Journal of Public Economics 107 (2013) 63–78

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Impact of bilingual education programs on limited English proficient students and their peers: Regression discontinuity evidence from Texas ☆

Aimee Chin [a], N. Meltem Daysal [b], Scott A. Imberman [c]

[a] Department of Economics, University of Houston, 204 McElhinney Hall, Houston, TX 77204-5019, United States
[b] Department of Business and Economics, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark
[c] Department of Economics, Michigan State University, Marshall-Adams Hall, 486 W Circle Dr Rm 110, East Lansing, MI 48824, United States

ARTICLE INFO

Article history:
Received 28 November 2012
Received in revised form 12 August 2013
Accepted 13 August 2013
Available online 30 August 2013

Keywords:
Bilingual education
Peer effects
Limited English proficient (LEP) students

ABSTRACT

Texas requires a school district to offer bilingual education when its enrollment of limited English proficient (LEP) students in a particular elementary grade and language is twenty or higher. Using school panel data, we find a significant increase in the probability that a district provides bilingual education above this 20-student cutoff. Using this discontinuity as an instrument for district bilingual education provision, we find that providing bilingual education programs (relative to providing only English as a Second Language programs) does not significantly impact the standardized test scores of students with Spanish as their home language (comprised primarily of ever-LEP students). However, we find significant positive impacts on non-LEP students' achievement, which indicates that education programs for LEP students have spillover effects to non-LEP students.

1. Introduction

One of the major challenges facing educators and policymakers today is the large and growing number of limited English proficient (LEP) children in U.S. public schools. About 1 in 9 students enrolled in pre-kindergarten to grade 12 were classified as LEP in 2008–09, a marked increase from the ratio of 1 in 13 recorded one decade earlier (National Clearinghouse for English Language Acquisition, 2011). These LEP students are present not only in big cities and other traditional immigrant-receiving areas, but also across the country; even by 2001–02, when U.S. immigrants were less geographically dispersed than they are today, about half of public schools in the U.S. had at least one LEP student (Zehler et al., 2003). Lack of proficiency in English presents a significant barrier to learning in U.S. schools, and given these recent trends in LEP student population and geographic dispersion, how to educate LEP students is likely to remain an important policy issue in the coming years.

School districts are required by federal law to provide special assistance to LEP students.[1] They typically offer Bilingual Education (BE) or English as a Second Language (ESL) to help LEP students. While there is considerable variation in how these programs are implemented in the classroom, a defining feature of BE is the use of the student's native language for at least some of the academic instruction; other programs such as ESL teach only in English. Given this feature, LEP students participating in BE tend to be placed in a self-contained classroom with classmates who share the same home language and a dedicated bilingual education teacher who can teach in that language. In contrast, LEP students participating in ESL tend to be placed in mainstream classrooms with pullout time with an ESL teacher to improve their English skills.

In this paper, we identify the causal effect of BE on the academic achievement of LEP students and their non-LEP peers using quasi-experimental variation in BE exposure generated by a policy rule governing the provision of bilingual education programs in Texas. The policy rule requires a school district to offer BE when its enrollment of LEP students in a particular elementary grade level and language is twenty or higher. Below this 20-student cutoff, districts are free to offer BE or ESL, with most choosing to offer only ESL.[2] This suggests a regression discontinuity (RD) design in which the effect of providing BE (relative to ESL) on student achievement can be obtained by comparing student outcomes in districts just above the 20-student cutoff (and therefore more likely to provide BE) and student outcomes in districts just below the cutoff. We elaborate on this RD strategy in Section 5.

This paper adds to a large literature evaluating educational programs for LEP students, which we briefly summarize in Section 2.2. It addresses two major gaps in this literature. First, this literature has focused exclusively on the impacts on the intended beneficiaries themselves (i.e., the

☆ We thank Erzo F.P. Luttmer, Chinhui Juhn, Danielle Li, and workshop participants at the 2010 and 2011 National Academy of Education Annual Meeting, 2011 ASSA Meetings, 2012 SOLE Meetings, and 2012 APPAM Fall Research Conference for helpful comments and discussion. Imberman gratefully acknowledges financial support from the National Academy of Education/Spencer Foundation Postdoctoral Fellowship. A University of Houston Small Grant was used to purchase the data, for which we are grateful. The authors bear sole responsibility for the content of this paper.

E-mail addresses: achin@uh.edu (A. Chin), meltem.daysal@sam.sdu.dk (N.M. Daysal), imberman@msu.edu (S.A. Imberman).

[1] The relevant laws are Title VI of the 1964 Civil Rights Act and the Equal Educational Opportunity Act of 1974. Section 2.1 provides a legislative background on educational programs for LEP students.

[2] There are cost advantages to offering only ESL, as we discuss in Section 5.1.

0047-2727/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jpubeco.2013.08.008

LEP students) and ignored any effects that these programs might have on non-LEP students. Yet, because these programs change the student composition of mainstream classrooms and school budgets, among other things, there is potential for spillover effects to non-LEP students. To our knowledge, our study is the first to test for spillover effects of educational programs for LEP students, and to the extent that they exist, to quantify them. Quantifying these spillover effects is necessary for a complete cost–benefit analysis of the various LEP programs; *all else equal*, policy makers might prefer the program that benefits non-LEP students more (or, stated differently, harms non-LEP students less).

Second, most of the studies in this literature do not address the potential problem of endogeneity in student exposure to the educational programs for LEP students. In general, student exposure to a program is not random, and instead is the result of decisions made by students, parents, schools and districts. Thus, it is likely correlated with unmeasured and unobserved characteristics of the students, parents, schools and districts, some of which might in turn be correlated with student achievement. Estimates of program effects that do not take this into account tend to be biased. Our research adds to the handful of studies that provide estimates of the impacts of LEP programs with a causal interpretation.[3]

We implement our RD strategy using panel data on elementary schools in districts near the 20-Spanish-LEP-student cutoff defined by the policy rule. We describe these data in Section 6. We restrict our attention to the policy rule vis-à-vis *Spanish* LEP students for a practical reason: Spanish is the home language of 90% of Texas' total LEP enrollment, and is the only language for which there is enough variation across districts to implement our empirical strategy. Due to this restriction, our results pertain to the effect of district provision of Spanish bilingual education programs (relative to providing only ESL for Spanish LEP students). However, considering that Spanish is the language of over three-quarters of total LEP enrollment in the U.S. and accounts for an even higher share of bilingual education programs operating in the U.S. (Zehler et al., 2003), it is especially policy relevant to understand the effects of Spanish BE programs.

To preview the results of Section 5, we find a significant increase in the probability that a district provides BE above the 20-Spanish-LEP-student cutoff. We do not find any significant jumps at the cutoff in covariates unrelated to BE provision, nor do we observe "stacking" of districts below the cutoff, which validates the interpretation of differences in student outcomes just above and just below the cutoff as due to district BE provision.

We proceed by using the variation in district provision of BE induced by the policy rule as an instrumental variable to identify the causal impact of district provision of BE on student achievement. While the impact of *school* provision of BE would also be of interest, we focus on *district* provision of BE because this is directly linked to the policy rule. These instrumental variable estimates provide the local average treatment effect of district provision of Spanish BE among districts whose decision to offer Spanish BE is constrained by the policy rule. Our main findings are as follows. First, district provision of BE raises the standardized math and reading test scores of students who are non-LEP and whose home language is not Spanish. In our preferred specification controlling for a linear spline of the running variable, the district-wide Spanish LEP count in a student's first grade cohort, the positive impacts on non-LEP achievement are statistically significant. Students who are non-LEP and whose home language is not Spanish would never have been candidates to participate in Spanish BE programs, thus this finding is indicative of spillover effects.

Second, district provision of BE has generally positive but smaller and statistically insignificant effects on students whose home language is Spanish. A vast majority of Spanish home language students (89%) are classified as LEP in first grade, and so would have been eligible to participate in educational programs for Spanish LEP students. Hence, our results suggest that the intended beneficiaries of the LEP programs fare similarly in BE and ESL programs. Finally, we find that district BE provision increases test scores on all students taken together. The positive net impact indicates that on average, the test score gains due to district BE provision exceed test score losses.

[3] These studies include Slavin et al. (2011), Matsudaira (2005) and Angrist, Chin and Godoy (2008), which we summarize in Section 2.2.

2. Background and related literature

2.1. Legislative background on educational programs for LEP students

The Bilingual Education Act passed in 1968 was the first federal law expressly addressing the educational needs of LEP students in American schools, and did so by providing a financial reward—federal grants awarded on a competitive basis—for providing help to LEP students.[4] Later federal laws made it a legal responsibility of school districts to provide such help. The two main pieces of legislation are the Title VI of the 1964 Civil Rights Act, which prohibited discrimination on the basis of race, color or national origin in federally-assisted programs, and the Equal Educational Opportunity Act of 1974, which basically extended Title VI to school districts not receiving federal funds. School districts faced termination of funding from the U.S. Department of Education or private lawsuits if they failed to provide LEP children with an equal educational opportunity.

The U.S. Supreme Court's *Lau v. Nichols* decision in 1974 made clear that the prevalent practice of "sink or swim" instruction, in which LEP students are placed in the same classrooms as non-LEP students without additional services, was a violation of LEP students' civil rights. To receive an equal educational opportunity, LEP students were entitled to special assistance. Guidelines on Title VI compliance issued by the Department of Education to school districts in 1970 called for "affirmative steps"[5] to help LEP students without specifying what educational programs to use, and new guidelines in 1975 specified bilingual education.

In an environment that demanded Title VI compliance, individual states passed laws mandating bilingual education programs for LEP students. Massachusetts was the first, with a 1971 law, followed by Alaska (1972), California (1972), Illinois (1973), Texas (1973), New York (1974), and others. The programs mandated by these laws tended to be *transitional* BE programs, which have the goal of mainstreaming the LEP students as soon as they acquire sufficient English-language skills; native language instruction is temporarily used so that the LEP students can keep up in academic subjects.[6] Additionally, these laws did not require *every* school district to provide bilingual education to *every* LEP student. Instead they specified the circumstances under which a school district would provide BE, and these circumstances generally involved the number and concentration of LEP students of a particular grade and of a particular language group in a school district. Below, we take advantage of the specifics of the Texas law to identify the effects of bilingual education.

Since the late 1990s, there has been a shift away from using bilingual education toward using English-only programs to help LEP students. Revealingly, in 2002, the Bilingual Education Act was renamed the English Language Acquisition, Language Enhancement, and Academic Achievement Act. Also, several states eliminated bilingual education in public schools through ballot initiatives: California (1998), Arizona (2000) and Massachusetts (2002). Reflecting this policy shift, 40% of LEP students in U.S. public schools were in a BE program in 2001–02, compared to 63% in 1991–92. The rest are in English-only programs, with the largest being ESL (Zehler et al., 2003).

[4] This legislation was Title VII of the Elementary and Secondary Education Act. Crawford (1989) provides a good history of BE in the U.S., and Nieto (2009) provides a more recent summary.

[5] 35 Fed. Reg. 11,595 (1970) as cited in U.S. Commission on Civil Rights (1975) , p. 71.

[6] Dual language programs that mix LEP and non-LEP students with the goal of proficiency in both English and another language (the LEP students' home language, which is a foreign language for non-LEP students) are rare and not the subject of this study.

To comply with U.S. civil rights laws, school districts must provide special assistance to LEP students, but neither history nor existing evaluations (which we discuss below) provide clear guidance for which programs are more effective. This study contributes to the debate on how to help LEP students by providing new empirical evidence on the effects of bilingual education program provision (compared to ESL program provision alone); it is at this very margin—BE or ESL—that many school districts are making policy decisions today.

2.2. Related literature

There is a substantial body of research evaluating the effect of educational programs for LEP students on the LEP students themselves.[7] Most of these papers do not address the potential problem of endogeneity in program exposure. A few recent studies use research designs that are more convincing for identifying causal relationships. Slavin et al. (2011) conduct a randomized evaluation in which LEP kindergartners in six schools were randomly assigned within school to either bilingual education or structured English immersion (an English-only approach). They find no statistically significant differences in tests measuring English skills by fourth grade (sample size was 92 students). Matsudaira (2005) uses a regression-discontinuity design taking advantage of a district policy rule governing which students are classified as LEP and therefore eligible to participate in educational programs for LEP students (which in this district is BE or ESL). He finds little difference in academic achievement between students who scored just below the English skills assessment cutoff (and are eligible for BE/ESL services) and just above (ineligible and placed in mainstream classrooms). Angrist, Chin, and Godoy (2008) look at policy shifts in Puerto Rican schools that changed the medium of instruction from English to Spanish and find no impact on English-speaking ability in adulthood. These three studies suggest that, while not helping LEP students, pedagogical approaches using native language instruction do not seem to hurt them either. Our study contributes by exploiting a different source of exogenous variation in exposure to educational programs for LEP students.

A notable gap in the literature evaluating LEP programs is that it ignores potential impacts of these programs on non-LEP students. To our knowledge, our study is the only one to examine potential effects of educational programs for LEP students on non-LEP peers. Closely related, though, are Cho (2011) and Geay, McNally and Telhaj (2012), which estimate the impact of LEP students on non-LEP peers. These studies are pertinent to ours because, as we discuss below, one channel through which BE provision may impact non-LEP students is through decreasing the exposure of non-LEP students to LEP students. Cho uses within-school, and sometimes within-student, variation in having an LEP classmate in the Early Childhood Longitudinal Study, Kindergarten Cohort dataset, and finds that having at least one LEP classmate is associated with lower reading test score gains but no significant differences in math test score gains. Geay et al. use variation in share of classmates who are non-native English speakers arising from within-school cohort-to-cohort fluctuations, as well as from European Union enlargement, and find no significant effects of exposure to non-native English speakers on the achievement of native English speakers in England.

Although few studies examine peer effects of LEP students on non-LEP students specifically, there exists a large literature on peer effects in education. Two strands within this literature are especially related to our study. One is on peer effects in primary and secondary education, including on the effect of being exposed to peers who have higher or lower achievement test scores (e.g., Angrist and Lang, 2004; Hanushek, Kain, Markman and Rivkin, 2003; Hoxby and Weingarth, 2006; Lavy, Paserman, and Schlosser, 2012; Imberman, Kugler, and Sacerdote, 2012) or who exhibit disruptive behavior (e.g., Figlio, 2005; Aizer, 2008; Carrell and Hoekstra, 2010). A second strand is on the impact of immigrants on the educational outcomes of natives (e.g., Betts, 1998; Hoxby, 1998; Liu, 2000; Neymotin, 2009). While many immigrants are LEP when they first enter school in the U.S., they exit LEP status with more time spent in the U.S., and moreover it should be noted that only half of LEP students enrolled in U.S. public schools are foreign-born, thus the impact of *LEP* students may well differ from the impact of *immigrant* students. We contribute to this broader literature on peer effects in education by examining a new source of variation in peer composition: exposure to LEP students induced by Texas' policy rule about bilingual education provision.

3. Empirical strategy

3.1. Conceptual framework

The direction of the impact of BE programs (compared to ESL programs) on academic achievement is theoretically ambiguous for both LEP and non-LEP students. LEP students are the students eligible to participate in BE and ESL. BE and ESL are the two most common programs offered by schools to address the learning needs of LEP students, and each has advantages and disadvantages. For example, native language instruction might delay English acquisition, but it might also enable LEP students to better keep up in math and other academic subjects while they are learning English. On the other hand, an English-only approach like ESL might improve learning in both English and other subjects by reallocating time that would have been spent on teaching LEP students the ability to read and write in their native language.

Non-LEP students do not participate in BE or ESL themselves, but they may experience spillover effects from these programs. The nature of these spillover effects likely differs by program. One reason is that the two programs result in dramatically different distributions of LEP students across classes within a grade. Mainstream classes have fewer LEP students when BE is offered because LEP students in BE tend to be grouped together to form a separate class while LEP students in ESL are in the same classes as non-LEP students (with ESL instruction provided on a pull-out basis, or with an in-class ESL aide). Exposure to LEP students in class could impact non-LEP students' academic performance through various mechanisms. For example, teachers with LEP students in their classes may need to provide extra assistance to these students, which would take time away from other students. Also, LEP students may be more prone to disruption due to frustration from difficulties understanding the material taught in English. Additionally, to the extent that achievement among LEP students is lower than the non-LEP students, there could be an achievement peer effect which could worsen non-LEP outcomes. Nonetheless, LEP students could generate positive peer effects because although they are less advanced (at least temporarily) along the English skills dimension, they could be more advanced along other dimensions (e.g., knowledge in other subjects, non-cognitive skills) that matter for student achievement or there could be benefits from having a more diverse classroom.

A second reason that BE and ESL could generate differential spillover effects is through the school budget. Whether BE or ESL is more costly for the school district depends on a number of factors, including the number of LEP students and their distribution across languages and grades. When the school district has few LEP students in the same language and grade, BE programs tend to be more expensive than ESL programs. Since a BE teacher is typically attached to a specific class on a full-time basis (serving LEP students of a common home language and grade), there is little possibility for schools to spread the cost of a BE teacher over LEP students of different home languages and grades as they could with an ESL teacher. Thus, to pay for BE, schools may have to reallocate resources, and this may impact non-LEP students' academic performance. While a simple story of BE programs crowding out programs for non-LEP students might suggest negative effects, schools may

[7] See, for example, Baker and de Kanter (1981), Willig (1985), Rossell and Baker (1996) and Greene (1998) for reviews. Slavin et al. (2011) offer a more recent discussion of this literature.

offset the higher costs with additional revenues, so these resource effects are ambiguous in direction.

3.2. Identification strategy

Given these theoretical considerations, how bilingual education provision affects academic achievement is ultimately an empirical question. Thus we turn to our estimation strategy. We wish to estimate the effect of exposure to bilingual education programs on achievement, which might be approximated as:

$$y_{idcg} = \alpha + \beta\, BE_{dc} + X_{idcg}\pi + \varepsilon_{idcg} \quad (1)$$

for student i in school district d who is a member of first grade cohort c and observed at grade g. y is an achievement measure, BE is an indicator for the student's school district providing a bilingual education program, X is a set of student, school and district characteristics, and ε is the error term.

The parameter of interest is β. The key impediment to interpreting the ordinary least squares (OLS) estimate of β as the causal effect of student exposure to BE stems from the potential endogeneity of district BE provision in Eq. 1. It is not random which school districts provide BE. For example, they may be the ones with more LEP students, more growth in LEP students, more wealth, LEP students with especially low English proficiency, and so on. Given these differences, they would likely have had different student outcomes regardless of BE provision. A priori we cannot even sign the direction of the bias. To address this endogeneity problem, we use a regression-discontinuity approach that exploits a policy rule governing the provision of BE in Texas. The State of Texas mandates provision of BE in a given language and elementary grade by a school district when the district-wide population of LEP students in that language and elementary grade is greater than or equal to 20.[8]

Essentially, our empirical strategy is to compare student outcomes in districts that have slightly less than 20 LEP students in a language-grade to those with slightly more. In practice, we focus only on the policy rule vis-à-vis Spanish LEP students, who represent about 90% of Texas' total LEP enrollment, because Spanish is the only language group for which there is enough variation across districts to implement our empirical strategy.[9] It is unlikely that districts with 19 Spanish LEP students differ that much from districts with 20 Spanish LEP students, but due to the policy rule, the latter districts must offer BE, and we can take the difference in outcomes between the districts with 20 and districts with 19 to learn about the effect of district BE provision. This example is meant to be illustrative only, as limiting our analysis to only those districts with exactly 19 or 20 Spanish LEP students would lead to very imprecise estimates. In implementing our empirical strategy therefore, we expand the neighborhood around the cutoff. With the wider bandwidth, it becomes possible that there exists a relationship between the number of Spanish LEP students and the outcome that is not solely due to the policy rule, so it is necessary to control for the number of Spanish LEP students. We therefore estimate the following eq. to study the impact of the policy rule on district BE provision:

$$BE_{dc} = \alpha^{FS} + \delta^{FS} Above20_{dc} + f(LEP\_Count_{dc}) + X^{FS}{}_{idcg}\pi^{FS} + \varepsilon^{FS}{}_{idcg} \quad (2)$$

for student i in school district d who is a member of first grade cohort c and observed at grade g. BE is an indicator for the student's school district offering a bilingual education program, LEP_Count is the district-wide Spanish LEP student count for student i's first grade cohort, $Above20$ is an indicator for LEP_Count being greater than or equal to 20, $f(LEP\_Count)$ is a continuous function of LEP count,[10] and X^{FS} is a set of student, school and district characteristics.[11]

Below, we find that δ^{FS} is positive and statistically significant, indicating that the policy rule induced some districts that otherwise would not have provided BE to provide BE. We proceed by instrumenting the potentially endogenous regressor in Eq. 1, BE, with $Above20$ (where X in Eq. 1 is comprised of X^{FS} and $f(LEP\_Count)$) in order to obtain an estimate of β with a causal interpretation. This strategy is often referred to as a fuzzy regression discontinuity design (Imbens and Lemieux, 2008; Lee and Lemieux, 2010). The first-stage equation associated with the 2SLS estimation of Eq. 1 is given by Eq. 2.

Below we also report the results of estimating the reduced-form equation,

$$y_{idcg} = \alpha^{RF} + \delta^{RF} Above20_{dc} + f(LEP\_Count_{dc}) + X^{FS}{}_{idcg}\pi^{RF} + \varepsilon^{RF}{}_{idcg}. \quad (3)$$

The reduced-form effect of being just above the 20-student cutoff, δ^{RF}, indicates the effect of increasing the likelihood of a school district offering BE on achievement. It is desirable to rescale this reduced-form effect to obtain the interpretation of the effect of school district provision of BE, which is what is given by the 2SLS estimate of β.

Note that our regression models measure the provision of BE at the district-cohort level, not at the individual or school-cohort level. Focusing on an aggregate measure of BE provision is natural when we use data on non-LEP students, who are never participants in BE programs but nevertheless could experience spillover effects from them. LEP students, on the other hand, are eligible to participate in BE programs if their school district offers it to their cohort but may choose not to take it up; thus, by using this district-cohort level measure of BE availability, we are capturing an intention-to-treat effect for LEP students rather than the effect of participation in BE. Focusing on potential exposure to BE, rather than actual take-up of BE, circumvents issues concerning non-random selection of individuals into BE programs. Moreover, the intention-to-treat effect is of direct interest for policy making, as school districts can only control whether to offer BE—students cannot be forced to participate in BE.

Our choice to focus on the provision of BE at the *district*-cohort level instead of the *school*-cohort level is motivated by the nature of the variation in BE exposure that the Texas policy rule provides. When the district provides BE in order to comply with the policy rule, it typically provides BE in one, not all, of the elementary schools within the district. If district BE provision affects student achievements in other ways than just through a student's own school providing BE (e.g., through changing the distribution of students and resources among schools in the same district), the policy rule would not be a valid instrument for *school* BE.[12]

The fuzzy RD strategy identifies the local average treatment effect (LATE) for school districts close to the 20-student cutoff. These school districts tend to be smaller, be less urban and (of course) have fewer LEP students than the average district. Thus, the effect of district BE provision on achievement estimated in this study may not reflect the average treatment effect or generalize to larger districts. However, since few

[8] This rule is part of Texas Administrative Code, Title 19, Part 2, Chapter 89, Subchapter BB, Rule §89.1205 (Commissioner's Rules Concerning State Plan for Educating Limited English Proficient Students).

[9] In contrast, LEP students with other home languages are both fewer and more concentrated, leaving too few observations of districts near the 20-student cutoff for a particular language and grade. The Spanish share of LEP students in 2001–02 is 77% in the U.S. (Zehler et al., 2003); given Texas' proximity to Mexico and Central America, it is not surprising that its Spanish share is somewhat higher.

[10] Our main results use a linear function that allows for different slopes above and below the cutoff.

[11] The results are similar whether we include or exclude the additional control variables (see Section 5.6), however our preferred specification includes them to improve the precision of the estimates.

[12] In practice, over 60% of the schools in our sample are in districts with only one elementary school with most of the others having two, so the distinction between using "district BE provision" and "school BE provision" is not large. In Section 5.5, we find that results using only schools in single elementary school districts are qualitatively similar, and somewhat stronger in terms of magnitudes and significance, relative to results using the full sample.

studies exist that convincingly identify the causal effects of educational programs for LEP students (and none of these consider spillover effects to non-LEP students), our new evidence is of interest even if it is estimating a LATE. Additionally, we believe this LATE is per se interesting because a majority of school districts in the U.S. with LEP students have relatively small LEP enrollments. In 2001–02, while 43% of school districts in the U.S. had at least one LEP student (Zehler et al., 2003), only 2.6% of school districts with LEP students enrolled 5000 or more LEP students 61% enrolled 1-99 LEP students. Thus, the vast majority of school districts with LEPs are contending with some, but not many, LEP students. As immigrants increasingly settle outside of traditional immigrant-receiving places, the number of these low-LEP-enrollment school districts will continue to grow.

4. Data

To implement our RD strategy, we use publicly-available data on the standardized test scores and demographic characteristics of students enrolled in Texas public elementary schools.[13] To maintain data confidentiality, the Texas Education Agency provided us with *grouped* student data rather than individual-level student data. In particular, we obtained mean data at the school-grade-year level for three mutually exclusive categories of students: (1) students who are not classified as LEP and do not have Spanish as their home language (below, we refer to them as the "non-LEP, non-Spanish home language" students, and we refer to the dataset containing these students' test scores as the "non-LEP, non-Spanish home language" sample); (2) students who have Spanish as their home language (analogously, the "Spanish home language" students and sample)[14]; and (3) remaining students. Below, we estimate Eqs. 1–3 separately for the first two categories of students, as well as for all students combined,[15] where *i* in these equations now indexes *school* rather than the *individual student*. Because we wish to quantify the impact of a policy mandating districts to offer BE on the average *school*, given that our observations are at the school-grade-year level, we use OLS rather than weighted estimation.[16] We note that the use of grouped student data, instead of individual-level student data, is appropriate in our context because the policy is based on an aggregate rather than an individual-level rule. In addition, our student category-school-grade-year level data are at a less aggregate level than the policy, which applies to the district-grade-year level.[17] Given the level of the policy, we use standard errors clustered by district for inference (Bertrand, Duflo and Mullainathan, 2004; Angrist and Pischke, 2009).

The "non-LEP, non-Spanish home language" sample is comprised of students who would never have been in educational programs for Spanish LEP students, so effects of district BE provision estimated using this sample only reflect spillover effects of Spanish BE programs. The "Spanish home language" sample is comprised of students who are currently LEP, formerly LEP (but since mainstreamed) and never LEP, so in theory, effects estimated using this sample reflect both the effect of Spanish BE programs on the individuals who were eligible to participate in Spanish LEP programs (i.e., the intended beneficiaries) as well as spillover effects. In practice though, 89% of "Spanish home language" students are classified as LEP in first grade in our analysis sample,[18] so results using data on these students can be interpreted as the effects of BE provision on the intended beneficiaries. Finally, the results using data on all students reflect the net effect of the Spanish BE programs.

To assess whether the policy rule is binding, we must examine whether the probability of a district providing BE increases at the 20-student cutoff. In our empirical analysis below, we use the district counts of Spanish LEP students from first grade as the relevant counts for determining district BE provision. Several points are worth noting. First, a district's count of Spanish LEP students for a given first-grade cohort is the sum of the Spanish LEP student count for that cohort across all schools within that district. Second, in our data, the Spanish LEP student count for a particular school-year-grade cell is masked if the number is between 1 and 4, so for the purpose of obtaining the district count of Spanish LEP students, we assigned the average value of 2.5 to those school-grade-year cells with masked values.[19] Third, LEP status is temporary, with LEP students exiting LEP status once they have learned English.[20] Consider a district that has 20 Spanish LEP students in a first grade class and hence offers BE. By the time these students reach third grade, there will likely be fewer than 20 Spanish LEP students in the cohort due to mainstreaming of initially LEP students who become proficient in English. Nonetheless, even though the LEP count falls, the district typically continues providing BE to this cohort for several years.[21] Thus, it is appropriate to use the district Spanish LEP count in the first grade class for a student's cohort rather than concurrent LEP counts.

Because the policy rule specifies a 20-student cutoff, we restrict the data to districts near this cutoff. Our main analysis uses observations in districts with 8 to 39 Spanish LEP students in a given first-grade cohort.[22] We further restrict the data to smaller districts, which we define as districts with total first-grade enrollment below 200 in the 2004–05 school year, in order to form a consistent sample of districts for which the policy rule is likely to bind.[23] Fig. 1 maps the districts satisfying our sample criteria using the average Spanish LEP counts in the district during the sample period, and indicates that the districts in our sample are located all over Texas, and that districts above the cutoff are often located next to districts below the cutoff. This provides some reassurance that on the basis of geographic location, the districts just above and just below the cutoff are comparable. We provide additional analysis on the comparability of districts above and below the cutoff in Section 5.3.

Our student outcome measures are standardized test performance for 2002–03 through 2009–10, aggregated to the school-grade-year for each of our three student categories. The Texas Assessment of Knowledge and

[13] We obtained the test score data through a public information request to the Texas Education Agency.

[14] The "Spanish home language" designation is based on a home language survey that parents fill out when their child first enrolls in a district. This is not the same as Hispanic status, as there are many Hispanics who do not use Spanish at home.

[15] Mean data for all students combined for a given school-grade-year is the weighted average across the three categories (weighted by the number of students in each category). For districts near the 20-Spanish-LEP-student cutoff, the third category (LEP students with a language other than Spanish as the home language) is very small—accounting for less than five percent of enrollment—thus we do not consider this category separately.

[16] If one wished to estimate the impact on the average *student*, one would use weighted least squares, weighting by enrollment (this is equivalent to performing OLS using individual student level data).

[17] Although we described the policy variation as occurring at the level of district-cohort in Section 5.2, this statement is also accurate because year and grade data are used to construct cohort data.

[18] In the full sample, without the restriction to small districts near the 20-Spanish-LEP-student cutoff, 78% of Spanish home language students are classified as LEP in first grade.

[19] Our results are robust to alternative ways of treating the masked values in forming the district count of Spanish LEP students variable, as we show in Section 5.6. For the student achievement measures, too, the true value for a given school-year-grade cell is masked when there are fewer than 5 students in that cell. We treat masked achievement data as missing achievement data, and drop the associated observations from the analysis.

[20] Due to mainstreaming, LEP status drops off considerably as students age. Using data on LEP counts by grade and district acquired from the Texas Education Agency, we observe that LEP share of enrollment in Texas fall from 11% in kindergarten and 1st grade to 4% by 8th grade.

[21] Texas supports a transitional BE program where LEP students are moved to mainstream classes once they acquire sufficient English skills. Typically students stay in BE several years before attaining the English skills to be mainstreamed, and so the district typically commits to providing BE to this cohort for this duration.

[22] We show in Section 5.6 that our results are robust to using other bandwidth selections.

[23] Larger school districts have more resources, and tend to offer bilingual education services even below the cutoff, thus they are less responsive to the policy. Regarding the use of 2004–05 enrollments to define district size, in fact we have used first grade enrollment totals from other years to define the set of districts to be included, and our empirical results are unchanged, as we show in Section 5.6. We do not impose the 200 cap on every year for sample inclusion because changes in school enrollment can be endogenous to student achievement.

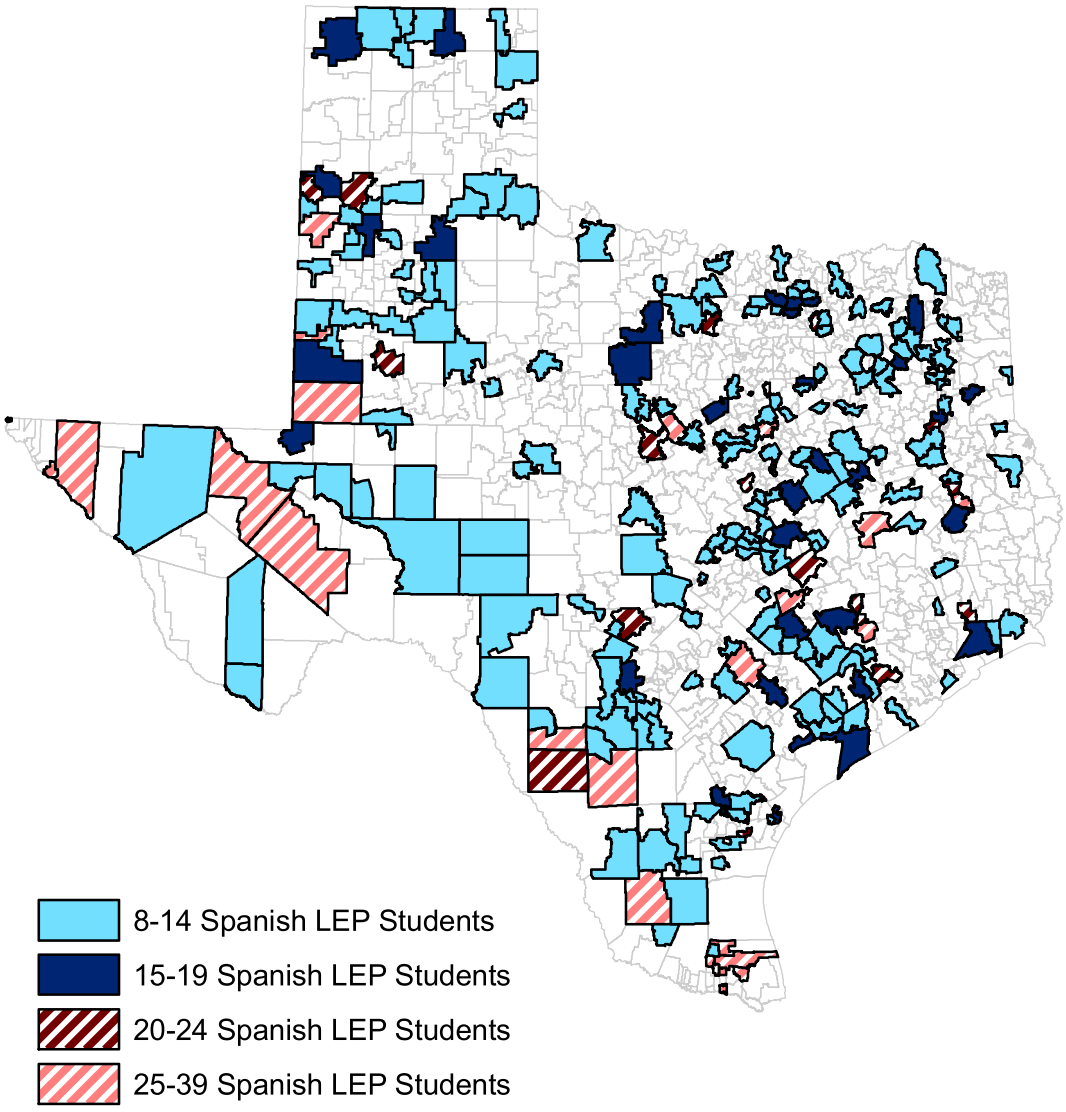

Fig. 1. Texas school districts with 8–39 Spanish LEP students in the first grade cohort. The shaded area corresponds to school districts with fewer than 200 students in the first grade cohort in 2004–05, and with between 8 and 39 Spanish LEP students in the first grade cohort on average between 2002–03 and 2009–10; this shaded area corresponds to the 261 districts present in the "non-LEP, non-Spanish language" sample described in Table 1.

Skills (TAKS) math and reading tests were introduced in 2002–03 (replacing the Texas Assessment of Academic Skills), and to avoid combining student achievement measures based on different tests, we do not use earlier test score data.[24] Students are tested from third grade onwards, and our policy rule concerns BE provision in elementary schools. Thus, we have three grades with test score outcomes: third, fourth and fifth. Our main analysis pools students in grades 3 to 5.[25] We link test takers to their first grade cohort's district-wide number of Spanish LEP students and district-wide provision of bilingual education, so we are using demographic data for 1998–99 to 2007–08.[26]

We use three measures of test performance for each subject, with the intention of capturing effects at different parts of the student academic achievement distribution. The mean standardized scale score, which is the average test score normalized to standard deviation units using the state-wide mean and standard deviation, captures movements in test scores from all parts of the distribution. The passing rate, which is the percent of students that met the minimum passing standard set by the State Board of Education, captures movements in test scores for students at the margin of passing/failing. The commended rate, which is the percent of students that met a much higher passing standard (on average, only the top third of passers satisfy this), captures movements in test scores for higher-achieving students. Table 1 provides the means and standard deviations of the variables for the "non-LEP, non-Spanish home language" sample overall and separately for those above and below the cutoff.[27]

[24] One potential critique of examining math and reading outcomes separately is the possibility of multiple hypothesis tests indicating one of the two indicators to be significant at the 5 level when in fact there is no true effect. In order to address such a concern, we estimated our models using the average of math and reading scores as a single outcome and find similar results.

[25] Our findings are similar when we estimate separate models by grade (see Appendix Tables 1–4).

[26] Thus, for the third, fourth and fifth graders observed at year t, we assign the first grade district-wide Spanish LEP count from year $t-2$, $t-3$ and $t-4$, respectively. For example, for fifth grade test takers in 2002–03, first grade cohort characteristics are taken from 1998–99.

[27] Appendix Tables 5 and 6 show the summary statistics for the "Spanish home language" and "all students" samples, respectively. The school-grade characteristics are basically the same across the three samples (which is to be expected, as the three samples cover different categories of students who are attending the same schools), however the test performance measures are lower on average among "Spanish home language" students.

Table 1
Summary statistics for "Non-LEP, Non-Spanish Home Language" sample, 2002–03 through 2009–10.

	Full sample (8 to 39 LEP in 1st grade cohort)	8 to 19 LEP in 1st grade cohort	20 to 39 LEP in 1st grade cohort
Mean (standard deviation) among all students in same school-grade			
% female	48.6	48.5	48.8
	(6.0)	(6.1)	(5.6)
% White	43.7	47.8	31.2
	(24.5)	(23.6)	(23.1)
% Hispanic	48.0	43.6	61.5
	(27.0)	(25.4)	(27.4)
% Black	7.8	8.0	7.0
	(10.5)	(10.8)	(9.3)
% economically disadvantaged	63.1	60.8	70.4
	(17.8)	(17.3)	(17.3)
% LEP	11.7	10.3	16.1
	(10.3)	(9.0)	(12.4)
% in bilingual program	3.5	1.3	10.2
	(8.7)	(4.8)	(13.2)
% in special education program	11.4	11.7	10.3
	(5.5)	(5.7)	(4.8)
% in gifted and talented program	7.2	7.2	7.3
	(4.9)	(5.0)	(4.6)
Mean (S.D.) among non-LEP, non-Spanish home language students in same school-grade			
TAKS math standardized scale score	−0.08	−0.07	−0.13
	(0.35)	(0.36)	(0.33)
TAKS math passing rate	80.0	80.5	78.5
	(14.0)	(13.8)	(14.4)
TAKS math commended rate	27.8	28.0	27.1
	(14.4)	(14.5)	(14.1)
TAKS reading standardized scale score	−0.03	−0.01	−0.08
	(0.32)	(0.33)	(0.30)
TAKS reading passing rate	83.0	83.5	81.4
	(12.3)	(12.3)	(12.2)
TAKS reading commended rate	28.5	28.8	27.4
	(13.9)	(14.1)	(13.2)
Observations (school-grade-year)	3761	2835	926
# of schools	413	375	148
# of districts	261	247	89

Notes: Each school-grade-year for grades 3 through 5 is a separate observation. Sample is limited to observations in districts with fewer than 200 students in the 1st grade cohort in 2004–05, in district-grade-years with between 8 and 39 LEP students in the 1st grade cohort, and with non-missing math or reading achievement variables.

5. Results

5.1. OLS estimates of the effect of district provision of bilingual education

In Table 2, we present the results from estimating Eq. 1 via OLS using the same three samples that we use below to implement our RD strategy.[28] The OLS coefficients for district BE provision are not significantly different from zero using the "non-LEP, non-Spanish home language" sample (columns 1 and 2), which if interpreted causally would indicate no spillover effects for BE programs. In columns 3 and 4, the OLS coefficients are negative, and sometimes significant at the 10% level, which if interpreted causally would indicate harmful effects of BE for Spanish home language students. The estimated effect for all students combined is shown in the last two columns; because about three quarters of all students are in the "non-LEP, non-Spanish home language" category, these estimates are close to those reported in columns 1 and 2.

These estimated coefficients for district provision of BE are unlikely to have a causal interpretation, however, for reasons discussed above. We present the OLS estimates to provide a counterpoint to the RD estimates of the effect of district provision of BE discussed below, which we find to be more positive. Comparing the RD estimates to the OLS estimates reveals that the OLS estimates in Table 2 tend to be downward biased. That is, districts that are observed to provide BE are worse in unmeasured/unobserved dimensions that positively correlate with student test performance.

5.2. The discontinuity in district provision of bilingual education

If the Texas policy rule governing BE provision is binding, then we should observe a discontinuity in district BE provision at the 20-student cutoff, with districts above the cutoff having higher rates of BE provision. To examine this, in Fig. 2, we plot the average share of schools that are in districts providing BE by the number of Spanish LEP students in the first grade cohort. Visually, there is a jump up at the 20-student cutoff. It is worth noting, though, that the jump is not from no provision at all to universal provision. On the one hand, to the left of the cutoff, some schools provide BE by their own choice. On the other hand, to the right of the cutoff, we see there is not perfect adherence to the policy rule (had this been the case, 100% of districts would provide BE above the cutoff). One reason is that participation in BE requires parental consent, and some parents choose ESL for their child even when BE is available. What is measured in the data is whether any students in the cohort are enrolled in BE; we cannot observe when a district offers BE but there is zero take-up. A second reason is that it is difficult to recruit certified bilingual education teachers, and school districts are allowed to delay providing BE if they are unable to find a qualified BE teacher. Finally, some districts are permitted to send their BE students to a neighboring district.

The first row of Table 3, Panel B provides the first-stage results corresponding to Fig. 2. The coefficient for being above the cutoff is positive and statistically significant across all samples. Districts with more than 20 Spanish LEP in the 1st grade cohort are 28 percentage points more likely to offer BE than districts below the cutoff. This is a sizable effect—a three-fold increase relative to the mean for district BE provision among observations below the cutoff in the full analysis sample (which is 9%). Clearly, the rule generates sizable variation in district provision of BE for a particular first grade cohort.

5.3. Tests of the validity of the regression discontinuity design

Because BE programs tend to be more expensive than ESL programs, and especially because school districts mandated by the Texas policy rule to provide BE also tend to provide ESL (because some parents do not wish to enroll their children in BE even when it is available), there is a financial incentive for school districts to manipulate their enrollment or LEP classifications to avoid having to provide BE. If districts have discretion over the number of students categorized as LEP, and exercise it to avoid having to provide BE, then potentially students on one side of the 20-student cutoff could be systematically different from those on the other side, invalidating the RD design. To assess this, in Fig. 3 we plot the distribution of district first grade Spanish LEP students. A discontinuity in the density of school districts around the 20-student cutoff would suggest manipulation of our running variable (McCrary, 2008). As the figure shows, there are no irregular heaps in the density of district first grade LEP counts. In particular, the number of districts declines smoothly as the number of Spanish LEP students increases. More formally, applying the test proposed by McCrary (2008), we do not find any significant change in the density at the 20-student cutoff.[29] Therefore it does not appear that districts manipulate LEP student numbers to avoid providing BE.

Next, we check whether there are differences in observable characteristics across the 20-student cutoff. Appendix Fig. 1 graphs the covariate means by number of Spanish LEP students in the first grade cohort

[28] Results based on the sample containing all schools, i.e., without the restriction to schools in small districts near the 20-Spanish-LEP-student cutoff, are similar.

[29] The coefficient for the above cutoff dummy is −0.03 (s.e. 0.14). Barreca, Lindo and Waddell (2011) show that heaping in the running variable can lead to biased estimates even if heaping occurs away from the cutoff. Fig. 3 shows that there is no evidence of heaping at any value of the running variable in our sample. We also conducted donut regressions in which we dropped LEP enrollment groups immediately around the 20-student cutoff and found results that are very similar to our main estimates.

Table 2
OLS estimates of relationship between district bilingual education provision and TAKS achievement.

	"Non-LEP, Non-Spanish Home Language" Students		"Spanish Home Language" Students		All Students	
	Math	Reading	Math	Reading	Math	Reading
	(1)	(2)	(3)	(4)	(5)	(6)
Mean standardized achievement	0.008	0.006	−0.055	−0.071*	−0.007	−0.014
	(0.029)	(0.026)	(0.042)	(0.042)	(0.032)	(0.029)
Passing rate	−0.257	−0.420	−1.677	−3.076*	−0.870	−1.254
	(1.057)	(0.850)	(1.835)	(1.741)	(1.179)	(0.973)
Commended rate	0.610	0.531	−2.015*	−1.128	0.302	0.210
	(1.012)	(0.866)	(1.125)	(1.046)	(1.035)	(0.880)
Observations	3759	3761	3247	3244	3819	3819

Notes: Observations are at the school-grade-year level, and cover grades 3 through 5 in the 2002–03 through 2009–10 school years. Sample is limited to observations in districts with fewer than 200 students in the 1st grade cohort in 2004–05, in district-grade-years with between 8 and 39 LEP students in the 1st grade cohort, and with non-missing achievement variables. The observations differ slightly across the panels because the incidence of masked achievement outcomes differs across the student categories. Each coefficient and associated standard error reported comes from a separate regression that also controls for grade-year fixed effects, the percent of the school-grade-year overall who are female, economically disadvantaged, white, black and Hispanic, and the percent of students in the student category listed at the top of the table in each school-grade-year who are female or economically disadvantaged. Standard errors clustered by district are in parentheses.

* Denotes significance at 10% level.
** Denotes significance at 5% level.
*** Denotes significance at 1% level.

for all students in the same school-grade-year cells as the students whose test scores we analyze below. These covariates include gender, race, economic disadvantage (free or reduced lunch), enrollment in gifted and special education programs, and total grade enrollment. For all covariates, the graphs show smooth distributions around the cutoff point. More formally, we estimate Eq. 3 using each of these observable characteristics as the dependent variable. These results are reported in Appendix Table 7, panel A, and confirm the visual evidence in Appendix Fig. 1—there are no changes in covariates at the cutoff that are statistically significant at the 5% level and only one at the 10% level.[30] In order to combine these multiple tests into a single test statistic, we also estimate a Seemingly Unrelated Regression and perform a Chi-squared test for the hypothesis that the coefficients of *Above*20 across these regressions are jointly equal to zero. The p-value from this test, reported below the regression estimates, is 0.46.

Since we estimate effects on different categories of students below, we conduct similar tests at the level of school-grade-year-student category, though it should be noted that fewer variables are available. Panels B and C of Appendix Table 7 show the results for "non-LEP, non-Spanish home language" students and "Spanish home language" students, respectively. There is no evidence of discontinuities in share of these students who are female, economically disadvantaged or in gifted programs.

Overall, these results indicate that there are no discontinuities in the underlying characteristics of the students. This supports the interpretation that observed discontinuities at the 20-student cutoff are only due to the policy rule governing BE provision, and not to differential student composition above and below the cutoff.

5.4. Effect of district provision of bilingual education on student achievement

Fig. 4 provides visual evidence on the reduced-form relationship between district Spanish LEP count in the first grade cohort and the achievement of "non-LEP, non-Spanish home language" students. The top three graphs show effects on math achievement, and the bottom three graphs show effects on reading achievement. In all six graphs, we observe higher achievement among students in schools in districts above the cutoff. We also observe achievement gains to the right of the cutoff for "Spanish home language" students (Fig. 5), though they are smaller than that observed for "non-LEP, non-Spanish home language" students. As far as the secular effects of the running variable on student achievement are concerned, the graphs in Figs. 4 and 5 indicate that student achievement steadily decreases as Spanish LEP count increases. The pattern for the secular effects does not suggest that a quadratic or cubic form would offer a better fit than a linear form. This observation leads us to control for the underlying effects of the running variable using a linear function that is free to have different slopes above and below the cutoff in our preferred specification.[31]

Table 3 presents results from a formal evaluation of the achievement outcomes around the cutoff, as described in Section 5.2. Each cell of Table 3 represents a coefficient from a separate regression. In Panel A, we report results from estimating the reduced-form relationship between student outcomes and the dummy for the district being above the 20-student cutoff, i.e., the estimates of δ^{RF} in Eq. 3. The coefficient in the first row of column 1 indicates that non-LEP students in schools in districts with at least 20 Spanish LEP students in the first grade cohort have a mean standardized math test score that is 0.059 standard deviations (SD) higher than non-LEP students in schools in districts with less than 20 Spanish LEP students in the first grade cohort. Similarly, the coefficient in the second (third) row of column 1 indicates that the math passing (commended) rate of non-LEP students is 0.984 (2.215) percentage points higher in schools in districts located to the right of the cutoff. There are also improvements in the reading achievement of non-LEP students above the cutoff, though only the reduced-form effect on mean reading standardized test is statistically significant (column 2). Turning to results in columns 3 and 4 of Panel A, we find reduced-form estimates that are generally positive and small, and never statistically significant for the "Spanish home language" students. These results for "non-LEP, non-Spanish home language" and "Spanish home language" students are in line with the graphical evidence presented in Figs. 4 and 5, respectively. For all students combined, the reduced-form effects are positive and sometimes statistically significant (columns 5 and 6 of Panel A), mirroring the impacts on the "non-LEP, non-Spanish home language" students who form about three quarters of the student body in the schools in our samples.

In Panel B of Table 3, we present results from the estimation of the first stage equation (Eq. 2), followed by the results from our structural model (Eq. 1) using *Above*20 to instrument for district provision of bilingual education. As discussed above, the first row indicates that the first

[30] The covariate tests reported in Appendix Table 7 have only the above cutoff dummy and the smoother as right-hand-side variables. The results are similar when we add controls (e.g., grade-year fixed effects, covariates aside from the one being used as the dependent variable).

[31] In Section 5.6, we present results from specifications controlling for the running variable in different ways.

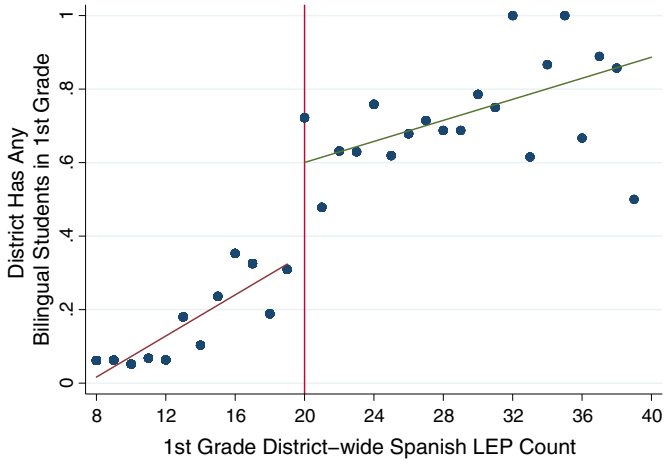

Fig. 2. District has a bilingual program in first grade, 1998–99 through 2007–08. Limited to districts with fewer than 200 students in 1st Grade in 2004. The school-year is the unit of observation. We use the years 1998–99 through 2007–08 because the 3*rd*–5*th* graders for whom we observe achievement outcomes in 2002–03 to 2009–10 map back to the first graders in 1998–99 to 2007–08.

stage relationship is strong and robust across all three student samples. The next three rows display the 2SLS effect of district BE provision on our measures of student achievement.[32] We discuss the 2SLS effects for each student category in turn. For "non-LEP, non-Spanish home language" students, we find positive effects of district BE provision on all three math outcomes. For mean standardized math achievement, the 2SLS coefficient is statistically significant at the 10% level, and indicates that district BE provision increases the mean math test score of non-LEP students by 0.21 SD.[33] While this seems large, we stress that this is a cumulative effect over 4 to 6 years. Thus, we can divide by 5 to get a very rough annual effect of 0.04 SD.

For math commended rate, the coefficient is statistically significant at the 5% level, and indicates that district BE provision increases the math commended rate of non-LEP students by 8.0 percentage points (row 3); this is a 30% increase when compared to the mean commended rate among schools in districts below the cutoff. The coefficient for the math passing rate of non-LEP students is 3.5 percentage points (row 4) and not statistically significant. It is interesting that the impact on the commended rate is larger than the impact on the passing rate. The positive point estimates for both suggest that students at the margin of failing/passing the exam and higher achievers, are contributing to the measured increase in average math scores but the larger commendable rate impacts indicate that the improvements are larger for high-achieving than for low-achieving students.

The 2SLS estimates of the impact of district BE provision on "non-LEP, non-Spanish home language" students' reading outcomes are reported in column 2 of Panel B. We again find that the estimated effects are positive, however only the 2SLS effect on mean standardized reading achievement is significant at the 10 percent level. Even though these 2SLS estimates are somewhat imprecise, and we cannot reject the null hypothesis that they are zero at the 5% level, we are able to reject relatively small negative effects on non-LEP students' reading achievement. The 95% confidence intervals for the effect on standardized reading score is [−0.01, 0.35], which means we can rule out district BE provision reducing standardized reading scores by more than a *hundredth* of a standard deviation at the 5% significance level. Similarly, for the reading passing rate, we can rule out negative effects over 1.6 percentage points at the 5% significance level; less than 2% of the mean passing rate (83%) for the sample below the cutoff. Also, for the reading commended rate, we can rule out district BE provision reducing reading commended rates by more than 1.6 percentage points at the 5% significance level.

As "non-LEP, non-Spanish home language" students would not have participated in the LEP programs themselves, the finding of some significant impacts of district BE provision on their academic performance points to the presence of spillover effects of educational programs for LEP students on their non-LEP peers. In Section 5.7 below, we discuss potential mechanisms for these spillover effects and compare the magnitudes to education peer effects estimated by other studies.

Next, we turn to the impacts of district BE provision on the standardized test performance of "Spanish home language" students, which are reported in Table 3, Panel B, columns 3 and 4. The 2SLS effects are generally positive, but never statistically significantly different from zero, mirroring the reduced-form effects found in Panel A. However, we are able to reject modestly sized negative effects on test scores. The 95% confidence intervals for the standardized math and reading score are [−0.20, 0.41] and [−0.27, 0.35], respectively. Thus, we can reject at the 5% significance level that district BE provision lowers mean math and reading scores by more than 0.2 and 0.3 SD, respectively, over a 4 to 6 year period. The coefficients for the passing rate and commended rate have wide confidence intervals, and thus we cannot rule out economically meaningful negative effects along these margins.

These results for "Spanish home language" students indicate that while district BE provision does not statistically significantly benefit "Spanish home language" students, it does not harm them either, at least in terms of test score outcomes. Since the Spanish home language student sample is primarily composed of ever-LEP students, these results also indicate that the impact of bilingual education provision on the intended beneficiaries is generally positive though not statistically significant. This finding is consistent with the Slavin et al. (2011), Matsudaira (2005) and Angrist, Chin and Godoy (2008) studies discussed in Section 2.2—LEP students' educational outcomes are not significantly different in bilingual programs compared to English-intensive programs.

Finally, in the last two columns of Table 3, Panel B, we present the 2SLS effects on the standardized test performance of all students combined. Arguably these net effects of district BE provision are the most relevant ones for policy making, as they weigh the test score gains and losses across all students equally. Even if in practice policy makers do not assign equal weights to each student—for example, to comply with U.S. civil rights laws, they might care more about the LEP students' educational progress—these estimates provide a benchmark to gauge the implicit efficiency gains or losses associated with a particular LEP program. These estimation results consistently point to net positive effects of district BE provision on student achievement, though the 2SLS estimates are significant at the 10% level or better only for mean standardized math achievement score and mean math commended rate. The positive net impact in the "all students" sample indicates that on average, test score gains due to district BE provision exceed test score losses experienced by students in a given cohort and school. That is, for the districts hovering around the 20-Spanish-LEP-student cutoff used in our analysis, provision of bilingual education actually led to net gains in student achievement without hurting either "Spanish home language" students as a group, or "non-LEP, non-Spanish home language" students as a group.

5.5. Single elementary school districts

Some school districts in our sample have multiple elementary schools, and offering BE in one of the elementary schools is enough for districts above the 20-student cutoff to be in compliance with the Texas policy rule. Because the Texas policy rule provides variation in district BE provision, we focus on the impacts of district BE provision in our main analysis. In general the impact of one's district providing BE may

[32] Given the first-stage coefficient of about 0.28, these 2SLS effects will be about 3.6 (=1/0.28) times the reduced-form coefficients in Panel A.

[33] The 95% confidence interval for the effect is [−0.01, 0.43], which means we can rule out district BE provision reducing standardized math score by more than a *hundredth* of a standard deviation at the 5% significance level.

Table 3
Estimates of effect of district bilingual education provision on TAKS achievement.

	"Non-LEP, Non-Spanish Home Language" Students		"Spanish Home Language" Students		All Students	
	Math	Reading	Math	Reading	Math	Reading
	(1)	(2)	(3)	(4)	(5)	(6)
A. Reduced form — OLS coefficient for "District has >= 20 Spanish LEP students in 1st Grade Cohort"						
Mean standardized achievement	0.059**	0.046*	0.030	0.011	0.056*	0.039
	(0.029)	(0.025)	(0.043)	(0.043)	(0.029)	(0.026)
Passing rate	0.984	1.378	0.455	−0.141	0.880	1.015
	(1.007)	(0.874)	(2.047)	(1.891)	(1.063)	(0.955)
Commended rate	2.215**	1.283	0.954	0.848	2.211**	1.313
	(1.075)	(0.894)	(1.313)	(1.222)	(1.006)	(0.859)
B. 2SLS — endogenous regressor is "District has Any Bilingual Program in 1st Grade Cohort"						
1st Stage, OLS coefficient for 1st Grade LEP Count >= 20	0.277***	0.275***	0.276***	0.277***	0.272***	0.272***
	(0.057)	(0.057)	(0.052)	(0.052)	(0.058)	(0.058)
2nd stage — mean standardized achievement	0.213*	0.169*	0.107	0.040	0.206*	0.143
	(0.114)	(0.092)	(0.155)	(0.156)	(0.112)	(0.095)
2nd stage — passing rate	3.547	5.010	1.652	−0.509	3.232	3.727
	(3.714)	(3.356)	(7.368)	(6.830)	(3.946)	(3.576)
2nd stage — commended rate	7.987**	4.666	3.462	3.058	8.124**	4.826
	(4.050)	(3.206)	(4.761)	(4.316)	(3.802)	(3.050)
Observations	3759	3761	3247	3244	3819	3819

Notes: Observations are at the school-grade-year level, and cover the 2002–03 through 2009–10 school years. Sample is limited to observations in districts with fewer than 200 students in the 1st grade cohort in 2004–05, in district-grade-years with between 8 and 39 LEP students in the 1st grade cohort, and with non-missing achievement variables. Each coefficient and associated standard error reported comes from a separate regression that also controls for district-wide Spanish LEP count in the relevant 1st grade cohort, district-wide LEP count interacted with a dummy for being above 20 Spanish LEP students, grade-year fixed effects, the percent of the school-grade-year overall who are female, economically disadvantaged, white, black and Hispanic, and the percent of students in the student category listed at the top of the table in each school-grade-year who are female or economically disadvantaged. Standard errors clustered by district are in parentheses.
* Denotes significance at 10% level.
** Denotes significance at 5% level.
*** Denotes significance at 1% level.

differ from the impact of one's own school providing BE. Among districts with only one elementary school, however, *district* provision of BE means *school* provision of BE, which allows us to investigate the impacts of school BE provision. In two respects, we might think of exposure to BE programs in single elementary school districts as more intense. First, since there is no scope for students to change schools within the same district, in this sample non-LEP students' exposure to LEP classmates unambiguously decreases above the cutoff.[34] Second, these schools may be especially responsive to the policy rule because not only are they small in terms of student enrollment (by construction of our analysis samples, they have fewer than 200 students in the first grade cohort), they also have less ability to shift resources. Besides within school shifts, only shifts across school levels (e.g., between elementary and secondary schools) are possible but these cross-level shifts are likely less feasible than shifts across schools at the same level.

In Table 4, we repeat the analysis of student achievement restricting our sample to observations in districts with a single elementary school from 1998–99 to 2009–10. Over 60% of the observations in the samples used in Table 3 are from single elementary school districts. The first stage estimates, reported in the first row of Panel B, are essentially the same as those obtained in our main analysis: the coefficient for *Above*20 is about 0.28 and statistically significant. In contrast, the reduced form estimates and the implied 2SLS effects are larger in magnitude and more often statistically significant. In particular, we find significant positive impacts of district BE provision on all six student achievement measures for "non-LEP, non-Spanish home language" students and all students combined. While the 2SLS effects on "Spanish home language" students are still not significant at conventional levels, they are now larger. The finding of stronger impacts among schools in single elementary school districts is consistent with these schools experiencing higher treatment intensity above the cutoff.

5.6. Sensitivity analysis

In Table 5, we examine the robustness of our results to several potential concerns.[35] Since our most novel contribution is investigating spillover effects of educational programs for LEP students, we present results using the "non-LEP non-Spanish home language" sample.[36] Each cell shows the 2SLS RD estimate of the effect of district BE provision from a separate regression. The first row displays the baseline 2SLS estimates from Table 3 for ease of comparison.

Our baseline model assumes that the underlying relationship between student achievement and district-wide Spanish LEP student count in the first grade cohort is piecewise linear (we control for a linear function of district-wide Spanish LEP student count that allows for different slopes above and below the cutoff). To the extent that this functional form is misspecified, the RD design leads to a biased estimate of the treatment effect. Our choice of a piecewise linear function was guided by the graphs of academic achievement by Spanish LEP count (Fig. 4 for "non-LEP, non-Spanish home language" students), which did not indicate nonlinearities in the underlying relationship. In rows 2 and 3 of Table 5, we test the sensitivity of our results to the assumed functional form of the running variable by choosing polynomials of different degrees. In row 2, when we use

[34] For multiple elementary school districts, a non-LEP student's exposure to LEP classmates can increase or decrease above the cutoff depending on which school offers the BE program, and the extent to which students change schools on the basis of BE program location.

[35] In Appendix Table 8, we provide additional evidence on the robustness of our results by examining the sensitivity of the estimates to the way we have treated masked values in the data, to the choice of school year for defining small school district and to restricting the sample to districts with different number of students in 2004. In addition, we have estimated our structural equation using several "placebo" cutoffs and found no evidence of an effect on non-LEP and LEP student achievement. We also did not find any evidence that district enrollment changed at the cutoff, alleviating concerns about cross-district migration.

[36] We also perform the same robustness checks for the "Spanish home language" and "all students" samples, and find that the baseline results stand (see Appendix Tables 9 and 10).

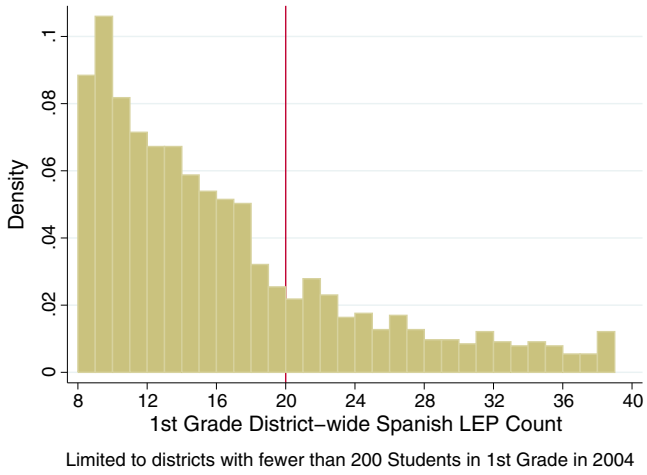

Fig. 3. Distribution of district first grade LEP counts, 1998–99 through 2007–08. We use the years 1998–99 through 2007–08 because the 3*rd*–5*th* graders for whom we observe achievement outcomes in 2002–03 to 2009–10 map back to the first graders in 1998–99 to 2007–08.

a quadratic smoother (i.e., we control for a quadratic function of Spanish LEP count that allows for different slopes above and below the cutoff), the point estimates still suggest a positive impact on student achievement. However, none of the coefficients are statistically significant because the standard errors are on average 1.6 times of those from the baseline model. The coefficient in row 2, column 1 indicates that district BE provision raises non-LEP, non-Spanish home language students' mean math achievement by 0.118 SD, but this coefficient is not significantly different from zero. When a cubic smoother is used (row 3), again none of the estimated effects are significant, though the point estimate indicates an effect of 0.117 SD on mean math achievement. With either the quadratic or cubic smoother, we note that the 95% confidence interval of the 2SLS effect of district BE provision contains the point estimates from the baseline specification. On the other hand, the lower bounds of the 95% confidence intervals are never below −0.3, which means we can rule out harmful effects on mean math and reading scores of greater than 0.3 SD over a 4 to 6 year period. Thus, regardless of which smoother we use, we find that district BE provision does not have large harmful impacts on non-LEP students' achievement.

In rows 4 and 5, we check the sensitivity of our estimates to the chosen bandwidth. Our main analysis was restricted to school-grade-year observations in districts with between 8 and 39 Spanish LEP students in the first grade cohort. Row 4 restricts the sample to observations in districts within a tighter range of the 20-Spanish-LEP-student cutoff: 10 to 29. The sample size is reduced, resulting in larger standard errors; however the results are qualitatively similar to the baseline model. In row 5, we expand the bandwidth to schools in districts with 8 to 49 Spanish LEP students in the first grade cohort and the results are again similar to the baseline model.

Finally, in row 6, we check the sensitivity of our estimates to the inclusion of control variables by estimating an RD model with only the above cutoff dummy and smoother as independent variables. If the Texas policy rule indeed provides exogenous variation in district BE provision, adding control variables should only improve the precision of the estimates while having no impact on the point estimates. This is borne out in our analysis: the point estimates are similar whether controls are used (row 1) or not (row 6), but the standard errors are considerably larger with no controls. This lends confidence that our estimated effects are indeed driven by the policy.

5.7. Potential mechanisms

One mechanism that could explain our findings is changes in class composition. District BE provision reduces the number of LEP students in mainstream classes as LEP students in BE form their own separate classes. In Fig. 6, we plot the average share of ESL students in the mainstream classrooms by the number of Spanish LEP students in the first grade cohort. LEP students are comprised of BE students and ESL students, with the former in separate classrooms and the latter remaining in the same classrooms as the non-LEP students, thus share ESL among students not in BE is a measure of exposure of non-LEP students to LEP classmates. Visually, there is a jump down, indicating that district BE provision lowers non-LEP students' exposure to LEP students inside the classroom. When we estimate Eq. 3 with share ESL among students not in BE as the dependent variable, the coefficient on *Above*20 is −5.26 (s.e. 1.43), i.e., district BE provision reduces LEP share of mainstream classes by 5 percentage points. If we attribute all of the reduced-form impacts on non-LEP student achievement to share of classmates who are LEP, then our estimates imply that a 10 percentage point increase in the share of a class that is LEP would reduce non-LEP math achievement by 0.11 SD and reading achievement by 0.09 SD.[37] We regard these as upper bounds for the peer effects of exposure to LEP students, since district BE provision likely affects non-LEP achievement besides through LEP class share.

It is instructive to compare these estimates of effects of exposure to LEP classmates to education peer effects estimates found in other studies. Before doing so, however, it is important to note that while the estimates discussed below are based off of one year of exposure, our estimates incorporate multiple years of exposure. Thus, if we make the (strong) assumption that LEP peer effects are cumulative and constant across grades, then with 5 years of exposure from kindergarten to 4th grade the upper-bound peer-effect estimates from LEP students are 0.02 SD in both math and reading.

Most directly related to our study are Cho (2011) and Geay, McNally and Telhaj (2012), which estimate the impact of LEP students on non-LEP peers. Using U.S. data, Cho finds that non-LEP students who have at least one LEP classmate have 0.04 to 0.06 SD lower reading achievement, and no difference in math achievement. The average non-LEP student in her sample had 2.6 LEP classmates, which is about 13% of the class. Using data from England, Geay et al. find effects of share LEP on non-LEP reading, writing and math test scores that are close to zero, and not statistically significant. Peer effects generated by other types of students could also provide an interesting contrast. For example, Lavy and Schlosser (2011) find that increasing the proportion of boys in a class by 10 percentage points reduces achievement by 0.02 SD while Lavy, Paserman and Schlosser (2012) find that a similar increase in the proportion of grade repeaters in a class reduces achievement by around 0.07 SD. Further, Carrell and Hoekstra (2010) find that a 10 percentage point increase in the share of students in a class with troubled home environments reduces achievement by around 0.05 SD. Thus, after adjusting for exposure time, the upper-bound on peer effects generated by LEP students on non-LEP students that we estimate are smaller than most of the preceding estimates and similar to the impacts of increasing the share of boys in a class.

While our findings are consistent with peer effects generated by LEP students, it is difficult to provide direct evidence for this interpretation. However, corroborative of this interpretation are the following. First is the finding of larger spillover effects for schools that are in single elementary school districts. These are the very schools where non-LEP students' exposure to LEP students clearly decreases above the cutoff. Second is the finding of larger positive spillover effects for the math commended rate than the math passing rate. This suggests that higher achieving children gain more when BE is provided. Positive effects of having more high-achieving peers have been documented by Imberman, Kugler and Sacerdote (2012), among others, and could arise from direct interaction between students or from indirect peer effects such as

[37] The effect of a one percentage change increase in share of a class that is LEP is calculated by dividing the reduced-form effect on non-LEP mean standardized achievement by the mean reduction in ESL rate of 5.26.

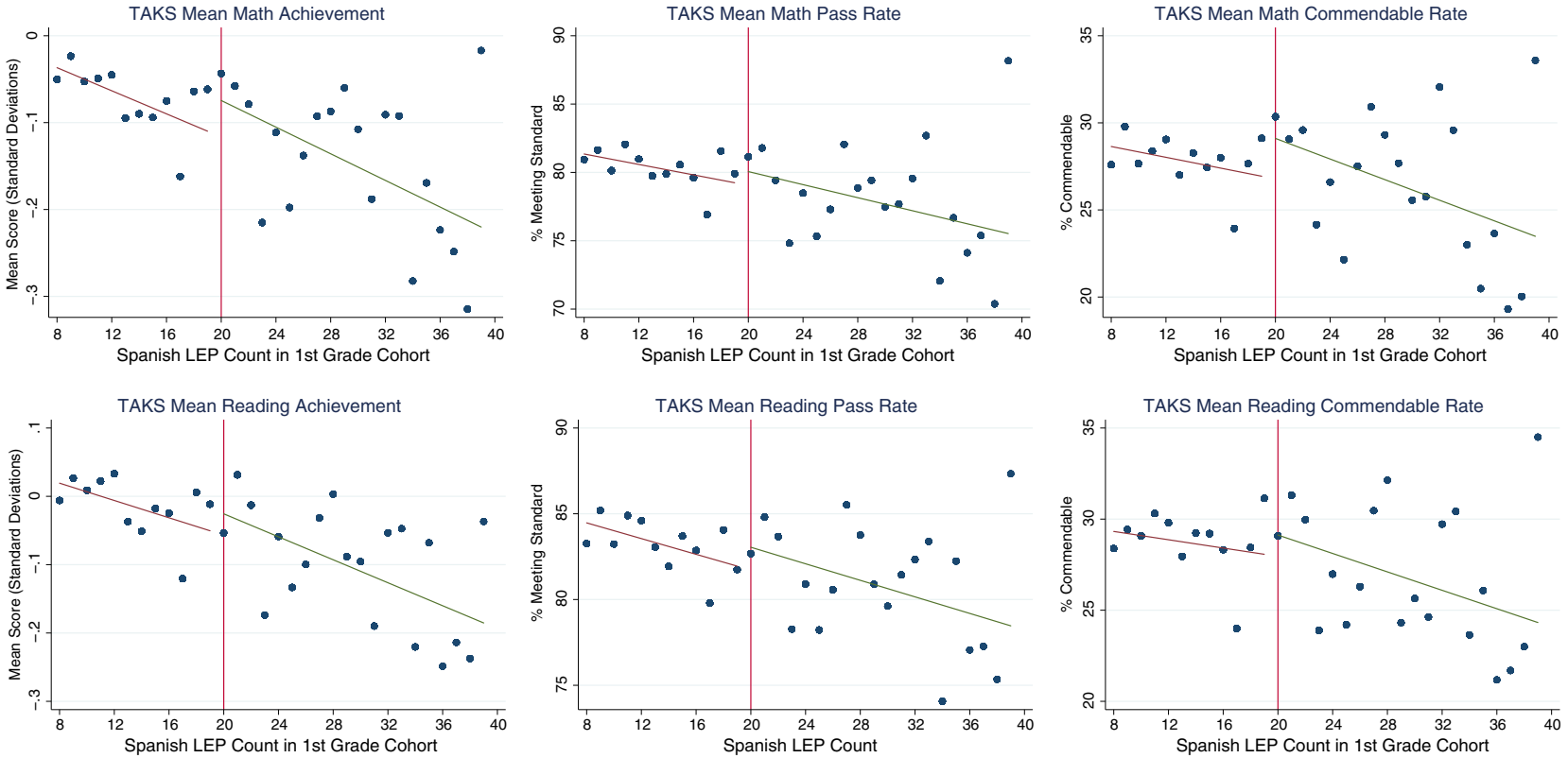

Fig. 4. TAKS performance for non-LEP, non-Spanish home language students. Covers 3*rd*–5*th* Grade from 2002–03 through 2009–10 using the "Non-LEP, Non-Spanish Home Language" sample.

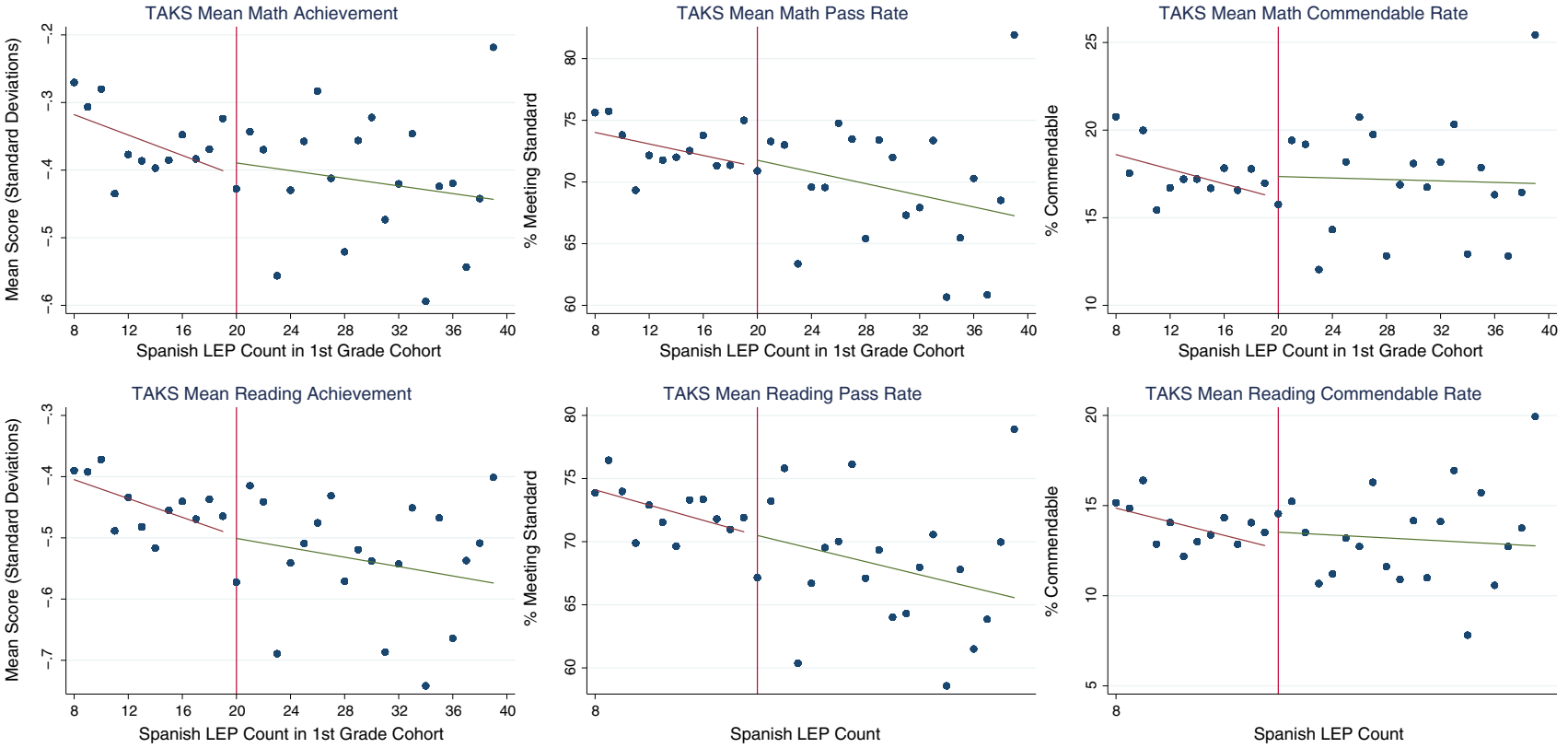

Fig. 5. TAKS performance for Spanish home language students. Covers 3*rd*–5*th* Grade from 2002–03 through 2009–10 "Spanish Home Language" sample.

Table 4
Estimates of effect of district bilingual education provision on TAKS achievement — districts with a single elementary school.

	"Non-LEP, Non-Spanish Home Language" Students		"Spanish Home Language" Students		All Students	
	Math	Reading	Math	Reading	Math	Reading
	(1)	(2)	(3)	(4)	(5)	(6)
A. Reduced form — OLS coefficient for "District has >= 20 Spanish LEP students in 1st Grade Cohort"						
Mean standardized achievement	0.107***	0.105***	0.088*	0.033	0.100***	0.089***
	(0.035)	(0.031)	(0.053)	(0.056)	(0.035)	(0.034)
Passing rate	2.817**	3.318***	3.286	1.184	2.549*	2.820**
	(1.280)	(1.017)	(2.486)	(2.543)	(1.378)	(1.224)
Commended rate	3.439***	2.848**	1.196	1.276	3.241***	2.665**
	(1.280)	(1.140)	(1.611)	(1.465)	(1.182)	(1.065)
B. 2SLS — Endogenous regressor is "District has Any Bilingual Program in 1st Grade Cohort"						
1st stage, OLS coefficient for 1st Grade LEP Count >= 20	0.285***	0.282***	0.268***	0.272***	0.279***	0.279***
	(0.065)	(0.065)	(0.066)	(0.066)	(0.065)	(0.065)
2nd stage — mean standardized achievement	0.374**	0.374***	0.329	0.120	0.360**	0.320**
	(0.152)	(0.134)	(0.202)	(0.203)	(0.148)	(0.133)
2nd stage — passing rate	9.878*	11.781***	12.264	4.357	9.134*	10.106**
	(5.008)	(4.399)	(9.093)	(9.318)	(5.219)	(4.709)
2nd stage — commended rate	12.062**	10.109**	4.462	4.692	11.616**	9.550**
	(5.356)	(4.486)	(6.041)	(5.340)	(4.929)	(4.193)
Observations	2296	2298	2105	2102	2350	2350

Notes: Observations are at the school-grade-year level, and cover the 2002–03 through 2009–10 school years. Sample is limited to observations in districts that had a single elementary school from 1998–99 to 2009–10 and with fewer than 200 students in the 1st grade cohort in 2004–05, in district-grade-years with between 8 and 39 LEP students in the 1st grade cohort, and with non-missing achievement variables. Each coefficient and associated standard error reported comes from a separate regression that also controls for district-wide Spanish LEP count in the relevant 1st grade cohort, district-wide LEP count interacted with a dummy for being above 20 Spanish LEP students, grade-year fixed effects, the percent of the school-grade-year overall who are female, economically disadvantaged, white, black and Hispanic, and the percent of students in the category listed at the top of the table in each school-grade-year who are female or economically disadvantaged. Standard errors clustered by district are in parentheses.
* Denotes significance at 10% level.
** Denotes significance at 5% level.
*** Denotes significance at 1% level.

Table 5
Sensitivity analysis of 2SLS estimates of effects on "Non-LEP, Non-Spanish Home Language" Students.

	Math			Reading		
	Mean achievement (1)	Passing rate (2)	Commended rate (3)	Mean achievement (4)	Passing rate (5)	Commended rate (6)
1) Baseline (from Table 3, columns 1 and 2)						
District	0.213*	3.547	7.987**	0.169*	5.010	4.666
Provides BE	(0.114)	(3.714)	(4.050)	(0.092)	(3.356)	(3.206)
Observations	3759	3759	3759	3761	3761	3761
2) Quadratic smoother						
District	0.118	1.065	1.396	0.114	4.304	3.041
Provides BE	(0.194)	(5.737)	(6.684)	(0.148)	(5.305)	(5.342)
Observations	3759	3759	3759	3761	3761	3761
3) Cubic smoother						
District	0.117	2.594	−0.250	0.006	2.563	−2.401
Provides BE	(0.201)	(5.930)	(6.225)	(0.140)	(5.446)	(4.835)
Observations	3759	3759	3759	3761	3761	3761
4) Bandwidth of 10 to 29						
District	0.151	1.979	4.829	0.151	3.782	5.500
Provides BE	(0.134)	(4.206)	(4.810)	(0.107)	(3.652)	(3.752)
Observations	2738	2738	2738	2740	2740	2740
5) Bandwidth of 8 to 49						
District	0.221*	8.394**	8.394**	0.158*	4.328	4.928
Provides BE	(0.118)	(4.006)	(4.006)	(0.090)	(3.299)	(3.033)
Observations	3899	3899	3899	3901	3901	3901
6) No controls or fixed effects						
District	0.144	3.455	7.983	0.108	4.600	4.044
Provides BE	(0.165)	(5.651)	(6.667)	(0.153)	(5.365)	(5.806)
Observations	3759	3759	3759	3761	3761	3761

Notes: See Table 3 notes regarding base sample and specification. Reported is the 2SLS coefficient for the endogenous regressor, "District has Any Bilingual Program in 1st Grade Cohort" where the identifying instrument is "District Has >= 20 Spanish LEP Students in 1st Grade Cohort". Standard errors clustered by district are in parentheses.
* Denotes significance at 10% level.
** Denotes significance at 5% level.
*** Denotes significance at 1% level.

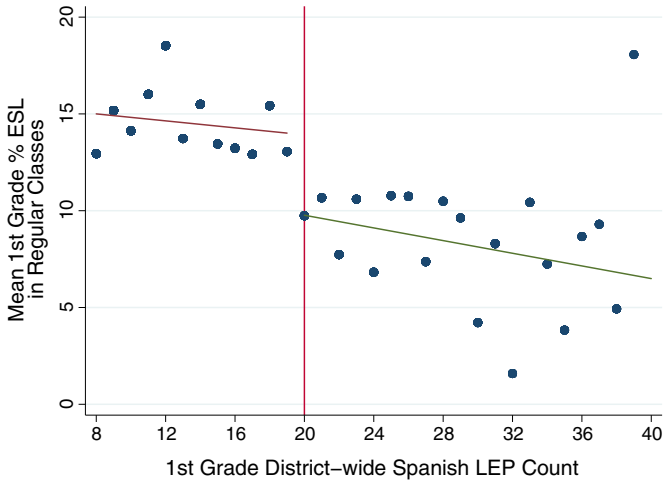

Fig. 6. Percent of regular class students in ESL, 1998–99 through 2007–08. Regular class students are defined as any student not in Bilingual Education. Limited to districts with fewer than 200 students in 1st Grade in 2004. The school-year is the unit of observation. We use the years 1998–99 through 2007–08 because the 3*rd*–5*th* graders for whom we observe achievement outcomes in 2002–03 to 2009–10 map back to the first graders in 1998–99 to 2007–08.

classroom instruction being adjusted in accordance with student composition (e.g., mainstream classes with fewer LEP students might cover more advanced academic content).

Besides changes in class composition, another mechanism that could explain our findings is changes in school resource allocation. Because BE programs cost more and state funding does not cover all of the incremental costs of BE in Texas, the main resource story a priori involved BE provision crowding out resources for non-LEP students, which would generate negative spillover effects. Our finding of positive spillover effects does not support this crowd-out story. A simple resource story that could reconcile the finding of positive spillover effects would be if district BE provision increased school resources for non-LEP students. Data limitations constrain our capacity to rigorously assess this, however in a crude analysis of available school expenditure data, we find a statistically significant increase at the cutoff in school per-pupil spending on programs for LEP students, but no significant change in total program spending.[38] While these results appear inconsistent with the simple resource story, it must be noted that they are quite imprecise, and we cannot rule out total spending increasing by at least the extra spending on LEP programs. Thus, while we do not find strong evidence in favor of a resource story, we acknowledge that the availability of richer data on resources would enable more powerful tests of this mechanism.

6. Conclusion

In this paper, we examine the effects of bilingual education programs (versus ESL programs alone) on the achievement of intended beneficiaries (LEP students) and their classmates (non-LEP students). To address potential bias due to the endogeneity of student exposure to BE, we use a regression-discontinuity approach that exploits a policy rule governing the provision of BE in Texas. We find that district BE provision has significant positive effects on the standardized test scores of "non-LEP, non-Spanish home language" students, indicating the presence of spillover effects for educational programs for LEP students on non-LEP peers. We do not find statistically significant effects on the standardized test scores of "Spanish home language" students, a group that is 89% ever-LEP students, indicating that the intended beneficiaries of the LEP programs fare similarly in BE and ESL programs. On net for all students, the effects are positive and significant.

Given the high rates of low-skilled immigration in recent decades and the dispersion in settlement patterns of immigrants away from traditional immigrant-receiving areas, the issue of how to help LEP students is likely to keep its place at the center stage of education policy debates. Our results contribute to this debate by providing a convincing research design to evaluate the relative merits of BE programs compared to ESL programs alone—a question relevant to many school districts. While our findings are obtained in the context of smaller, less urban school districts in Texas, which may limit their external validity, nonetheless they have broad implications. In particular, any cost–benefit analysis on the value of LEP programs should take spillovers into account. In our context, BE programs provide more benefits than what the estimated effects on intended beneficiaries alone indicate, because these estimates ignore the benefits accruing to LEP students' peers. Although BE programs are adopted with the intention of helping LEP students, in our context it appears that the non-LEP students gain more from them, at least in terms of standardized test scores.

[38] Ideally, we would like resource data at the school-grade level for the same years as the achievement data. Unfortunately, the available data are only at the school level and only beginning in 2004–05. We obtain the school-grade-year level dataset for analyzing the impact of district BE provision on school resources by merging in the resource data by school and first grade cohort. For more details, see Chin et al. (2012).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jpubeco.2013.08.008.

References

Aizer, Anna, 2008. "Peer Effects and Human Capital Accumulation: The Externalities of ADD". NBER Working Paper No. 14354. National Bureau of Economic Research, Cambridge, MA.
Angrist, Joshua, Lang, Kevin, 2004. Does schooling integration generate peer effects? Evidence from Boston's Metco Program. Am. Econ. Rev. 94 (5), 1613–1634.
Angrist, Joshua, Pischke, Jorn-Steffen, 2009. Mostl Harmless Econometrics. Princeton University Press, Princeton, NJ.
Angrist, Joshua, Chin, Aimee, Godoy, Ricardo, 2008. Is Spanish-only schooling responsible for the Puerto Rican language gap? J. Dev. Econ. 85 (1–2), 105–128.
Baker, K.A., de Kanter, A.A., 1981. Effectiveness of Bilingual Educational: A Review of the Literature. Washington, D.C, Office of Planning, Budget and Evaluation, U.S. Department of Education.
Barreca, Alan, Lindo, Jason, Waddell, Glen, 2011. "Heaping-Induced Bias in Regression Discontinuity Designs". NBER Working Paper No. 17408. National Bureau of Economic Research, Cambridge, MA.
Bertrand, Marianne, Duflo, Esther, Mullainathan, Sendhil, 2004. How much should we trust differences-in-differences estimates. Q. J. Econ. 119 (1), 249–275.
Betts, Julian R., 1998. Educational crowding out: do immigrants affect the educational attainment of american minorities? In: Hamermesh, Daniel S., Bean, Frank D. (Eds.), Help or Hindrance? The Economic Implications of Immigration for African-Americans. Russell Sage Foundation, New York.
Carrell, Scott, Hoekstra, Mark, 2010. Externalities in the classroom: how children exposed to domestic violence affect everyone's kids. Am. Econ. J. Appl. Econ. 2 (1), 211–228.
Chin, Aimee, Meltem Daysal, N., Imberman, Scott A., 2012. Impact of Bilingual Education Programs on Limited English Proficient Students and Their Peers: Regression Discontinuity Evidence from Texas." NBER Working Paper No. 18197. National Bureau of Economic Research, Cambridge, MA.
Cho, Rosa, 2011. Are there Peer Effects Associated with Having English Language Learner (ELL) Classmates?: Evidence from the Early Childhood Longitudinal Study Kindergarten Cohort (ECLS-K). Brown University, mimeo.
Crawford, James, 1989. Bilingual Education: History, Politics, Theory and Practice. Crane Publishing Company, Inc., Trenton, NJ.
Figlio, David, 2005. "Boys Named Sue: Disruptive Children and Their Peers". NBER Working Paper No. 11277. National Bureau of Economic Research, Cambridge, MA.
Geay, Charlotte, McNally, Sandra, Telhaj, Shqiponja, 2012. Non-native Speakers of English in the Classroom: What are the Effects on Pupil Performance?" IZA Discussion Paper No. 6451.
Greene, Jay P., 1998. A Meta-Analysis of the Effectiveness of Bilingual Education. Tomas Rivera Policy Institute.
Hanushek, Eric, Kain, John, Markman, Jacob, Rivkin, Steven, 2003. Does peer ability affect student achievement? J. Appl. Econ. 18 (5), 527–544.
Hoxby, Caroline M., 1998. Do immigrants crowd disadvantaged american natives out of higher education? In: Hammermesh, Daniel S., Bean, Frank D. (Eds.), Help or Hindrance? The Economic Implications of Immigration for African-Americans. Russell Sage Foundation, New York.
Hoxby, Caroline, Weingarth, Gretchen, 2006. Taking race out of the equation: school reassignment and the structure of peer effects. Presented at the 2006 American Economic Association Annual Meetings, 2006.

Imbens, Guido, Lemieux, Thomas, 2008. Regression discontinuity designs: a guide to practice. J. Econ. 142 (2), 615–635.
Imberman, Scott, Kugler, Adriana, Sacerdote, Bruce, 2012. Katrina's children: evidence on the structure of peer effects from hurricane evacuees. Amer. Econ. Rev. 102 (5).
Lavy, Victor, Schlosser, Analia, 2011. Mechanisms and impacts of gender peer effects at school. Am. Econ. J. Appl. Econ. 3 (2), 1–33.
Lavy, Victor, Paserman, Daniele, Schlosser, Analia, 2012. Inside the black box of ability peer effects: evidence from variation in high and low achievers in the classroom. Econ. J. 122 (559), 208–237.
Lee, David S., Lemieux, Thomas, 2010. Regression discontinuity designs in economics. J. Econ. Lit. 48 (2), 281–355.
Liu, Samuel T., 2000. Essays on the Effects of Immigration on Education and Crime (Ph.D. thesis, MIT).
Matsudaira, Jordan D., 2005. Sinking or Swimming? Evaluating the Impact of English Immersion versus Bilingual Education on Student Achievement. University of California, Berkeley, mimeo.
McCrary, Justin, 2008. Manipulation of the running variable in the regression discontinuity design: a density test. J. Econ. 142 (2), 698–714.
National Clearinghouse for English Language Acquisition (NCELA)., 2011. The Growing Numbers of English Learner Students 1998/99–2008/09. U.S. Department of Education Office of English Acquisition mini-poster, Washington DC (February).
Neymotin, Florence, 2009. Immigration and its effect on the college-going outcomes of natives. Econ. Educ. Rev. 28, 538–550.
Nieto, Diego, 2009. A Brief History of Bilingual Education in the United States. Perspectives on Urban Education Spring 61–72.
Rossell, Christine H., Baker, Keith, 1996. The educational effectiveness of bilingual education. Res. Teach. Engl. 30, 7–74 (February).
Slavin, Robert E., Madden, Nancy, Calderón, Margarita, Chamberlain, Anne, Hennessy, Megan, 2011. Reading and language outcomes of a five-year randomized evaluation of transitional bilingual education. Educ. Eval. Policy Anal. 33 (1), 47–58.
U.S. Commission on Civil Rights, 1975. A Better Chance to Learn: Bilingual Bicultural Education, U.S. Commission on Civil Rights Clearinghouse Publication No. 51.
Willig, Ann C., 1985. Meta-analysis of selected studies on the effectiveness of bilingual education. Rev. Educ. Res. 55, 269–317 (Fall).
Zehler, Annette M., Fleischman, Howard L., Hopstock, Paul J., Stephenson, Todd G., Pendzick, Michelle L., Sapru, Saloni, 2003. Descriptive Study of Services to LEP Students and LEP Students with Disabilities, Volume IA Research Report — Text. Department of Education, Washington, DC.