language Development Test (CELDT) was the formal assessment used to evaluate a student's English language skills from 2001 to 2017. Approximately 10% of the total U.S. student population was classified as EL in 2015 (U.S. DOE, 2017). At the time, more than three quarters of the "current EL" population, or students with active EL classification, identified as Latino/a, yet the current EL population overall was extremely diverse with regard to race, ethnicity, nationality and languages spoken (U.S. DOE, 2017). The 10 districts enrolling the highest proportions of current ELs were located in California, Alaska, New Jersey, Arizona and Washington (U.S. DOE, 2017).[1] Of the current EL population, approximately 14% qualified for SPED (U.S. DOE, 2017).

Students who classify as EL (i.e., not English-proficient) are entitled to educational services for English language development (ELD) (U.S. DOE, 2016). Services for EL students vary across schools, districts, and states. Some ELs are enrolled in bilingual or dual language programs, in which their primary language is used in addition to English for content instruction; others are enrolled in courses taught only in English but have designated ELD time in their weekly schedules. In our partner district, dual language programs that provide home language instruction from kindergarten to Grade 8 are offered in Spanish, Cantonese, Mandarin, and Korean. Additional shorter-term bilingual and biliteracy programs in these and other languages are also available in select schools. Students who are unable to or choose not to attend a dual language program receive designated and integrated ELD, which aims to support linguistic progress as well as academic core content skills. Designated and integrated ELD take the form of courses focused on developing communicative and academic English proficiency or content-based courses with language development embedded.

Students in Special Education

Students in SPED receive services to enable them to access a free appropriate public education (FAPE). Since 1975, federal law and related judicial rulings have required that all children nationwide have access to FAPE. This means that any student with needs due to a disability are to receive individually tailored supports and accommodations. The Individuals with Disabilities in Education Act (2004) part B covers students aged 3 to 21 and requires schools to provide services in the least restrictive environment (Carnock & Silva, 2019). As a result, schools must provide necessary support while also ensuring the students are not unnecessarily diverted from typical educational settings. IDEA defines 13 distinct disability categories.[2] When a student is identified as having a disability in any of these 13 categories, the student is entitled to SPED services and an Individualized Education Program (IEP) is established.

In compliance with federal and state policies, our partner district follows a multistep process to identify student eligibility for SPED services. For students who are identified as having difficulty with academics, health, behavior, or attendance, schools may first suggest a variety of interventions such as a Student Success Team. If evidence suggests the interventions are insufficient for addressing a student's needs, family, school staff, and other adults in the community can initiate a referral for an evaluation to find out if the student has a disability. The district guards against inappropriate basis for referral by requiring data-based documentation that the challenges in the student's educational progress are not primarily due to lack of appropriate instruction in reading or math, English language acquisition, cultural, environmental, economic, or socio-linguistic factors.

During an evaluation, existing information is reviewed and new assessments can be conducted to identify all of the student's needs for SPED and other services. IDEA specifies that tests and interviews given in the evaluation must be in the student's primary language and must not be conducted in a way that discriminates against any student based on a disability or on their cultural or racial background. In this district, assessments are available in English, Spanish, Mandarin, and Cantonese. Informed by the evaluation results, decisions regarding the student's eligibility for SPED services are made by a team that includes the students' family, school staff in general education and SPED, and sometimes other educators and/or caregivers. For eligible students, a meeting is held to develop an IEP, which states the services the district will provide.

In addition, the district conducts a meeting at least once a year with the student's family to review the student's progress toward goals in the current IEP and make appropriate changes. Under IDEA, students must be reevaluated at least every 3 years to see if they continue to be "with a disability" as defined by law and to update their educational needs and services.

Nationwide, since the 2007-08 school year, approximately 13% of all students have been placed in SPED after being identified as having a disability within one of the 13 categories specified in IDEA. This translates to more than 6 million students annually receiving SPED services under IDEA (U.S. DOE, 2013). Students identified with either a "specific learning disability" or "speech or language impairment" made up more than half of all disability classifications.[3]

Dually-Identified Students

Around 700,000 students nationwide in the 2014-15 school year were dually identified, or eligible for both EL and SPED services, when counted using a current-EL framework, which considers only ELs who have not yet attained fluent English proficiency or exited EL services. Furthermore, current-EL students were more likely to be identified for either the "specific learning disability" or "speech or language impairment" disability categories than non-EL students (U.S. DOE, 2017). However, these statistics mask substantial complexity in defining the presence of EL and dually identified students nationwide by failing to account for students who were ELs at one time but have since reclassified to English-proficient. In contrast, an "ever-EL" framework encompasses a broader set of students who are either current ELs or students who have exited EL services. A major strength of using the ever-EL framework is that the underlying sample remains consistent over time, retaining all students who ever are identified for EL status regardless of reclassification status (Umansky, 2016b). Using the ever-EL framework is especially important for studying dually identified students because it helps to compare SPED placement rates for students who were similar at baseline, or prior to any EL intervention (Umansky et al., 2017).

Dually identified students are an important student population, and their unique intersection of needs for educational services calls for greater study and evaluation (Carnock & Silva, 2019; Fagan & Pentón Herrera, 2022; Park et al., 2016). The provision of both EL and SPED services is required by federal law but implemented at the local level. Federal appropriations for SPED have historically only covered a limited portion of the actual costs to provide such services to districts (National Council on Disability, 2018). In recent years, federal appropriations for SPED have provided just over 15% of the actual cost districts experience when implementing these services. For EL services, real funding levels (i.e., adjusted for inflation) have recently dropped below the per pupil amount appropriated in 2002 (Carnock & Silva, 2019). The lack of sufficient funding for both EL and SPED programs makes concerns regarding the provision of services for dually identified students who rely on both programs even more stark.

Furthermore, extant qualitative research on bilingual students with needs for SPED services (also referred to as ELs with disabilities, emergent bilinguals with disabilities, or emergent bilinguals labeled as disabled [EBLAD]) highlights considerable limitations in their access to learning opportunities. In a review of the literature on normalcy, disability, and race, Cioè-Peña (2017a) pointed out that in relation to an ideology of normalcy, for which a command of the English language is a prerequisite, EBLAD students are perceived as doubly disabled. These inaccurate perceptions can have a major impact on the inclusion of EBLADs, often resulting in segregation and ostracization of students from their peers and/or limited access to multilingual learning environments (Cioè-Peña, 2017a, 2017b). Cioè-Peña (2020a) found that even in schools that offer bilingual or dual language education, EBLAD children tended to be placed in English-only programs because they are viewed as less likely to be able to attain, sustain, and benefit from a multilingual identity. A related study focusing on EBLAD teachers and school staff found that these program placement decisions are often made solely by the school without input from students and their families, and the disconnect between school and home leads to the families' having limited capacity to support their children's learning (Cioè-Peña, 2020b). In a review of literature focused on bilingual children with autism spectrum disorders, Park (2014) also highlighted empirical findings that challenged the notion that bilingual education is too difficult for autistic children. Park (2014) showed that developing dual-language abilities is essential for facilitating communication with parents, increasing social interactions, and identity formation.

In our partner district, IEPs for EL students must (a) state that the student is to be assessed annually using the English language proficiency assessment in California or an alternate (with modification and/or accommodation if appropriate) and the results considered; (b) include the student's current level of academic achievement and functional performance; (c) indicate whether general education, SPED, or both will provide ELD services and ensure teachers have appropriate primary language support, ELD, and SPED credentials; (d) create ELD goals and ensure all students receive at least 30 min of daily ELD appropriate for their proficiency level; (e) state the language of instruction for other goals; and (f) offer the family translated documents.

Associations Between EL Classification and SPED Placement

Existing research suggests that EL classification can affect later student placement in SPED (Burr, 2019; Burr et al., 2015; Hibel & Jasper, 2012; Umansky et al., 2017). To the extent disproportionate placement occurs, SPED placement may be a key moderator of the effect of EL classification on students' short- and long-term outcomes. Evidence on this topic is mostly associative in nature and documented using regression or hazard analyses. In this section, we discuss extant literature describing the challenging work of disentangling disabilities from developing language proficiency, the prior focus on disproportionate representation, how both underidentification and overidentification of EL students for SPED can be harmful, and recent methodological advances in the work exploring the effects of EL classification and reclassification.

Disentangling Disabilities From Developing Language Proficiency

Prior studies highlighted the challenge of differentiating disabilities from needs for ELD. Poorly designed language assessments with weak psychometric properties, for example, can create problems for discerning between language needs and disability needs (Macswan & Rolstad, 2006). In addition, an early study noted that a disproportionate number of Latino/a students were labeled as having a learning disability solely due to limited English proficiency (Ortiz & Polyzoi, 1986). More recent literature suggests that difficulty differentiating between a disability and language proficiency continues to challenge educational institutions and staff (Carnock & Silva, 2019; Park, 2019b). This can be especially true for students in the early grades and frequently results in diagnoses for a language need earlier than a disability (Burr, 2019; Carnock & Silva, 2019). Qualitative research found educators' beliefs about whether and when to identify EL students for SPED evaluation to cluster around two themes: "wait to be sure" and "the sooner the better" (Park, 2020). Policies pertaining to district SPED identification processes may be particularly relevant and important to consider relative to this phenomenon (Burr, 2019).

Disproportionate Identification

Multiple studies highlighted the issue of a potential disproportionality (i.e., either underrepresentation or overrepresentation) of EL student participation in SPED. Crucially, underidentification of EL students for SPED can be harmful for students academically if EL students with disabilities are not receiving necessary services (Greenberg Motamedi et al., 2016). Such a phenomenon could be occurring as a result of delayed testing for EL students (Samson & Lesaux, 2009). A delay may be stimulated by some form of explicit or implicit bias against EL students (Figueroa & Newsome, 2006). For example, the model-minority stereotype has been found to create biases that affect SPED identification for Asian-American students in different ways (Park, 2019a). One assumption identified by that study was that Asian American students were obedient and achieved academic success without need for support, so those who showed deviant behavior or struggled academically must have had a disability. An alternative explanation is that EL students may be more difficult to identify for SPED due to difficulty differentiating between language proficiency and a disability (Burr, 2019; Carnock & Silva, 2019). Regardless, underidentification merits concern because some students with needs for SPED lack access to essential support services.

On the other hand, overidentification of EL students for SPED can also be harmful to students (Burr, 2019; Burr et al., 2015) if inappropriate placement in SPED limits EL students' inclusion in general education classrooms (Samson & Lesaux, 2009). A key component of federal law establishing protections for both EL and SPED students dictates that students must be placed in the *least* restrictive educational environment (Carnock & Silva, 2019). Accordingly, placement in SPED without a need for the service can stymie a student's ability to participate in general education class settings, which may be essential for the student's growth and development. Furthermore, prior work has highlighted that SPED participation can lead to harmful stigmatization (Shifrer, 2013).

Proportionate representation at the margin, therefore, would suggest that services are being adequately provided with students still having access to less restrictive classroom settings. This represents the appropriate middle ground that districts are seeking to reach (Burr et al., 2015).

Umansky and colleagues (2017) used discrete-time hazard analyses and an ever-EL framework to examine the likelihood that a student subsequently participates in SPED (i.e., becomes dually identified). They found that ever-EL students were less present in SPED overall and within most disability categories (Umansky et al., 2017). However, an important limitation of this study was that a causal link was not identified. In other words, it is unclear whether participation in EL services *led* students to be underclassified in SPED. As such, outstanding questions about whether EL status *causes* disproportionate classification for SPED remain.

Other existing work on the intersection of EL classification and SPED classification also suggests that ELs tend to be both disproportionately identified for most disability categories and identified later than non-ELs for SPED services (Hibel & Jasper, 2012; Morgan et al., 2015; Samson & Lesaux, 2009). Notably, however, these three studies did not report findings for the subgroup of students that were initially classified for EL, but rather examined samples of students who either spoke another language at home or were identified as children of immigrants (i.e., students who *may or may not* have been eligible for EL services). Still other work has looked exclusively at subgroups of *current-EL* students to analyze disproportionality (e.g., Artiles et al., 2005; Sullivan, 2011; Sullivan & Bal, 2013; Wagner et al., 2005). However, as has been previously noted in the literature, a key consideration when reviewing these analyses is that they did not address reclassification. In other words, the results reported in these articles did not account for the way that the sample of EL students changed as students reclassify as English-proficient and exit EL status. Retaining reclassified students in the sample is appropriate as it enables a full and consistent comparison of students over time. Not accounting for these students may lead to inaccurate evaluations of SPED placement rates. A further limitation to some extant literature is the reliance on repeated cross-sectional data instead of panel data (e.g., Klingner et al., 2005; Morgan et al., 2015; Samson & Lesaux, 2009). Using only cross-sectional data inhibits the ability to precisely identify if observed relationships were due to policy interventions or changes in the underlying sample. In sum, a substantial amount of prior research has emphasized the importance of understanding the disproportionality of EL students in SPED. However, to date, the empirical methods applied to this topic, while consistently becoming more advanced, have remained unable to explore a causal link.

Our study, combining panel-based research designs, the ever-EL approach (which compares ever-ELs to never-ELs, both of whom entered kindergarten at the cusp of fluent English proficiency) and fine-grain student-level EL and SPED participation data, allows for a rigorous quantitative analysis of the causal relations between EL status and SPED placement near the English proficiency margin. In particular, using an RD design, we examine the probability of SPED participation between 1 and 6 years after being designated for EL status for students who scored similarly on the CELDT initial assessment during kindergarten.[4]

Regression Discontinuity Evidence on EL Classification and Reclassification

Up to this point, an RD design has not been used to study how EL status affects SPED placement. However, recent research on how EL status affects other educational outcomes illustrates the value of this approach. Several studies have leveraged student scores from the CELDT or other English proficiency assessments to employ RD designs that estimate the effect of EL status on academic achievement and attainment (Johnson, 2019; Shin, 2018; Umansky, 2016a, 2016b; Umansky et al., 2021). More specifically, Umansky et al. (2021) explored the effect of initial EL classification on Alaska Native student outcomes using an RD with a standardized screener assessment score as the forcing variable. They found negative impacts on academic outcomes for Alaska Native students in Grades 3 and 4. Johnson (2019) used a binding-score RD framework and found the effects of initial EL classification on outcomes such as high school graduation and college attendance to be limited. Shin (2018) found weak positive effects of initial EL classification on standardized test scores. Two other RD studies that used initial student CELDT scores as the forcing variable found that EL classification was in fact harmful to the likelihood of taking rigorous academic coursework and student achievement on standardized tests (Umansky, 2016a, 2016b).

In addition to these RD studies that examined the effects of *initial* EL classification, this research builds on an important body of work on EL reclassification. Robinson (2011) highlighted that assessment-based criteria are powerful levers for guiding reclassification decisions and demonstrated approaches to evaluating assessment score thresholds. This paper showed that when it comes to determining an appropriate threshold for transitioning ELs out of language services, rather than a positive effect, a null effect (which signals a smooth transition with no improvement or detriment to subsequent academic outcomes) is the desirable result. Following this work, several studies estimated the effects of shifting reclassification criteria and of reclassification decisions on students' downstream outcomes. For example, Robinson-Cimpian and Thompson (2016) found that using more stringent assessment-based criteria (thus making reclassification more difficult to attain) eliminated negative effects that reclassification had previously had on students' English language arts test scores and high school graduation rates. Carlson and Knowles (2016) estimated the effects of reclassifying at the end of 10th grade and found positive effects on students' ACT scores and some evidence for positive effects on high school graduation and postsecondary enrollment. These studies suggested that school districts (a) monitor student outcomes following the reclassification decision and make changes to the threshold if necessary and (b) examine their provision of academic opportunities and ensure that resources are available to all students regardless of their reclassification decision.

Our study advances the literature examining disproportionate representation of EL students in SPED by applying three key components shared by these studies on EL initial classification and reclassification: an RD design, student-level data, and a known cut score to determine EL eligibility.

Data

We partner with a large urban school district with a long history of serving a diverse EL population through an exceptionally rich variety of EL services. We leverage the district's longitudinal data from the California Longitudinal Pupil Achievement Data System (CALPADS). The data include four essential sets of information: (a) SPED placement, (b) EL classification, (c) the results (overall and by domain) that students obtained during their *initial* CELDT, and (d) students' demographic characteristics. Our study is based on panel data for SPED placement, EL classification, and demographic characteristics from School Year (SY) 2006-07 through SY 2018-19. Results from the initial CELDT are available from SY 2006-07 through 2016-17.[5]

To understand the effect of EL status on a consistent set of SPED outcomes, our sample focuses exclusively on students who took the initial CELDT during their kindergarten year.[6] The students in our sample, therefore, were those whose families reported speaking a language other than English at home and entered the district during their kindergarten year.[7] Critically, our analytical sample includes all students who took the CELDT, whether they were classified as EL or English-proficient.[8]

To gain clear insight into the SPED placement outcomes across the elementary school timespan, we follow seven cohorts of initial CELDT takers for 7 years (i.e., during the initial CELDT year and in the six subsequent years). Figure 1 provides an illustration of the cohorts included in our sample. For the main analyses, we keep cohorts of initial CELDT takers from SY 2006-07 through SY 2012-13 (N = 12,607).[9] For each of these cohorts, we observe SPED placement for students in each subsequent year.[10] Table 1 provides summary statistics.

Our principal outcome variable is an indicator that the student was first placed in SPED between Grades 1 and 6 after the initial CELDT in kindergarten. Students who were identified for SPED in the *same* year (kindergarten) as the initial CELDT are not flagged by this outcome. This is because we do not observe the precise start date (day and month) of SPED placement during kindergarten and cannot identify whether SPED placement occurred before or after the CELDT. Also, many students enter kindergarten having been flagged as needing SPED services through an early childhood education program (e.g., Head Start; Pre-kindergarten).[11] Therefore, our analysis focuses on those students identified for SPED *following* the initial CELDT so that we can directly understand the influence of EL status.[12]

In addition, our data include baseline demographic characteristics that make up our student-level controls. We include a flag for whether the student identifies as female. Furthermore, we have race/ethnicity-based flags for individuals that identify as (a) Hispanic, (b) Chinese, or (c) Decline to State Race/Ethnicity. Our data also

Kindergarten Cohort	*t=0*	*t=1*	*t=2*	*t=3*	*t=4*	*t=5*	*t=6*	*t=7*	*t=8*	*t=9*	*t=10*	*t=11*	*t=12*
2006-07	m	m	m	m	m	m	m	s	s	s	s	e	e
2007-08	m	m	m	m	m	m	m	s	s	s	s	e	
2008-09	m	m	m	m	m	m	m	s	s	s	s		
2009-10	m	m	m	m	m	m	m	s	s	s			
2010-11	m	m	m	m	m	m	m	s	s				
2011-12	m	m	m	m	m	m	m	s					
2012-13	m	m	m	m	m	m	m						
2013-14	s	s	s	s	s	s							
2014-15	s	s	s	s	s								
2015-16	s	s	s	s									
2016-17	e	e	e										
2017-18	e	e											
2018-19	e												

FIGURE 1. *CELDT cohort data.*

Note. Student-level data for individuals who took their initial California English Language Development Test (CELDT) assessment in kindergarten are from CALPADS for SY 2006-07 through SY 2018-19. Our main analysis includes cohorts who took the initial assessment from SY 2006-07 through SY 2012-13, observed annually during the assessment year and the six subsequent years. Our supplemental analyses include cohorts who took the initial assessment from SY 2006-07 through SY 2015-16 observed in 3 to 10 subsequent years. We exclude data from cohorts after SY 2015-16 cohorts and we exclude data after the 10th subsequent year for SY 2006-07 and SY 2007-08 cohorts. In the graphic above, "m" indicates that the data were used in the main analysis, "s" indicates that the data were used in supplemental analyses, and "e" indicates that the data were excluded.

include measures of the highest level of education received by the students' mother or father. We synthesize this information into a flag for whether the highest-educated parent had at least a high school diploma.[13] Our final baseline student characteristic approximates the student's age at the time of the initial assessment.[14]

Our data from the initial CELDT also include information about the primary language spoken by the student. More than 40 different languages were represented in the sample, with large groups of students speaking Spanish, Mandarin, or Cantonese. As shown in Table 1, approximately 32% of students taking the initial CELDT assessment indicated speaking Spanish. Another 36% of students spoke Mandarin or Cantonese, and about 33% of students indicated speaking another non-English language. Using this information, we explore effect heterogeneity by language group. This analysis was of interest because the district has SPED assessments available in English, Spanish, Mandarin, and Cantonese, but not in other languages. Students for whom home language assessment was not available were assessed in English. As a result, we consider the possibility of differential experiences and outcomes across languages.[15]

Methods

One approach to understanding the relation between EL status and SPED placement is to observe raw participation rates and differences between ELs and non-ELs. Another is to simply regress EL status and other covariates on our SPED outcome variables of interest to estimate an association between the two. Prior work on this topic largely relied on these types of descriptive

TABLE 1

Summary Statistics

Variable	*M*	*SD*	Minimum	Maximum
Student characteristics in year of initial assessment				
Female	0.50	0.50	0	1
Hispanic	0.35	0.48	0	1
Chinese	0.38	0.48	0	1
Decline to state race/ethnicity	0.04	0.20	0	1
Parent's highest education level ≥ high school diploma	0.49	0.50	0	1
Age at initial assessment (approximation)	6.4	0.3	4.6	7.4
Primary language: Spanish	0.32	0.47	0	1
Primary language: Mandarin/Cantonese	0.36	0.48	0	1
Primary language: Other	0.33	0.47	0	1
California English language development test				
Binding Score	−56	52	−176	151
Centered Overall Scale Score	−55	52	−176	151
Centered Listening Scale Score	−18	56	−147	161
Centered Speaking Scale Score	3	60	−156	225
Binding section: Overall	0.92	0.27	0	1
Binding section: Listening	0.06	0.25	0	1
Binding section: Speaking	0.01	0.12	0	1
English learner and special education				
Assigned EL by district	0.836	0.371	0	1
Eligible for EL (binding score < 0)	0.863	0.344	0	1
First placed in SPED between Grades 1 and 6	0.070	0.255	0	1
First placed in SPED between Grades 1 and 6 (if binding score < 0)	0.077	0.266	0	1
First placed in SPED Between Grades 1 and 6 (if binding score ≥ 0)	0.028	0.166	0	1
First placed in SPED in Grade 1	0.016	0.124	0	1
First placed in SPED in Grade 2	0.012	0.111	0	1
First placed in SPED in Grade 3	0.014	0.117	0	1
First placed in SPED in Grade 4	0.012	0.109	0	1
First placed in SPED in Grade 5	0.010	0.097	0	1
First placed in SPED in Grade 6	0.006	0.079	0	1

Source. California Longitudinal Pupil Achievement Data System (CALPADS), SY 2006-07 through SY 2018-19.
Note. The full sample includes students from this district who took the California English Language Development Test (CELDT) for the first time in Kindergarten between SY 2006-07 and SY 2012-13, had valid covariate and outcome data, and scored above the minimum score Overall and in Listening and Speaking domains ($N = 12{,}607$). Excluded from the sample are students who took the initial CELDT assessment for the first time in Grade 1 or later. The intent-to-treat group includes students who scored above the cutoff ($N = 10{,}880$). The control group includes students who scored below the cutoff ($N = 1{,}727$). EL = English Learner; SPED = Special Education.

approaches. Following this line of research, we report raw SPED placement rates and OLS regression results for our data.

A key limitation of such results, however, is the inability to determine how EL status *affects* SPED placement without making strong and likely invalid assumptions (e.g., selection on observables) about the relation. The central contribution of our study is the application of a more advanced quasi-experimental research design that can rigorously estimate a causal impact of EL status on SPED placement for students at the English proficiency margin. Leveraging CELDT scale score data, we differentiate between treatment and control groups by first determining whether the student was eligible for EL status based on their CELDT score. Then, we account for the imperfect take-up of treatment assignment by relying on an instrumental variables (IV) approach that uses the level of compliance with EL assignment based on CELDT scores as the instrument. This type of fuzzy RD or RD-IV approach has been previously applied to EL research in California (e.g., Robinson, 2011; Robinson-Cimpian & Thompson, 2016).

Using this approach, we leverage a core underpinning concept from the RD literature—the idea that students near the EL threshold were similar to one another in expectation and provide a strong counterfactual group (Thistlewaite & Campbell, 1960). Based on the assumption that students were unable to precisely manipulate their score (which appears to hold according to numerous empirical tests presented in Supplementary Tables A1, A2, and A3 and Supplementary Figures A1, A2, and A3 in the online version of the journal), EL status can be considered as good as randomly assigned for ranges of the sample near the threshold. Leveraging this natural experiment that occurs near the arbitrary EL threshold enables the use of a stronger counterfactual comparison group than other quasi-experimental research designs, such as propensity score matching (Angrist & Pischke, 2009; Thistlewaite & Campbell, 1960). Updating our results to account for noncompliance of EL assignment through the RD-IV approach allows us to sharpen our examination and narrowly consider the effects of EL status on SPED placement.

Kindergarten students needed to meet a predetermined cut score across multiple language domains to be classified as English-proficient. Students with an overall scale score below the "Beginning Advanced" level, a listening scale score below the "Intermediate" level, or a speaking scale score below the "Intermediate" level were classified as ELs. The initial step to implement the RD approach in this context is to construct a "binding score" forcing variable that accounts for these three ways in which a kindergarten student could have been classified as an EL (Papay et al., 2011; Porter et al., 2017; Reardon & Robinson, 2012). To do this, we create variables for the overall scale score and the listening and speaking scale scores that are centered around the cut scores for each domain and based on *initial* CELDT assessment results in each particular year. This allows us to put scores from each kindergarten cohort on the same scale despite minor adjustments to the CELDT assessment year to year. For each student, we then take the minimum value of these three variables to create the binding score for student *i* in school *s* and cohort *c*:

$$BindingScore_{isc} = MIN \begin{Bmatrix} CenteredOverall_{isc}; \\ CenteredListening_{isc}; \\ CenteredSpeaking_{isc} \end{Bmatrix}$$

As reported in Table 1, for 92% of students, the binding section was the overall score. For 6% of students, the binding section was listening. For the remaining students, speaking was the binding section.

First, we define the point at which we expect there to be a discontinuous jump in the probability of treatment:

$$EL\,Eligible_{isc} = 1(BindingScore_{isc} < 0)$$

We then apply the first stage equation of our RD design to understand how well the $EL\,Eligible_{isc}$ indicator predicts actual EL status:

$$\begin{aligned} EL\,Status_{isc} = \alpha\left(EL\,Eligible_{isc}\right) \\ + f\left(BindingScore_{isc}\right) \\ + \lambda_{sc} + \beta\mathbf{X_{isc}} + \epsilon_{isc} \end{aligned}$$

In this specification, α signifies the discrete jump at the cutoff for EL assignment and is our coefficient of interest. The indicator $EL\,Eligible_{isc}$ flags observations that were below the cutoff based on the binding score. The $f\left(BindingScore_{isc}\right)$ term represents a flexible function of the binding score.[16] We implement this as specifications that include linear splines of the forcing variable.[17] λ_{sc} represents school-cohort fixed effects, which allow us to remove variation that is consistent across groups of students testing from the same school-cohort combination. $\mathbf{X_{isc}}$ is a vector of student-level covariates, including an approximation of the student's age at initial assessment and indicators for being female, Hispanic, Chinese, having declined to state race/ethnicity, and having the most educated parent being at least a high school graduate. ϵ_{isc} is the mean-zero error term.

Through this specification, we test how our binding score forcing variable influenced the probability of actual EL assignment. We also probe for manipulation of the forcing variable around the cutoff through a variety of different diagnostic tests.[18]

Supplementary Table A1 provides point estimates of the first-stage relationships across a variety of specifications at the Calonico, Cattaneo, and Titiunik (CCT, 2014) optimal bandwidth. We observe large and statistically significant relationships between our discontinuity parameter, α, and EL assignment. In all cases, we estimate the jump in probability of being assigned EL status to be more than 75 percentage points. Importantly, this indicates that in most instances our binding score forcing variable is effectively flagging students that ultimately entered EL status (i.e., were "compliers" with treatment assignment). Further, the F-statistic for our main instrument, $EL\,Eligible_{isc}$, is over 1000 for each reported specification.[19] Still, these results highlight that our binding score forcing variable does not perfectly identify treatment. In some cases, a student may have scored below the threshold but was classified as English-proficient (i.e., was a "never-taker" of treatment assignment) or scored above the threshold but was still classified as EL (i.e., was an "always-taker" of treatment assignment). Figure 2 illustrates the high probability of EL classification based on the CELDT binding score. The visual provides binned averages for the CCT bandwidth.

However, we are most prominently interested in how EL status impacted SPED placement rates. Therefore, the specification of primary interest for our key outcomes relies on the following second-stage equation:

$$Y_{isc} = \gamma \widehat{EL\,Eligible}_{isc} + f\left(BindingScore_{isc}\right) + \lambda_{sc} + \boldsymbol{\beta}\mathbf{X}_{\mathbf{isc}} + \epsilon_{isc}$$

In this fuzzy RD specification, γ is our coefficient of interest on the predicted values from our first stage equation, $\widehat{EL\,Eligible}_{isc}$. In contrast to the first-stage equation, rather than using actual EL status as the outcome measure, we use Y_{isc} to look at our key SPED placement outcomes overall and by grade-level. The remaining terms of this second-stage equation are consistent with the terms from the first-stage equation.

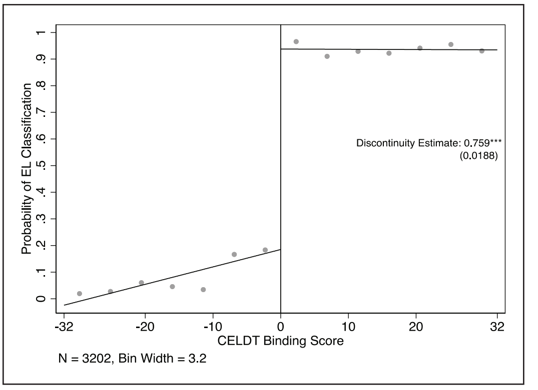

FIGURE 2. *First-stage: Probability of EL classification by CELDT binding score, CCT optimal bandwidth.*

Note. This figure depicts the probability of EL classification by the transformed CELDT Binding Score forcing variable for the CCT optimal bandwidth (+/− 32 points) analytical sample (N = 3,202). Bin widths are 3.2 points. In order to conform to the RD structure required by -rdrobust-, binding scores were transformed linearly in the following way: ([−1*BindingScore_isc] − 1). Positive values denote the treatment group (students who scored below the English proficiency threshold). Negative values denote the control group (students who scored above the English proficiency threshold). Each plotted point provides a binned average; a line of best fit is separately provided for data above and below the threshold. Discontinuity estimate is based on model that includes student covariates and school-cohort fixed effects. CELDT = California English Language Development Test; CCT = Calonico, Cattaneo, & Titiunik (2014).
***$p < .01$. **$p < .05$. *$p < .1$.

The key differences in this specification, therefore, are that (a) by using the $\widehat{EL\,Eligible}_{isc}$ indicator, we are focusing on how EL assignment affected SPED placement for the subset of "complying" students; and (b) by replacing EL assignment with Y_{isc}, we examine SPED placement, our key outcomes of interest.

Given that our primary interest is in how EL status affected SPED placement (i.e., the effect for "compliers"), the fuzzy RD or RD-IV design, in which the $\widehat{EL\,Eligible}_{isc}$ indicator serves as an instrument for the students' actual assignment to EL status, is most appropriate. Importantly, in order to implement the fuzzy RD design, we estimate two-stage-least-squares (2SLS) models. Results from those models are reported in Tables 3, 4, and 5. The reported results, focused exclusively on those individuals who complied with treatment, represent the Treatment on the Treated (TOT) estimand.

In addition, in Table 3 and in Supplementary Tables C1 and C2 (online version of the journal), we also report results from the following "reduced-form" specification for reference:

$$Y_{isc} = \omega\left(EL\ Eligible_{isc}\right) + f\left(BindingScore_{isc}\right) + \lambda_{sc} + \boldsymbol{\beta}\mathbf{X}_{\mathbf{isc}} + \epsilon_{isc}$$

Here, our primary interest is for ω, the coefficient on the $EL\ Eligible_{isc}$ indicator. The key distinction with this model is that we do not rely on the predicted indicator (i.e., $\widehat{EL\ Eligible_{isc}}$), but rather solely examine results for our SPED outcomes of interest based on whether the CELDT binding score was below the specified cutoff. These results are inclusive of compliers, always takers and never takers and provide evidence for the effects on SPED placement in contexts where the compliance rate for EL assignment is similar to our district.

Results

Descriptive Evidence

Raw SPED placement rates for the main sample are presented in Table 1. Supplementary Tables B1 through B4 provide extended descriptive statistics, including the raw placement patterns for EL students and English-proficient students for the overall sample and for primary language subsamples. In our main sample, 7.4% of EL students were placed in SPED between Grades 1 and 6 while 4.8% of English-proficient students were placed in SPED (see Table B1). These raw placement rates suggest overidentification of EL students in SPED by 2.6 percentage points. However, this relationship does not account for any observable or unobservable characteristics.

Table 2 presents OLS regression results for the relation between EL status and SPED placement. The bivariate regression results with no controls presented in column (1) align with the raw difference in rates described above. In columns (2) and (3), we add relevant student-level controls and school-cohort fixed effects. Across specifications, the relationship in our context is significant and positive. The model with all controls and fixed effects suggests that being assigned EL in kindergarten is associated with a 3.4 percentage point increase in the probability of SPED placement between Grades 1 and 6. Thus both the unconditional and conditional (on observable student-level controls and school-cohort fixed effects) differences indicate that EL students were *over*represented in SPED in this district overall.

RD Evidence

Table 3 presents our main RD results examining the effect of EL eligibility and EL status on subsequent SPED placement between Grades 1 and 6. In columns (1) to (3), we report reduced-form RD results, which estimate the intent-to-treat (ITT) estimand. In columns (4) to (6), we report two-stage-least-squares (2SLS) RD results, which estimate the TOT estimand. In all models we account for heteroskedasticity of the error term by implementing Eicker-Huber-White robust standard errors. We present three versions of our main specification. Columns (1) and (4) provide results from RD specifications that includes a linear spline of the forcing variable, but no other controls. Columns (2) and (5) report results from RD specifications that retain the linear spline of the forcing variable and include a set of baseline student-level demographic controls. Columns (3) and (6) present results from RD specifications that retain the linear spline of the forcing variable and the student-level controls and add school-cohort fixed effects.

While we report reduced-form estimates (columns 1–3) of EL eligibility on SPED placement, our preferred specification is the 2SLS estimation of EL status on SPED placement reported in column (6). We prefer the 2SLS estimate over the reduced-form estimate because it considers how EL status rather than EL eligibility impacted SPED placement. In addition, we prefer the specification in column (6) over the specifications reported in columns (4) and (5) due to the inclusion of both student-level covariates and school-cohort fixed effects, which modestly aids precision and controls for other relevant factors.[20]

A crucial consideration for the RD design is the choice of bandwidth. Rows in Table 3 report estimates of the discontinuous jump at the threshold for different bandwidth samples for each of our RD specifications. Row (2) reports results for +/− 32

TABLE 2

OLS: Estimated Association Between EL Assignment and SPED Placement

	Dependent variable: First placed in SPED between Grades 1 and 6		
Independent variable	(1)	(2)	(3)
Assigned EL by district	0.027***	0.023***	0.034***
	(0.0053)	(0.0055)	(0.0059)
Controls	No	Yes	Yes
School-Cohort FE	No	No	Yes
R^2	.002	.026	.077

Note. Robust standard errors are in parentheses. Each model reports the OLS relationship between being assigned EL by the District and placement in SPED between Grades 1 and 6. The analytical sample (N = 12,607) includes students who took the CELDT initial assessment in kindergarten between SY 2006-07 and SY 2012-13, had covariate and outcome measures available, and scored above the minimum score Overall and in Listening and Speaking domains. The model in column (2) includes the following student-level controls: an indicator for female; an indicator for the student's most educated parent being at least a HS graduate; an indicator for being Hispanic; an indicator for being Chinese; an indicator for declining to state race/ethnicity; and the student's age at initial assessment (coefficients suppressed). The model in column (3) includes the same set of student-level controls and adds school-cohort fixed effects (coefficients suppressed). The OLS estimate in column (3) is used to compare to subsequent Regression Discontinuity estimates. OLS = Ordinary Least Squares; EL = English Learner; SPED = Special Education.
***$p < .01$. **$p < .05$. *$p < .1$.

TABLE 3

Reduced-Form and 2SLS Estimates of EL Eligibility and EL Status on SPED Placement, by Bandwidth

	Dependent variable: First placed in SPED between Grades 1 and 6						
	Reduced-form			2SLS			
Bandwidth sample	(1)	(2)	(3)	(4)	(5)	(6)	N
2*CCT (+/− 64)	−0.012	−0.013	−0.015	−0.016	−0.017	−0.019	6,793
	(0.011)	(0.011)	(0.010)	(0.014)	(0.013)	(0.013)	
CCT (+/−32)	−0.006	−0.006	−0.01	−0.008	−0.008	−0.013	3,202
	(0.015)	(0.015)	(0.013)	(0.019)	(0.019)	(0.017)	
0.5*CCT (+/− 16)	−0.01	−0.01	−0.002	−0.015	−0.014	−0.002	1,534
	(0.022)	(0.022)	(0.017)	(0.031)	(0.031)	(0.025)	
Student-level controls	No	Yes	Yes	No	Yes	Yes	—
Cohort-school FE	No	No	Yes	No	No	Yes	—

Note. Robust standard errors are in parentheses. All models include a linear spline of the transformed CELDT binding score forcing variable (see Note 16 for details). Each cell represents a separate regression. The CCT optimal bandwidth is +/− 32. Columns (1) through (3) provide reduced-form estimates of the intent-to-treat (ITT) estimand. Columns (4) to (6) provide 2SLS estimates of the treatment on the treated (TOT) estimand. Models reported in columns (2), (3), (5), and (6) include the following student-level controls: an indicator for female; an indicator for the student's most educated parent being at least a HS graduate; an indicator for being Hispanic; an indicator for being Chinese; an indicator for declining to state race/ethnicity; and the student's age at initial assessment (coefficients suppressed). The models reported in columns (3) and (6) also include school-cohort fixed effects (coefficients suppressed). The point estimate for our preferred specification, model (6), is statistically different ($p < .01$) from the OLS point estimate (i.e., from Table 2, column [3]) for all bandwidth samples. 2SLS = two-stage-least-squares; EL = English Learner; SPED = Special Education; FE = fixed effects.
***$p < .01$. **$p < .05$. *$p < .1$.

points, which is the optimal RD bandwidth suggested by CCT approach. Row (1) reports results for twice the CCT optimal bandwidth, or +/− 64 points from the cut score. Row (3) reports results using half of the CCT optimal bandwidth, or +/− 16 points.

For the CCT optimal bandwidth, +/− 32, we observe small and non-significant estimates of between −0.006 and −0.013 across specifications for ITT and TOT effects. This suggests that students eligible for or classified as ELs were similarly likely to be placed in SPED between Grades 1 and 6 than those eligible for or classified as English-proficient. Results for all models at the other two bandwidths similarly indicate that EL eligibility or classification had no effect on SPED placement between Grades 1 and 6 after initial EL classification in kindergarten, with non-significant estimates between −0.002 and −0.019. Figure 3 presents a graphical illustration for the CCT optimal bandwidth of +/− 32 from the cutoff. Examination of both the visual relationships and the regression results indicates that students scoring below the English proficiency threshold were about as likely as their English-proficient peers to be identified for SPED between Grades 1 and 6. These RD results stand in stark contrast to the positive association from the OLS regression analysis.

We further probe the robustness of our main results across a larger number of bandwidth samples using our preferred specification (i.e., the RD-IV specification with a linear spline, student-level controls, and school-cohort fixed effects). Figure 4 presents the 2SLS point estimates and their corresponding 95% confidence intervals. For relatively small bandwidth samples, the point estimates are somewhat volatile and the confidence intervals are wider. This is what we would expect given the smaller sample sizes. As bandwidths increase, confidence intervals narrow and point estimates stabilize around −0.019, or 1.9 percentage points less than English-proficient students. Importantly, our point estimates are largely indistinguishable from zero (the blue dotted line), indicating that we have a precisely estimated null result. However, these estimates are statistically different from the regression analysis estimates (the red dotted line). For all bandwidth samples presented, we reject the null hypothesis (when $\alpha = 0.05$) that the RD coefficient of interest is equal to our OLS regression coefficient.

Table 4 reports 2SLS results for our preferred specification for first being placed in SPED during each grade separately. Column (1) replicates the RD result for our main outcome, first SPED placement between Grades 1 and 6. Column (2) reports the RD estimate for first placed in SPED in Grade 1; Column (3) for first placed in SPED

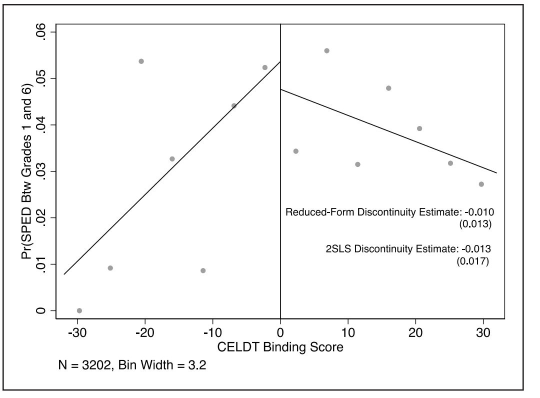

FIGURE 3. *Probability of SPED placement between Grades 1 and 6 by binding score forcing variable, reduced-form and 2SLS estimates.*
Note. This figure depicts the probability of being first placed in SPED between Grades 1 and 6 by the transformed CELDT Binding Score forcing variable for the CCT optimal bandwidth (+/− 32) analytical sample (N = 3,202). Bin widths are 3.2 points. In order to conform to the RD structure required by -rdrobust-, binding scores were transformed linearly in the following way: ([−1*BindingScore_isc] – 1). Positive values denote the treatment group (students who scored below the English proficiency threshold). Negative values denote the control group (students who scored above the English proficiency threshold). Each plotted point provides a binned average; a line of best fit is separately provided for data above and below the threshold. Discontinuity estimate is based on model that includes student covariates and school-cohort fixed effects. SPED = Special Education; 2SLS = two-stage-least-squares.
***$p < .01$. **$p < .05$. *$p < .1$.

in Grade 2; and Columns (4) to (7) follow in cognate form. In identical form to Table 3, the rows of this table correspond to different bandwidth samples (i.e., +/− 64, +/− 32, and +/− 16).

Results from columns (2–5) and (7) indicate no major difference in SPED placement probability in Grades 1 to 4 or Grade 6 for students classified as EL and English-proficient. Results from column (6), however, show marginally-significant ($p < .1$) point estimates of about −0.01 in bandwidth samples +/− 64 and +/− 32. This provides suggestive evidence that EL status was leading to modest underidentification for SPED in Grade 5.

In Table 5, we examine heterogeneous effects by students' primary language for our main outcome of interest.[21] Column (1) presents results for students whose primary language was Spanish. Here, we observe negative point estimates across bandwidths, and a statistically-significant −0.090

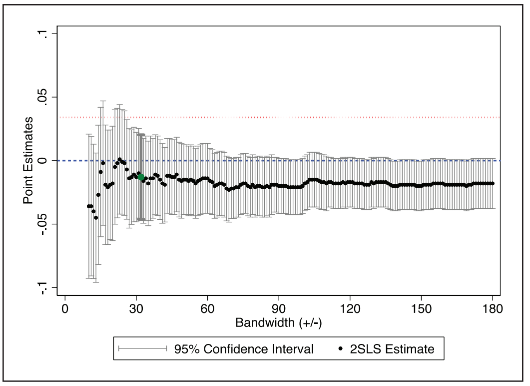

FIGURE 4. *Sensitivity of 2SLS estimates of EL status on SPED placement, across bandwidths.*
Note. The dependent variable is a flag for being placed in SPED between Grades 1 and 6. All models are the 2SLS specification and include a linear spline of the transformed forcing variable (see Note 16 for details), school-cohort fixed effects and the following student-level controls: an indicator for female; an indicator for the student's most educated parent being at least a HS graduate; an indicator for being Hispanic; an indicator for being Chinese; an indicator for declining to state race/ethnicity; and the student's age at initial assessment. We exclude point estimates for samples with bandwidths less than or equal to +/− 9. The 95% confidence interval around each estimate is also graphed. The lower dotted horizontal line is 0 and the upper dotted horizontal line is the full analytical sample OLS benchmark estimate of 0.034. The 2SLS estimate and confidence interval for the CCT optimal bandwidth (+/− 32) are bolded. 2SLS = two-stage-least-squares; EL = English Learner; SPED = Special Education.

for the CCT optimal bandwidth. These results suggest that Spanish-speaking students classified as ELs were less likely to be placed in SPED between Grades 1 and 6 than Spanish-speaking students classified as English-proficient. To contextualize these results, about 10.8% of Spanish-speaking students in the sample were placed in SPED between Grades 1 and 6. Relative to this rate, the magnitude of our RD estimate is substantial. Column (2) of Table 5 presents RD results for the subsample of students who primarily spoke Mandarin or Cantonese. Point estimates are close to zero and not statistically significant across all reported bandwidths. This suggests proportionate representation of these EL students placed in SPED between Grades 1 and 6. Column (3) of Table 5 shows estimates for students who primarily spoke languages other than English, Spanish, Mandarin, or Cantonese. Here, too, none of the point estimates are statistically significant.[22]

Figure 5 provides graphical illustrations of these heterogeneous results: Panel A for ELs whose primary language was Spanish; Panel B for ELs whose primary language was Mandarin or Cantonese; and Panel C for ELs whose primary language was a language other than English, Spanish, Mandarin, or Cantonese. Based on the point estimates and graphical results, we find suggestive evidence of underidentification for students whose primary language was Spanish and no evidence of disproportionate representation for the other two subgroups.

Discussion

Prior research has documented concerns specifically about EL students' being misidentified for SPED services and focused the analyses on disproportionate representation of EL students in SPED (Morgan et al., 2015, 2018; Umansky et al., 2017). However, largely due to data restrictions, these efforts have mainly resulted in descriptive findings. In this study, we present descriptive as well as causal evidence pertaining to the interaction of EL status and SPED placement in our partner district.

Descriptive and Causal Evidence

First, we report unconditional SPED placement rates for all students in our analytical sample, by EL eligibility, and by assigned EL status (see Table 1 and Supplementary Tables B1–B4). These are raw SPED placement rates for all students in the English proficiency score distribution without any statistical controls, separated by either the cut score or assigned EL status. If we subtract the rate for EL students from the rate for English-proficient students, we get a raw difference in means between these groups without any controls. By and large, EL students had higher SPED placement rates.

Then, we present OLS estimates (Table 2), which enable the assessment of the difference in probability of SPED placement when classification changes from English-proficient to EL. Importantly, in computing our OLS estimates, we are able to control for a variety of student characteristics (i.e., race/ethnicity; parental education level; and age) and school-cohort fixed effects (i.e.,

TABLE 4

2SLS Estimates of EL Status on Other SPED Placement Outcomes, by Bandwidth

	Dependent variable: First placed in SPED in							
	Grades 1 through 6	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	
Bandwidth sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	*N*
2*CCT (+/− 64)	−0.019	−0.004	0.010	0.004	0.003	−0.009*	−0.002	6,793
	(0.013)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	
CCT (+/− 32)	−0.013	−0.000	−0.013	0.012	0.008	−0.012*	−0.008	3,202
	(0.017)	(0.006)	(0.008)	(0.008)	(0.006)	(0.007)	(0.007)	
0.5*CCT (+/− 16)	−0.002	−0.005	−0.004	0.001	0.007	−0.000	0.002	1,534
	(0.025)	(0.008)	(0.012)	(0.011)	(0.009)	(0.011)	(0.011)	

Note. Robust standard errors are in parentheses. Each cell represents a separate regression. The +/− 32 bandwidth is the CCT optimal bandwidth. All models provide 2SLS estimates of the treatment on the treated (TOT) estimand and include a linear spline of the transformed forcing variable (see Note 16 for details), student-level controls and school-cohort fixed effects (coefficients suppressed). The student-level controls include the following: an indicator for female; an indicator for the student's most educated parent being at least a HS graduate; an indicator for being Hispanic; an indicator for being Chinese; an indicator for declining to state race/ethnicity; and the student's age at initial assessment. 2SLS = two-stage-least-squares; EL = English Learner; SPED = Special Education.

***$p < .01$. **$p < .05$. *$p < .1$.

TABLE 5

2SLS Estimates of EL Status on SPED Placement, by Bandwidth and Primary Language

	Dependent variable: First placed in SPED between Grades 1 and 6					
	Spanish		Mandarin or Cantonese		All other languages	
Bandwidth sample	(1)	*N*	(2)	*N*	(3)	*N*
2*CCT	−0.045	2,185	0.008	2,296	−0.010	3,826
	(0.029)		(0.014)		(0.016)	
CCT	−0.090**	906	0.007	1,028	−0.023	2,451
	(0.036)		(0.018)		(0.020)	
0.5*CCT	−0.017	411	0.007	471	−0.010	1,239
	(0.043)		(0.028)		(0.029)	

Note. Robust standard errors are in parentheses. Each cell represents a separate regression. The CCT optimal bandwidth for the Spanish-speaking sample is +/− 37; for the Mandarin- or Cantonese-speaking sample is +/− 32; and for the sample of all other languages is +/− 57. All models provide 2SLS estimates of the treatment on the treated (TOT) estimand and include a linear spline of the transformed forcing variable (see Note 16 for details), student-level controls and school-cohort fixed effects (coefficients suppressed). The student-level controls include the following: an indicator for female; an indicator for the student's most educated parent being at least a HS graduate; an indicator for being Hispanic; an indicator for being Chinese; an indicator for declining to state race/ethnicity; and the student's age at initial assessment. The mean SPED placement rate between Grades 1 and 6 after the initial CELDT assessment for key subgroups are as follows: 0.108 for Spanish speakers overall; 0.114 for Spanish speakers scoring below the threshold; 0.048 for Spanish speakers score above the threshold; 0.037 for Cantonese or Mandarin speakers overall; 0.040 for Cantonese or Mandarin speakers below the threshold; 0.008 for Cantonese or Mandarin speakers above the threshold; 0.070 for speakers of all other languages overall; 0.082 for speakers of all other languages below the threshold; and 0.029 for speakers of all other languages above the threshold. 2SLS = two-stage-least-squares; EL = English Learner; SPED = Special Education.

***$p < .01$. **$p < .05$. *$p < .1$.

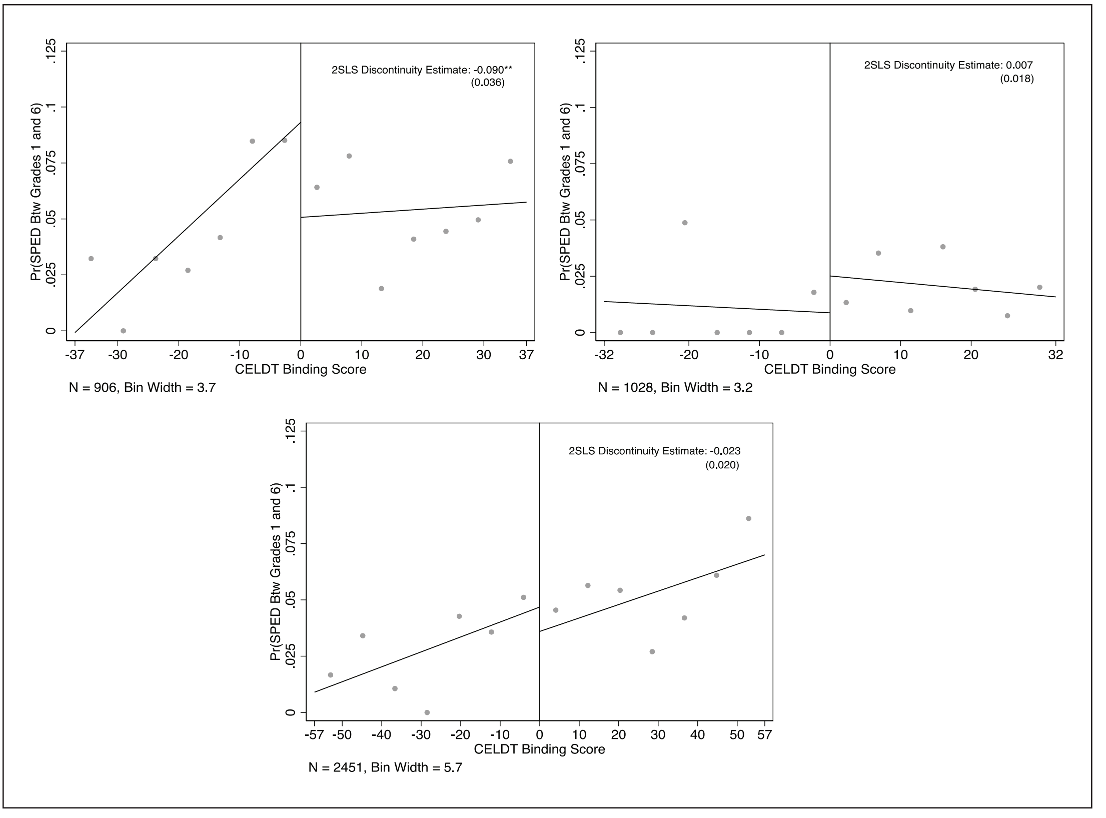

FIGURE 5. *Probability of SPED placement between Grades 1 and 6 by binding score forcing variable and primary language at the CCT optimal bandwidth, 2SLS estimates.*

Note. These figures illustrate the probability of first being placed in SPED between Grades 1 and 6 by the transformed CELDT Binding Score forcing variable. The CCT optimal bandwidths for analytical samples by primary language are used. The leftmost graph is for students with a primary language of Spanish (CCT optimal bandwidth: +/− 37; N = 906; bin width: 3.7 points); the middle graph is for students with a primary language of Mandarin or Cantonese (CCT optimal bandwidth: +/− 32; N = 1,028; bin width: 3.2 points); and the rightmost graph is for students with a different primary language (CCT optimal bandwidth: +/− 57; N = 2,451; bin width: 5.7 points). In order to conform to the RD structure required by -rdrobust-, binding scores were transformed linearly in the following way: ([−1*BindingScore_isc] – 1). Positive values denote the treatment group (students who scored below the English proficiency threshold). Negative values denote the control group (students who scored above the English proficiency threshold). Each plotted point provides a binned average; a line of best fit is separately provided for data above and below the threshold. Discontinuity estimate is based on model that includes student covariates and school-cohort fixed effects. SPED = Special Education; 2SLS = two-stage-least-squares.

common changes for students attending the same school in the same cohort). The contrast between our raw difference reported in Table 2, column 1 and OLS estimates with controls (Table 2, columns (2) and (3)) can be interpreted as the part of the relation that is explained away by the controls in the OLS models. Both OLS estimates and raw placement rates facilitate the estimation of average SPED participation rates for students who scored above and below the cut score. An advantage of the OLS estimates is that the estimation of the conditional mean difference in SPED classification rates between students who scored above and below the CELDT threshold is computed using data for all students in the score distribution. A downside to the OLS approach is that the results would be biased estimates of the causal effect if (observable or unobservable) variables that are correlated with both the treatment and the outcome are omitted. Since we only observe a few control variables, we treat our OLS estimates cautiously and interpret them as descriptive rather than causal evidence. In addition, the effect of EL status on SPED placement may be heterogeneous across the full range of English language proficiency, which our single OLS estimate is unable to capture.

Both the raw SPED placement rates and the OLS estimate indicate that in this district overall, EL students are overrepresented in SPED services. This has important implications for teacher education and professional development. SPED teachers need to be trained for supporting EL students; EL teachers also need to be prepared to support students with a disability. To ensure that both support services are aligned and integrated with core academic content, close collaboration is required for school and district staff who specialize in EL and SPED services and core subject instruction.

Finally, we report RD results (Tables 3–5), which allow us to estimate the probability of SPED placement in our partner district separately for observations just above the cut score and observations just below the cut score. Our coefficient of interest indicates the size of the gap between these estimated probabilities, or the jump in the probability of SPED placement at the cut score. Due to the predetermined nature of the CELDT cut scores and the inability for students to precisely manipulate their scores, these RD results can be interpreted as credibly causal estimates of the effect of scoring below the cut score (a strong proxy for EL status) on SPED placement. However, this LATE is only valid if assumptions for the RD design are met (e.g., the continuity of the forcing variable through the threshold). We probe these assumptions in Supplementary Appendix A and do not find clear evidence suggesting that they are invalid. By using the RD design, the estimation of the causal impact is not prone to omitted variable bias like in OLS, illustrating a substantial strength of this approach. Furthermore, the RD design provides an important and highly rigorous snapshot of differentials in SPED placement rates at the English proficiency margin.

RD Findings at the English Proficiency Margin

Using this rigorous quasi-experimental research design and data from a large school district, we provide new evidence about the effect of EL status on SPED placement at the English proficiency threshold. In contrast to positive correlations suggested by regression analyses, our main RD results at the English proficiency margin indicate a null or slightly negative overall effect of EL status on SPED placement between Grades 1 and 6 after initial EL classification in kindergarten. Our results indicate that differences are modest when comparing SPED placement rates for students who barely reach and students who barely miss the English proficiency threshold in kindergarten. EL status, therefore, led to proportionate identification of EL students for SPED in our partner district when compared with their English proficient peers. This finding suggests that qualitative analyses about the districts' identification protocols and practices would likely be informative and valuable.

Suggestive Evidence for Heterogeneity by Grade Level and Primary Language

Furthermore, we were able to consider disproportionate representation over time and by primary language subgroups. Our results indicate that EL underrepresentation by about 1 percentage point arose during Grade 5. One potential explanation for underidentification of ELs in Grade 5 is the preparation for elementary to middle school transition, which typically takes place after Grade 5 in the district. It is possible that to prepare for this transition, school staff are paying extra attention to students' learning challenges and needs. Teachers may be more likely to refer students who are native English users or initially English-proficient to be assessed for disabilities. In contrast, EL students may be less likely to be referred for a disability assessment if the teacher believes the challenges EL students experience with learning are due to developing language proficiency and not due to a potential disability.

We were also able to examine effect heterogeneity by primary language. Practices for identifying disabilities used by the district factor prominently into understanding these results. In our partner district, many tools for assessing needs due to disabilities are provided in English, Spanish, Cantonese, and Mandarin. In other words, when ELs who are Spanish and Mandarin/Cantonese speakers are assessed for disabilities, the assessments can often be conducted in their primary language, while ELs who speak other languages are assessed in English. We would expect the availability of primary language assessment to lead to more accurate placement that matches students' needs. But this could either increase or decrease the rate of SPED placement relative to using an English assessment. Although both Spanish-speaking and Mandarin/Cantonese-speaking ELs had access to primary-language SPED assessment, we find suggestive evidence that Spanish-speaking ELs were underidentified for SPED relative to Spanish-speaking English-proficient students while Mandarin/Cantonese-speaking ELs were proportionately identified. It is possible that Spanish-speaking and Mandarin/Cantonese-speaking ELs were given primary-language SPED assessments at different rates, or that SPED assessments performed differently with respect to primary language—either of which may have had an impact on referral outcomes. However, we do not observe the language of the assessment administered to each student and are unable to analyze the effect of primary-language assessments. The role of primary-language assessment in SPED placement is worth investigating in future research.

Limitations and Future Directions

An important limitation of the RD estimates described above is that they are most relevant and valid for the subset of students close to the English proficiency cut score. In other words, our RD results likely do not provide causal estimates for students who are far below or far above the cut score. Students with very low English proficiency may have SPED placement rates, as well as SPED service experiences and outcomes, that are different from students at the English proficiency margin. Unfortunately, the RD design does not allow us to interrogate the causal relations among English proficiency, EL classification, and SPED placement for this group. This is an important limitation to our RD approach and highlights the ongoing value of the descriptive results. In this case, raw placement rates and OLS estimates help us to contextualize the RD estimates, and vice versa. In presenting descriptive and causal findings, we aim to provide a more nuanced picture of the complex relations among English proficiency, EL status, and SPED placement.

Several other limitations to this study merit consideration. First, the data come from one district with a long history of serving a large, diverse EL student population and a mature research-practice partnership with a large research university, so the findings may be influenced by this context. In other district settings, ELs may be overrepresented, proportionally represented, or underrepresented in SPED, and variation by primary language may be different. This district's demographic composition and extensive experience in addressing the needs of immigrant and EL students are unique. For example, the district has a tradition of close collaboration between its large EL and SPED departments, and multiple staff members specialize in identifying, serving, and reclassifying students for *both* English language and SPED needs. As a result, the district's SPED identification process for emergent multilingual students is somewhat distinct from English-proficient students. Thus, our results may not be generalizable to other districts which serve smaller, less diverse EL populations or apply the same SPED identification procedures to all students. District leaders should consider the demographic composition of their own EL and general student populations as well as their EL and SPED classification policies and services when evaluating the applicability of these findings to their context. Future work can consider exploring the placement of EL students in SPED across different districts contexts.

Second, the data did not support further disaggregation by disability type or a robust exploration of results for students taking the initial CELDT assessment after kindergarten. For example, the proportionality of SPED placement by EL status may have differed among the disability categories, but we did not have full access to student-level disability categories. Thus, our results may mask important heterogeneity along this line. This topic should be explored further in future work. Also, the sample size did not allow a robust standalone analysis of ELs who entered the district between Grades 1 and 5. Supplementary Tables D1 and D2 provide a preliminary analysis for this sample, but future research should consider differential effects of later district entry on dual identification when more data become available.

Finally, this study was not able to identify the precise mechanisms driving the observed effect of kindergarten EL classification or the source of heterogeneous treatment effect by students' primary language. Detailed mixed-methods inquiry into this topic could prove valuable for identifying such mechanisms in future research.

Conclusion

For students with disabilities, timely and appropriate placement into educational support services is essential (Burr, 2019). Delayed identification or misidentification can be academically and psychologically harmful for students (Carnock & Silva, 2019). Our results offer a causal assessment of how EL status in kindergarten affected subsequent SPED placement in our partner district using an RD design. These findings advance our understanding of the representation of EL students in SPED around the English proficiency margin. They demonstrate the viability of the RD approach in this context and suggest that ongoing research can use RD in addition to descriptive analyses to better understand the interaction of these two major educational service programs. Our results also shed light on *when* and *for whom* disproportionate representation occurred. Such information is valuable to our partner district as they look to improve policy pertaining to the SPED placement process for EL students.

Having discussed the findings with the research team, the district has begun multiple initiatives. First, both SPED and EL departments have set out to analyze qualitative data on SPED identification, with a focus on ELs with Spanish as their primary language. Second, the partnership plans to examine additional data to analyze EL pathways through SPED identification and programs. This study has thus motivated a series of mixed-methods inquiries aimed at continuing to develop more equitable practices around SPED identification and services.

Scholarship that advances existing literature and that provides directly usable results for practitioners highlights the valuable contributions of research conducted through research-practice partnerships. Furthermore, the application of the RD design to explore this topic sets the stage for ongoing research studying the interaction of EL and SPED services in other contexts.

Acknowledgments

We are grateful for the collaborative efforts of our partner school district and all participating staff. We appreciate helpful comments from Thomas Dee, Eric Bettinger, David Plank, and Kalena Cortes. We also thank seminar participants from the Stanford Center for Education Policy Analysis, the John Garner Center for Youth and Their Communities, the University of Hawai'i at Mānoa Department of Educational Psychology, the Association for Education Finance and Policy, and the Association of Public Policy Analysis and Management.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: We acknowledge funding from the California Education Partners through grant 162875 and support from the Institute of Education Sciences, U.S. Department of Education through grant R305B140009 to the Board of Trustees of the Leland Stanford Junior University. The opinions expressed are those of the authors and do not represent the views of IES the U.S. Department of Education, the Board of Trustees of the Leland Stanford Junior University, the Board of Regents of the University of Hawai'i, or NWEA.

ORCID iDs

Mark Murphy https://orcid.org/0000-0002-5709-8390

Angela Johnson https://orcid.org/0000-0003-4114-2446

Supplemental material

Supplemental material for this article is available online.

Notes

1. Overall, during the 2014-15 school year, the western and southwestern regions had substantially larger current-EL populations than most other parts of the United States (U.S. DOE, 2017).

2. Specifically, the categories are (a) autism, (b) deaf-blindness, (c) developmental delay, (d) emotional disturbance, (e) hearing impairment, (f) intellectual disability, (g) multiple disabilities, (h) orthopedic impairment, (i) other health impairment,

(j) specific learning disability, (k) speech or language impairment, (l) traumatic brain injury, and (m) visual impairment (including blindness) (Carnock & Silva, 2019).

3. The next most commonly experienced categories of disability were "other health impairment" and "autism," which together accounted for just under one quarter of all disability classifications. The next three most prominent categories are "developmental delay," "intellectual disability" and "emotional disturbance." Together, these three categories made up approximately 18% of all disabilities classified. Finally, hearing impairments, multiple disabilities, orthopedic impairments, traumatic brain injury and visual impairments consistently made up less than 5% of all disabilities classified (U.S. DOE, 2013). In addition, approximately 10% of SPED students also classified as being an EL (National Council on Disability, 2018).

4. From this point forward, we refer to 1 year after being designated for EL status as Grade 1, 2 years after as Grade 2, and so on. However, it is possible that a student was retained in a particular grade for a second year. Due to missingness in the grade level variable in our dataset, we cannot exactly estimate the frequency of this occurrence. The district reports a low level of grade retention overall, however, suggesting that it was quite infrequent.

5. After 2016-17, California began implementing the English Language Proficiency Assessments for California (ELPAC) and discontinued the use of the CELDT.

6. In Supplementary Tables D1 and D2, we report results for an additional sample of students who took the initial CELDT assessment between Grades 1 and 5. We privilege the kindergarten sample in our main analysis because students who enter the district and are assessed at kindergarten entry are more comparable to one another than students who enter the district in later grades. Of the students taking the initial CELDT assessment between kindergarten and Grade 5, more than 75% were assessed during the kindergarten year. The focus on kindergarten CELDT takers is also consistent with other recent RD studies relying on the CELDT forcing variable (e.g., see Johnson, 2019; Shin, 2018; Umansky, 2016b).

7. We exclude students that scored the minimum score overall or in the speaking or listening domains because the CELDT assessment simply gives these students the lowest raw score and does not differentiate between abilities at these levels. We also exclude students with missing outcome or covariate data. More details about the analytical sample can be found in Supplementary Appendix A.

8. California refers to students scoring above the cut score as Initially Fluent English Proficient (IFEP); in other contexts, the term "English-proficient" is more commonly used.

9. A critical tradeoff in our sample construction was between the number of cohorts to include and the length of time for which we would observe their outcome. We chose this analytical sample in an effort to study the most relevant time window with the greatest statistical power.

10. In OLS and RD results reported in Supplementary Tables C4 and C5, we examine more cohorts of students across shorter periods of time, as well as fewer cohorts of students across longer periods of time.

11. Supplementary Table C6 reports a robustness check in which we calculate 2SLS RD estimates using being placed in SPED between kindergarten and Grade 6 as the outcome. We observe slightly larger negative effects, but a pattern that is quite similar to our main results.

12. Supplementary Table B1 provides extended descriptive results for the main analytical sample for (a) students above and below the CELDT threshold and (b) students classified as EL and English-proficient by the district.

13. Supplementary Appendix A describes our process for constructing the parental education variable.

14. Since our data only provide the birth month and birth year of each student, we approximated student age at the initial CELDT assessment.

15. Supplementary Tables B2, B3, and B4 provide extended descriptive statistics for the subsamples by primary language for (a) students above and below the CELDT threshold and (b) students classified as EL and English-proficient by the district.

16. In order to conform to the RD structure required by the -rdrobust- command in Stata (which we rely on for implementing our analyses), binding scores needed to be transformed linearly in the following way: $([-1 * BindingScore_{isc}] - 1)$. Positive values denote the treatment group (students who scored below the English proficiency threshold) and the difference when subtracting their score from the cut score. Negative values denote the control group (students who scored above the English proficiency threshold) and the difference when subtracting their score from the cut score. We rely on this transformed binding score forcing variable for all manipulation tests and RD analyses (i.e., including first stage, reduced-form, and 2SLS analyses).

17. While we examined models that incorporated quadratic splines of the forcing variable, the Akaike Information Criteria (AIC) indicated that linear specifications should be privileged. As a result, we principally report specifications with linear splines of the transformed forcing variable. We also explored allowing various higher order polynomial specifications of the transformed forcing variable. Supplementary Figure C1 provides the visuals from these examinations, which tend to show a similar relation to what we observe using the linear specification of the forcing variable.

18. Supplementary Figures A1, A2, and A3 and Supplementary Tables A1 and A2 provide evidence pertaining to the continuity of the forcing variable. Supplementary Figure A1 presents raw histograms of the transformed forcing variable (see Note 18) for full and CCT samples. Supplementary Figure A2 shows results from the McCrary (2008) density test. Supplementary Figure A3 illustrates results from the Cattaneo et al. (2018) density test. Supplementary Table A1 provides the first stage analysis and Supplementary Table A2 reports covariate balance across the threshold. The combined evidence does not suggest a violation of the continuity assumption. In addition to these checks, Supplementary Figures A4, A5, and A6 provide first stage analyses at the relevant CCT optimal bandwidth, McCrary (2008) density tests and raw histograms for the forcing variable for each primary language subsample: Spanish, Mandarin/Cantonese, and all other languages. For these subsamples, we similarly observe no evidence of a discontinuity at the threshold.

19. We also test the first stage relationship for three primary language subsamples (i.e., Spanish, Mandarin/Cantonese, and all other languages) and find quite consistent results across language groups. As shown in Supplementary Figures A4, A5, and A6, we observe large and statistically significant jumps at the threshold (all greater than 0.7) with the first bin to the left of the threshold exhibiting the largest rate of non-compliance: a likelihood of EL classification between 0.1 and 0.2. The observed fuzziness is quite similar across subsamples.

20. However, we know that reduced-form estimates may be of interest. Supplementary Tables C1 and C2 present reduced-form estimates that parallel Tables 4 and 5 from the main analysis.

21. Each language subgroup is only about one-third of the main analytical sample, which reduces the precision of our estimates.

22. Supplementary Table C3 provides grade-by-grade 2SLS RD results for the CCT bandwidth across the three primary language subgroups.

References

Albus, D., & Thurlow, M. (2013). *2012-13 publicly reported assessment results for students with disabilities and ELLs with disabilities*. National Center on Educational Outcomes.

Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Artiles, A. J., Rueda, R., Salazar, J. J., & Higareda, I. (2005). Within-group diversity in minority disproportionate representation: English language learners in urban school districts. *Exceptional Children*, *71*(3), 283–300.

Berkeley, S., Scruggs, T. E., & Mastropieri, M. A. (2010). Reading comprehension instruction for students with learning disabilities, 1995-2006: A meta-analysis. *Remedial and Special Education*, *31*(6), 423–436.

Burr, E. (2019). *Guidance manuals for educators of English learners with disabilities: Ideas and lessons from the field*. National Center on Educational Outcomes.

Burr, E., Haas, E., & Ferriere, K. (2015). *Identifying and supporting English learner students with learning disabilities: Key issues in the literature and state practice* (pp. 1–62). http://ies.ed.gov/ncee/edlabs/projects/project.asp?projectID=4483

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, *14*(4), 909–946.

Carlson, D., & Knowles, J. E. (2016). The effect of English language learner reclassification on student ACT scores, high school graduation, and postsecondary enrollment: Regression discontinuity evidence from Wisconsin. *Journal of Policy Analysis and Management*, *35*(3), 559–586. https://doi.org/10.1002/pam.21908

Carnock, J. T., & Silva, E. (2019). *English learners with disabilities: Shining a light on dual-identified students* (pp. 1–25). https://d1y8sb8igg2f8e.cloudfront.net/documents/English_Learners_with_Disabilities_Shining_a_Light_on_Dual-Identified_Students_TaYvpjn.pdf

Cattaneo, M. D., Jannsson, M., & Ma, X. (2018). Manipulation testing based on density discontinuity. *Stata Journal*, *18*(1), 234–261. https://doi.org/10.1177/1536867x180180011

Cioè-Peña, M. (2017a). Bilingualism, disability and what it means to be normal. *Journal of Bilingual Education Research & Instruction*, *19*(1), 138–160.

Cioè-Peña, M. (2017b). The intersectional gap: How bilingual students in the United States are excluded from inclusion. *International Journal of Inclusive Education*, *21*(9), 906–919. https://doi.org/10.1080/13603116.2017.1296032

Cioè-Peña, M. (2020a). Planning inclusion: The need to formalize parental participation in individual education plans (and meetings). *Educational Forum*, *84*(4), 377–390. https://doi.org/10.1080/00131725.2020.1812970

Cioè-Peña, M. (2020b). Raciolinguistics and the education of emergent bilinguals labeled as disabled. *Urban Review*, *53*, 443–469. https://doi.org/10.1007/s11256-020-00581-z

Every Student Succeeds Act of 2015 Pub. L. No. 1177 (2015-2016). (2015).

Fagan, D. S., & Pentón Herrera, L. J. (2022). Supporting English learners with supporting English learners with disabilities. *The State Education Standard*, *22*(1), 26–31.

Figueroa, R. A., & Newsome, P. (2006). The diagnosis of LD in English learners: Is it nondiscriminatory? *Journal of Learning Disabilities*, *39*(3), 206–214. https://doi.org/10.1177/00222194060390030201

Greenberg Motamedi, J., Cox, M., Williams, J., & Deussen, T. (2016). *Examining the representation of English learners in special education in Washington State*. Education Northwest, Regional Educational Laboratory Northwest.

Hibel, J., & Jasper, A. D. (2012). Delayed special education placement for learning disabilities among children of immigrants. *Social Forces*, *91*(2), 503–529. https://doi.org/10.1093/sf/sos092

Individuals with Disabilities in Education Act. (2004).

Johnson, A. (2019). The effects of English learner classification on high school graduation and college attendance. *AERA Open*, *5*(2), 1–23.

Klingner, J. K., Artiles, A. J., Kozleski, E., Harry, B., Zion, S., Tate, W., Durán, G. Z., & Riley, D. (2005). Addressing the disproportionate representation of culturally and linguistically diverse students in special education through culturally responsive educational systems. *Education Policy Analysis Archives*, *13*, 1–43.

Macswan, J., & Rolstad, K. (2006). How language proficiency tests mislead us about ability: Implications for English language learner placement in special education. *Teachers College Record*, *108*(11), 2304–2328. https://doi.org/10.1111/j.1467-9620.2006.00783.x

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698–714. https://doi.org/10.1016/j.jeconom.2007.05.005

Morgan, P. L., Farkas, G., Cook, M., Strassfeld, N. M., Hillemeier, M. M., Pun, W. H., & Schussler, D. L. (2017). Are black children disproportionately overrepresented in special education? A best-evidence synthesis. *Exceptional Children*, *83*(2), 181–198. https://doi.org/10.1177/0014402916664042

Morgan, P. L., Farkas, G., Cook, M., Strassfeld, N. M., Hillemeier, M. M., Pun, W. H., Wang, Y., & Schussler, D. L. (2018). Are Hispanic, Asian, Native American, or language-minority children overrepresented in special education? *Exceptional Children*, *84*(3), 261–279. https://doi.org/10.1177/0014402917748303

Morgan, P. L., Farkas, G., Hillemeier, M. M., Mattison, R., Maczuga, S., Li, H., & Cook, M. (2015). Minorities are disproportionately underrepresented in special education: Longitudinal evidence across five disability conditions. *Educational Researcher*, *44*(5), 278–292. https://doi.org/10.3102/0013189X15591157

Morgan, P. L., Frisco, M. L., Farkas, G., & Hibel, J. (2010). A propensity score matching analysis of the effects of special education services. *Journal of Special Education*, *43*(4), 236–254.

National Council on Disability. (2018). *Broken promises: The underfunding of IDEA*.

Ortiz, A. A., & Polyzoi, E. (1986). *Characteristics of limited English proficient Hispanic students in programs for the learning disabled: Implications for policy, practice and research. Part 1. Report summary*. US Department of Education.

Papay, J. P., Willett, J. B., & Murnane, R. J. (2011). High-school exit examinations and the schooling decisions of teenagers: A multi-dimensional regression-discontinuity analysis. *Journal of Research on Educational Effectiveness*, *7*(1), 1–27.

Park, S. (2014). Bilingualism and children with Autism Spectrum Disorders: Issues, research, and implications. *NYS TESOL Journal*, *1*(2), 122–129.

Park, S. (2019a). Beyond underrepresentation: Constructing disability with Young Asian American children to preserve the "model minority" stereotype. *Asia-Pacific Journal of Research in Early Childhood Education*, *13*(3), 73–95. https://doi.org/10.17206/apjrece.2019.13.3.73

Park, S. (2019b). Disentangling language from disability: Teacher implementation of Tier 1 English language development policies for ELs with suspected disabilities. *Teaching and Teacher Education*, *80*, 227–240. https://doi.org/10.1016/j.tate.2019.02.004

Park, S. (2020). Demystifying disproportionality: Exploring educator beliefs about special education referrals for English learners. *Teachers College Record*, *122*(5), 210–234.

Park, S., Shafer Willner, L., Consortium Jen Paul, W., & Evers, T. (2016). *English language learners with disabilities: A call for additional research and policy guidance* (pp. 1–31). https://www.oregon.gov/ode/schools-and-districts/grants/ESEA/EL/Documents/webpage.pdf

Pasternack, R. (2014). *Reinventing special education: Closing the gap between disabled and non-disabled students*. Accelify Consulting.

Porter, K. E., Reardon, S. F., Unlu, F., Bloom, H. S., & Cimpian, J. R. (2017). Estimating causal effects of education interventions using a two-rating regression discontinuity design: Lessons from a simulation study and an application. *Journal of Research on Educational Effectiveness*, *10*(1), 138–167.

Reardon, S. F., & Robinson, J. P. (2012). Regression discontinuity designs with multiple rating-score variables. *Journal of Research on Educational Effectiveness*, *5*(1), 83–104. https://doi.org/10.1080/19345747.2011.609583

Robinson, J. P. (2011). Evaluating criteria for English learner reclassification: A causal-effects approach using a binding-score regression discontinuity design with instrumental variables. *Educational Evaluation and Policy Analysis*, *33*(3), 267–292. https://doi.org/10.3102/0162373711407912

Robinson-Cimpian, J. P., & Thompson, K. D. (2016). The effects of changing test-based policies for reclassifying English learners. *Journal of Policy Analysis and Management*, *35*(2), 279–305. https://doi.org/10.1002/pam.21882

Samson, J. F., & Lesaux, N. K. (2009). Language-minority learners in special education: Rates and predictors of identification for services. *Journal of Learning Disabilities*, *42*(2), 148–162. http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ830432&site=ehost-live https://doi.org/10.1177/0022219408326221

Schwartz, A. E., Hopkins, B. G., & Stiefel, L. (2021). The effects of special education on the academic performance of students with learning disabilities. *Journal of Policy Analysis and Management*, *40*(2), 480–520. https://doi.org/10.1002/pam.22282

Shifrer, D. (2013). Stigma of a label: Educational expectations for high school students labeled with learning disabilities. *Journal of Health and Social Behavior*, *54*(4), 462–480. https://doi.org/10.1177/0022146513503346

Shin, N. (2018). The effects of the initial English language learner classification on students' later academic outcomes. *Educational Evaluation and Policy Analysis*, *40*(2), 175–195. https://doi.org/10.3102/0162373717737378

Sullivan, A. L., & Bal, A. (2013). Disproportionality in special education: Effects of individual and school variables on disability risk. *Exceptional Children*, *79*(4), 475–494. https://doi.org/10.1177/001440291307900406

Thistlewaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex-post facto experiment. *Journal of Educational Psychology*, *51*, 309–317.

Thurlow, M. L., Quenemoen, R. F., & Lazarus, S. S. (2006). Meeting the needs of special education students: Recommendations for the race to the top consortia and states. *College and Research Libraries News*, *67*(9), 548–550.

Umansky, I. M. (2016a). Leveled and exclusionary tracking: English learners' access to academic content in middle school. *American Educational Research Journal*, *53*(6), 1792–1833. https://doi.org/10.3102/0002831216675404

Umansky, I. M. (2016b). To be or not to be EL: An examination of the impact of classifying students as English learners. *Educational Evaluation and Policy Analysis*, *38*(4), 714–737. https://doi.org/10.3102/0162373716664802

Umansky, I. M., Thompson, K. D., & Díaz, G. (2017). Using an ever-English learner framework to examine disproportionality in special education. *Exceptional Children*, *84*(1), 76–96. https://doi.org/10.1177/0014402917707470

Umansky, I. M., Vasquez Cano, M., & Porter, L. M. (2021). *Resource for self-determination or perpetuation of linguistic imposition: Examining the impact of English learner classification among Alaska Native students* (No. 21-420; EdWorkingPaper). Annenberg Institute at Brown University.

U.S. Department of Education, National Center for Education Statistics. (2013). Table 204.30: Children 3 to 21 years old served under Individuals with Disabilities Education Act (IDEA), Part B, by disability: Selected years, 1976-77 through 2011-12. In U.S. Department of Education, National Center for Education Statistics (Ed.), *Digest of Education Statistics* (2013 ed.).

U.S. Department of Education. (2016). *Non-regulatory guidance: English learners and Title III of the Elementary and Secondary Education Act (ESEA), as amended by the Every Student Succeeds Act (ESSA)*.

U.S. Department of Education. (2017). *Our nation's English learners: What are their characteristics?* https://www2.ed.gov/datastory/el-characteristics/index.html

Wang, M., & Lam, Y. (2017). Evidence-based practice in special education and cultural adaptations: Challenges and implications for research. *Research and Practice for Persons With Severe Disabilities*, *42*(1), 53–61.

Authors

MARK MURPHY is an assistant professor at the University of Hawai'i at Mānoa's College of Education, 1776 University Avenue, Wist Hall 214, Honolulu, HI 96822; mmurph@hawaii.edu. He uses descriptive and quasi-experimental approaches to study issues of educational equity, with particular emphasis on multilingual students, enrollment patterns, and immigration.

ANGELA JOHNSON is a research scientist at NWEA, 121 NW Everett Street, Portland, OR 97209; angela.johnson@nwea.org. She applies quantitative and mixed methods to research focused on multilingual students and students with disabilities.

Manuscript received July 10, 2020
First revision received July 1, 2021
Second revision received June 13, 2022
Accepted July 29, 2022