The Effect of Bilingual Education: Long-Run Student Outcomes and Spillover Effects*

Kevin Hunt† and Hyunkyeong Lim‡
University of Wisconsin-Madison

March 8, 2024

Abstract

A large and growing share of students arrive at school as English language learners (ELL). In Texas, schools provide bilingual education or an English as a second language programs for these students. Using a regression discontinuity design based on a quasi-experimental policy variation in Texas, this paper examines the effect of exposure to bilingual education in 1st grade on students' test scores and college enrollment. We find no effect of bilingual education in first grade on elementary school standardized test scores but large significant effects on college enrollment for both ELL and non-ELL students: bilingual education increases four-year university enrollment for ELL and non-ELL students by 6.4 and 6.8 p.p., respectively. The positive effects for both student groups in the long-run suggests that bilingual education has positive effects that are not well measured by short-run test scores for ELL students, and these effects spill-over to the outcomes on non-ELL students.

*The conclusions of this research do not necessarily reflect the opinions or official position of the Texas Education Agency, the Texas Higher Education Coordinating Board, the Texas Workforce Commission or the State of Texas.

†kghunt@wisc.edu.
‡hyunkyeong.lim@wisc.edu

1 Introduction

A large and growing share of students in the United States arrive at school as English language learners (ELL), typically because they speak another language at home. In Fall 2020, 15% of public school first-graders were classified as ELL, and 76% of these students spoke Spanish as a first language (NCES (2023)). Large achievement gaps between ELL and non-ELL students make it crucial for policy-makers to understand how to best educate these students (NCES (2018)). These policies directly impact the outcomes of participating ELL students, and non-ELL students may also be affected by changing classroom compositions and peer effects. While previous papers have considered the causal effect of ELL education programs in elementary school on contemporaneous test scores (Chin et al. (2013), Pope (2016)), their effects on long-run outcomes such high school graduation and college enrollment are not well understood.

This paper aims to fill the gap in the literature by estimating the causal effect of bilingual education compared to English as a Second Language (ESL) programs in Texas elementary schools on the long-run outcomes for both ELL and non-ELL students. Bilingual education programs place ELL students with the same home language into a separate classroom where the core curriculum is taught in their native language alongside English instruction. Conversely, ESL programs keep ELL and non-ELL students in the same classroom for core curriculum instruction in English, and ELL students learn English during a designated pull-out time. While ELL students may learn the core curriculum more effectively via their native language in a bilingual education program, this also may slow their English acquisition because of reduced exposure compared to ESL programs. Both ELL and non-ELL student may also be affected by the different composition of their classmates under each program.

Using administrative records from the Texas Education Research Center (ERC), we are able to observe which elementary schools provide bilingual education and link elementary school students to future education and unemployment insurance records. Following Chin et al. (2013), we leverage quasi-experimental policy variation in Texas for whether an elementary school offers a bilingual education or ESL program. Texas law requires school districts to offer bilingual education in at least one school for an elementary grade if there are more than 20 ELL students with the same home language enrolled in that grade. School districts near the cut-off are likely to be similar in unobservable characteristics but have a differential probability of providing bilingual education. By comparing the outcomes of students just above and below the cut-off in a fuzzy regression discontinuity design, we can causally estimate the effect of bilingual education compared to ESL. While Chin et al. (2013) considered only publicly available elementary school test scores as an outcome, our rich administra-

tive data allows us to consider additional long-run outcomes including post-elementary test scores, high school graduation, college enrollment, and post-educational labor earnings. We find that bilingual education in elementary school has a large positive effect on both ELL and non-ELL students' four-year university enrollment rates.

To begin, we verify that the 20 student cut-off is a valid instrument for the provision of bilingual education in elementary school following the empirical strategy of Chin et al. (2013). We focus on the number ELL students with Spanish as a home language enrolled in a district in first grade as the instrument and whether a school district provides a bilingual education program in first grade as the treatment variable. Additionally, we restrict the sample to smaller districts with less than 200 total students because the policy cut-off is more likely to bind for these districts; larger districts with more resources are likely to provide bilingual education programs even when they have relatively few ELL students. Although this means that our results are only for a selected sample of schools, they are still policy relevant. Our sample districts cover a wide geographic range in Texas, and changing migration patterns have led to an increase in ELL students in smaller, more rural districts similar to our sample. Our sample covers first grade cohorts for academic years 1993-94 to 2008-09 linked to outcomes through 2019. Within our sample, districts that have more than 20 ELL students with Spanish as a home language are 0.25 p.p. more likely to offer bilingual education in first grade. The first stage F-statistic is always greater than 22. This is similar to the first-stage results from Chin et al. (2013).

Next, we attempt to replicate the second stage results from Chin et al. (2013) for elementary school test scores. While Chin et al. (2013) used only aggregated data for each school-grade-year cell for various demographic groups, we have access to the individual level records in the administrative files. In contrast to their results, our estimates for the effect of bilingual education are noisy and not statistically significant.

In the long-run, however, we find large and statistically significant positive impacts of bilingual education on student outcomes. The provision of bilingual education in first grade increases the future probability that students enroll in a four-year public university in Texas by 5.9 p.p. (27.8%). This is true for both ELL first graders, whose four-year enrollment increases by 6.4 p.p. (49.2%), and for non-ELL students, whose enrollment increases by 6.8 p.p. (30.0%). All these effects are statistically significant at the 10 percent level. The existence of a positive effect for both ELL and non-ELL students suggests that bilingual education for ELL students has positive spill-over effects to non-ELL students that persist into post-secondary education. We do not find any evidence of an effect of bilingual education on high school graduation rates or community college enrollment.

These results are consistent with previous work that finds early life interventions can

have large long-run effects on students outcomes in adulthood. (Chetty et al. (2014), Chetty et al. (2011), Jackson (2018), Petek and Pope (2016)) but reappear in students' later life outcomes. Mechanisms that could generate such delayed impacts could be that bilingual education programs improve students' beliefs in their abilities or expectations of the return to education. Additionally, bilingual education may provide skills and knowledge that is not well measured by standardized tests but improves performance in later grades and post-education outcomes.

Our research contributes to two literatures. First, our project contributes to the literature on the causal impact of education programs for ELL students by considering the future impacts of bilingual education on outcomes such as high school test scores, high school graduation, and college enrollment. Previous research has focused on short-run outcomes such as elementary school test score and found mixed results. Studies show that ELL students perform worse when enrolled in bilingual education (Pope (2016)), that bilingual education has negligible impacts on ELL student outcomes (Matsudaira (2005); Chin et al. (2013)), or that bilingual education improves ELL student outcomes (Collier and Thomas (2017)). Chin et al. (2013) also highlights the importance of considering the impact of ELL student education programs on non-ELL students; this paper finds that non-ELL students perform better on standardized tests when ELL students are taught in separate classrooms.

Second, our research contributes to the literature on the impacts of early life interventions such as kindergarten and first grade teachers. This literature has highlighted the importance of considering long-run outcomes to evaluate education programs that happen at young ages. For example, the literature on the impact of pre-school enrollment has often found that test-score gains from pre-school enrollment do not persist into future grades, but that these benefits reappear in future outcomes such as high school graduation, college enrollment, and labor income (Deming (2009); Pages et al. (2020); Bailey et al. (2021); Gray-Lobe et al. (2021)). Many programs for ELL students are targeted at earlier grades because ELL student numbers are largest in earlier grades. Our research shows the importance of considering long-run outcomes for evaluating ELL student programs. While we find little effect of first grade bilingual education on first grade test scores, students who enrolled at a first grade school district that provided bilingual education have higher four year college enrollment rates.

2 Policy Background

The education of ELL students and use of non-English languages in public schools have been the subject of legal debate in the United States dating back to the 19th century.[1] Schools

[1]See Wright (2010) for a summary of court rulings and laws effecting ELL student education in USA.

must strike a delicate balance between providing an adequate education for ELL students and not segregating these students from their non-ELL peers. While courts and federal policy have avoided specifying an exact education model, schools are required to provide a curriculum for ELL students that both teaches English as well as general academic content.

School's instruction of ELL students typically falls into one of two categories. English as a second language (ESL) is the most common form, accounting for the education of 77.6% of ELL public school students in 2019-20 (Williams and Zabala (2023)). In these programs, ELL students and non-ELL students are taught in the same classroom via English for the core curriculum, and teachers for these classrooms do not necessarily have any additional skills beyond standard requirements. During a pull-out time, ELL students are taught English by a certified ESL teacher (TEA, 2022). In addition to standard teacher requirements, this certification in Texas requires 120-150 hours of coursework as well as passing an exam on topics including basic linguistics and ESL teaching methodologies. ESL certifications need to be renewed each five years, and renewal requires at least 150 hours of professional development (Anantha (2024)). ELL students enrolled in ESL will be exposed to a substantial amount of English during the core curriculum as well as with their ESL teacher, who may not even speak their home language. This may improve the speed at which they learn English but also hamper them in understand material in the core curriculum. Non-ELL students will also interact with ELL students during the core curriculum, and this may impact how the core curriculum is taught.

In contrast to ESL, a bilingual education program separates ELL students into a different classroom taught by a certified bilingual educator; the core curriculum for ELL students is taught in their native language while students also learn English. Programs may either be classified as early-exit, meaning the students are expected to be reclassified as non-ESL in 2-5 years, or late-exit, where reclassification is expected in 6-7 years (TEA, 2022). In addition to standard teaching requirements, bilingual education teachers in Texas must pass exams on teaching methodologies as well as proficiency in the language of instruction. After passing exams, a one-year internship of teaching with a field supervisor must be completed for full certification (TTOT, 2023). Hiring and retaining a bilingual certified teacher is the main hurdle for schools to implement a bilingual education program. While the number of ELL students in Texas has grown dramatrically, the number of bilingual certified educators has actually declined (Lopez (2022)). Relative to ESL, bilingual education exposes ELL students to less English; this may slow their English acquisition but also improve their understanding of the core curriculum. Non-ELL students will also interact less with ELL students, and teachers in the standard classroom may adjust their teaching in response to changing classroom composition.

Texas state law has several stipulations about when a school district must provide bilingual education.[2] In elementary school (grades 1 to 5), a school district with 20 or more ELL students with the same home language must provide a bilingual education program in that grade. School districts with less than 20 ELL students with the same home language are free to choose between providing a bilingual education or ESL program. Not all school districts above the 20 student threshold are able to provide bilingual education, largely because they do not have a bilingual education certified teacher. These districts must apply for and receive a waiver from the state board of Texas to not face sanctions. As part of the waiver, these districts must outline their plan for providing bilingual education in the future (for example, by hiring or training a bilingual teacher). After elementary school, districts can choose to provide either a bilingual education or ESL program regardless of the number of ELL students. The requirement that elementary school districts with 20 or more ELL students in a grade-home language cell must provide a bilingual education program forms the basis of our empirical strategy discussed in section 3.

In Texas students are assigned an ELL status based on their home language and a English proficiency test.[3] Students who report English as their only home language classified as English proficient. If a student speaks a language other than English at home, they are required to take an English proficiency test specific to their grade level. The Texas Education Administration determines the passing threshold for this test, and students who score below the threshold are classified as ELL. If the ELL student is enrolled in a district that provides bilingual education, the student will be offered enrollment in the bilingual education course. Parents, though, may request that their child is not enrolled in the bilingual education program. Students whose parents request to not be enrolled in bilingual education or who are enrolled in districts that do not provide bilingual education are offered enrollment in the ESL program. Again, parents can request that their child does not participate in the ESL program. Students lose their ELL status once they are able to pass the English proficiency test, which is administered each year to all ELL students. After passing the proficiency test, students who were previously classified as ELL are monitored for an additional two years to confirm English proficiency.

[2]Detailed instructions for requirements for education of ELL students can be found on the Texas Education Administration's website (https://tea.texas.gov/academics/special-student-populations/english-learner-support/bilingual-education-exception-and-esl-waiver-resources).

[3]A detailed flow chart for ELL students classification and progression can be found at https://tea.texas.gov/sites/default/files/eb-el-decision-chart-for-lpac-accessible-version.pdf.

3 Empirical Strategy

To estimate the effect of exposure to bilingual education on student outcomes, we will need to compare students who did and did not attend elementary school at a district that provided bilingual education. This exposure to bilingual education is not random, though. For example, school districts may decide to provide bilingual education if their ELL students are performing particularly poorly in a standard classroom setting. In this scenario, we may estimate that ELL students do worse when a bilingual education program is provided, but in fact bilingual education is provided because these ELL students are doing poorly. Moreover, school districts with larger numbers of ELL student are more likely to provide bilingual education, and these school districts are different than school districts that do not provide bilingual education (larger, more urban, more economically disadvantaged students). If the likelihood of bilingual education provision increases with the number of ELL students, this can coincide with negative peer effects from academically less-prepared economically disadvantaged students, which can create negative bias to the causal effect of bilingual education.

Thus, we will leverage policy variation in Texas that generates quasi-random exposure to bilingual education in elementary school to measure the causal impact of bilingual education. As discussed in section 2, Texas requires that school districts with 20 or more ELL students in an elementary grade (1-5) with the same home language must provide bilingual education. School district grades with 19 and 20 ELL students are likely to be very similar on average; the only difference will be that the school district is required to provide bilingual education if there are 20 ELL students. By comparing school districts just above and just below the cut-off, we will be comparing similar sets of students except for the exposure to bilingual education. Any differences in outcomes between these two sets of students is likely to be caused by the bilingual education program. This policy variation suggests a regression discontinuity design as discussed in Imbens and Lemieux (2008) and Lee and Lemieux (2010).

Previous research by Chin et al. (2013) leverages this 20 ELL students cut-off in Texas to estimate the impact of bilingual education programs on elementary school test-scores. We extend this strategy to document the impact of bilingual education on student's non-test score and long-run outcomes including middle, high school graduation, and college enrollment. We estimate the effect of bilingual education at the individual level as our main interest is to measure the impact of bilingual education on each *student*'s long-run outcomes. The goal is to understand how bilingual education helps or hurts students in developing language and academic skills.

Under Texas law, the 20 ELL student cut-off is not always strictly binding as discussed in section 2. Thus, we use fuzzy regression discontinuity design where bilingual education

provision is instrumented by the 20 ELL student cutoff. In order for the policy variation to be a valid instrument, the 20 ELL student threshold must be a strong predictor of bilingual education programs. We first verify that the cutoff is a valid instrument through the following first-stage regression:

$$BE_{idc} = \delta^{FS}\mathbf{1}\{\#ELL_{dc} \geq 20\} + f^{FS}(\#ELL_{dc}) + X'_{dc}\gamma^{FS} + \epsilon^{FS}_{idc}. \quad (1)$$

BE_{idc} is an indicator for whether school district d of individual i provided bilingual education for cohort c in their first grade. BE_{idc} will take 1 if there was any student who received bilingual education in school district d in the first grade for cohort c based on each individual's bilingual education status. $\#ELL_{dc}$ is the number of ELL students in this district-cohort where f^{FS} is a continuous function of $\#ELL_{dc}$.[4] We additionally control for demographic characteristics of the first grade cohort c in school district d to alleviate the concern of ELL students sorting into school districts with certain demographic composition.

In addition to being a strong predictor of bilingual education provisions, the instrument (i.e. the number of ELL student being above the threshold) must not be correlated with other variables that impact student achievement. In the regression discontinuity framework, it is also important to check for bunching below the threshold, which would suggest that schools actively try to reduce ELL counts to not provide bilingual education. We show below that the density of school districts is smooth below and above the 20 ELL student cutoff. Finally, we only use schools with 8-39 ELL students in the 1st grade cohort to accurately measure the local average treatment effect of bilingual education following the previous work by Chin et al. (2013).

After confirming that the cut-off is a valid instrument for bilingual education provision, we will compare student outcomes at school districts just above and just below the threshold using the following regression:

$$y_{idc} = \delta^{SS}\mathbf{1}\{\#ELL_{dc} \geq 20\} + f^{SS}(\#ELL_{dc}) + X'_{dc}\gamma^{SS} + \epsilon^{SS}_{idc}. \quad (2)$$

Here, δ^{SS} captures the effect of a change in student outcomes y above the policy threshold. As mentioned previously, students just below and above this threshold will be very similar on average except that students above the threshold are much more likely to have been at a school district that provided bilingual education. Thus, δ^{SS} will provide evidence of the causal impact of bilingual education on student outcomes. Furthermore, re-scaling δ^{SS} by the first stage coefficient will give an estimate that can be interpreted as the effect of bilingual

[4]We use a linear function of $\#ELL_{dc}$ where the slopes are allowed to be different below and above the 20 ELL cutoff for f^{FS}.

education provision.

4 Data

Our project leverages detailed longitudinal administrative data from the Texas Education Research Center (Texas ERC) about Texas K-12 students. The sample contains Texas public school students enrolled for the academic years 1993-94 to 2018-2019.

The data contains information about a student's home language and ELL status. Therefore, we are able to see which students are eligible for additional help from schools based on their English proficiency. We focus on ELL students whose home language is Spanish. Additionally, the data identifies which program the ELL students are enrolled in (bilingual education or ESL). Using this enrollment information, we can classify school districts as providing bilingual education or not.[5] The longitudinal structure of the data allows us to analyze a rich set of student outcomes.

First, we have standardized test scores in math and reading for students from grades 3 to 12. ELL students in elementary school are allowed to receive these tests in their native language, while ELL students in older grades receive the test in English but with linguistic accommodations. These test scores will allow us to analyze how bilingual education programs impact academic achievement both when they are receiving the treatment (in elementary school) and in future grades.

Second, we have high school graduation information for students who graduated from a Texas public school system. Unfortunately, we cannot disentangle whether a student leaves the data because they dropped out of school or because they have transferred to a school district outside of the Texas public school system. Because Texas is a large state, a majority of students who start their schooling in Texas public schools will also finish their schooling in Texas; therefore, we are able to connect a large number of high school graduates to their first grade school district and exposure to bilingual education. Furthermore, if we assume that the bilingual education program has little or no influence on the decision to transfer to schools not in the Texas public school system, then any high school graduation differences can be assumed to be caused by changing drop out rates.

Third, we have college enrollment and graduation information for Texas public colleges. Similar to high school graduation, many Texas students decide to enroll in Texas public colleges for their tertiary education. Thus, we are able to match a larger number of Texas college students in the data to their first grade school district and exposure to bilingual

[5]We do not observe directly whether a school district provides bilingual education or not. We identify a school district to provide bilingual education if there is any student who is enrolled in bilingual education.

education. If we assume that bilingual education exposure does not influence the decision of students to enroll in colleges outside of Texas public school system, then we can estimate the impact of bilingual education on college enrollment decisions.

Finally, the data contains information on individuals labor income for workers in Texas who are covered by unemployment insurance. Unfortunately, because our treatment occurs in first grade, we need to follow a student for at least twelve years to see labor income in the year after an on-time high school graduation, and 16 years for an on-time college graduation. Even 16 years after first grade, many students are still in school or establishing their careers, and it would be ideal to use labor income at even later ages. Because of this need to observe students over long time frames to measure accurate labor income, our data time frame makes it difficult to measure accurate effects, so we have omitted these results here. In the future, we hope to include results for labor income because we do observe a few first grade cohorts into their late 20s.

Following Chin et al. (2013), we focus our analysis on first-graders whose district had between 8 and 39 ELL students because the regression discontinuity is identified in the neighborhood of the 20 student cut-off. Additionally, we drop school districts with more than 250 students in the first grade cohort in 2004-2005 academic year; larger school districts have more financial resources and are more likely to provide a bilingual education program regardless of whether or not they are above the cut-off. Restricting our sample to school districts with relatively small numbers of ELL student and small total number of students is a limitation of our empirical strategy as these school districts are substantially different than the average school district. Despite this limitation, the clean identification of bilingual education vs. ESL will yield interesting insights.

5 Results

5.1 Validity of the Empirical Strategy

As discussed in section 3, our main empirical setting is the policy variation in Texas that requires school districts with 20 or more ELL students in an elementary grade (1-5) with same home language to provide bilingual education environment. However, the 20 ELL student cutoff was not always enforced and some schools were waived from providing a separate classroom for ELL students where they are taught in their home language (see section 2 for details). Thus, for the policy variation to be a valid instrument, there must be a significant jump in the likelihood of providing bilingual education at the 20 ELL student cutoff.

Figure 2 illustrates the fraction of districts which provide bilingual education as a function

of the number of Spanish ELL first graders in the district using the 1994-2008 first grade cohorts. Visually, it confirms that there is a discontinuous jump in the probability at the 20 ELL student cutoff. The relationship between the probability of providing bilingual education and the number of Spanish ELL students in the district becomes noisier above the cutoff. To be more statistically precise, we run equation 1 and check if the probability of bilingual education provision changes statistically significantly at the 20 ELL student cutoff. Table 3 presents the results. The F-statistics of our instrument (the 20 ELL student cutoff) in the first-stage in Table 3 range from 10.56 to 26.87 during 1994-2008, which confirms that the policy variation provides us a valid empirical setting to evaluate the effect of bilingual education. In addition, we do not see any bunching of the districts or school campuses around the 20 ELL student cutoff. A potential threat to the empirical strategy is that the districts might strategically adjust the number of ELL students so they may or may not have to provide bilingual education. However, figure 3 shows no evidence of such a strategic behavior both at the district and school campus level. The distribution of the districts is smooth along the number of ELL students in each district.

Returning to Table 3, we observe that the probability of complying to the policy has changed over time. After 2008, which is the last first grade cohort in our sample, the jump in the probability at the cutoff falls significantly from around 0.2-0.3 to less than 0.06. Figure 1 plots the point estimate of δ^{FS} in equation 1 and its 90% confidence interval. In 2010's, we cannot reject that the jump in the probability at the cutoff is statistically significantly away from zero. In some cases, the point estimate is even below zero, which means it is less likely that a district provides bilingual education when it has more than 20 ELL students who speak Spanish as their home language.

A potential explanation for the decrease in the compliance to the policy lies in funding deficiency. In Texas, the source of state funding for all public school districts is the Texas Foundation School Program (FSP). FSP allocates the state funding primarily based on the average daily attendance of students. Moreover, FSP adjusts the allocation based on the district and student characteristics to meet the needs of districts. For instance, having more ELL students will allocate more funding to the district. However, Rolle and Jimenez-Castellanos (2014) concludes that the FSP formula was not enough to balance out another big source of revenue to the school districts: the revenue generated from local property values. According to their analyses, the districts with the highest percentage of ELL students generated on average $1,300 less per pupil than districts with the lowest percentage of ELL students[6]. In fact, Texas recently passed House Bill 3 (HB 3) by the 86th Texas Legislature

[6]The total difference amounts to approximately $5.9 million, which translates to funding the salaries and benefits for 90 Texas teachers annually (Rolle and Jimenez-Castellanos (2014)).

in 2019, where the State upward adjusted the weight for ELL students in the FSP formula, which was originally set in 1984 and never changed, to better fund bilingual education.

5.2 Test Scores

In this section, we examine the effect of exposure to bilingual education in the first grade on elementary and middle school test scores. Table 7-9 report the reduced form and IV effects of bilingual education exposure on test scores, along with the first-stage results for IV. Throughout Section 5, we will focus on the IV estimates (δ^{SS}), as they are the most relevant from the policy-maker's perspective. Different columns correspond to different schooling levels (elementary or middle school) and different subjects. Tables 8-9 dissect the sample into different groups of students (Spanish speaking ELL and non-Spanish non-ELL) to see if there are heterogeneous effects across the groups.

Across the schooling levels and different groups of students, we do not find any statistically significant effect of bilingual education on test scores[7]. Given that the results are noisy, if anything, bilingual education has a negative effect on elementary school test scores and a positive effect on middle school test scores. Based on the IV estimates, bilingual education on average decreases elementary math test scores by about 0.082 standard deviation, and decreases elementary reading test scores by about 0.021 standard deviation. However, in middle school, bilingual education on average increases math test scores by about 0.051 standard deviation, and increases reading test scores by about 0.052 standard deviation.

The effects are larger to ELL students than to non-ELL students, although we cannot reject that the effects are statistically significantly different from zero. For ELL students (Table 8), based on the IV estimates, bilingual education decreases elementary math scores by 0.105 standard deviation, and elementary reading scores by 0.130 standard deviation. In middle school, however, the positive effect is also larger for ELL students, where the math scores increase by 0.299 standard deviation, and the reading scores increase by 0.261 standard deviation. For non-ELL students (Table 9), the IV estimates for elementary math scores and reading scores are -0.082 and -0.019, respectively. In middle school however, the estimates become positive: 0.058 for math scores and 0.049 for reading scores.

[7] Chin et al. (2013) find more optimistic results of bilingual education on elementary school test scores. The average elementary math scores increase by about 0.206 standard deviation and reading scores increase by about 0.143 standard deviation, on average. Their estimates are also noisy, where the effect on math test scores is statistically significant at the 10% level. They cannot reject that the effect on reading test scores is statistically different from zero at the 10% significance level. The effect is larger for non-ELL students in their work. A few potential factors that contribute to these differences are that (1) we look at a larger sample (from 1994 cohort rather than from 2003 cohort), and that (2) our analyses are at the individual level, so individuals are assigned to the 1st grade school they actually went to.

5.3 Long-Run Outcomes

We now turn to how the exposure to bilingual education in the first grade affect long-run educational outcomes, such as high school graduation and college enrollment. Although bilingual education may have limited effect on short-run test scores, the effect may manifest in the long-run either in terms of academic skills or through intangible social capital. We discuss the potential mechanisms of bilingual education which may differ for the short-run and long-run outcomes in Section 6. Table 4-6 present the reduced form and IV estimates of the effect of bilingual education exposure in the 1st grade on high school graduation and college enrollment. Again, each column corresponds to a different outcome variable and we focus on the IV estimates which are the most policy relevant. Table 5-6 are the results for the separate subgroups of the sample, Spanish speaking ELL and non-Spanish non-ELL students.

To begin with, there is little to no effect of bilingual education on high school graduation. Column (1) of Table 4 reports, on average, a less than 1% increase in high school graduation rate as a result of bilingual education, which is statistically insignificant. However, the unconditional mean of high school graduation rate of the sample is pretty high, which hovers around 75%. Even for the subgroup of ELL students, high school graduation rate is over 74%. The small gap in high school graduation rate between ELL and non-ELL students implies that the ability to communicate in English is perhaps not the crucial factor that determines high school graduation, and thus it is not surprising that bilingual education has negligible effect on high school graduation.

On the other hand, bilingual education has a sizable effect on college enrollment, especially on a 4-year university enrollment. The IV estimate of column(2) of Table 4 reports about 12% increase in college enrollment rate, although statistically insignificant. This positive effect is mostly driven by the 4-year university enrollment rate. Column (3) of Table 4 shows that the enrollment to a 4-year university statistically significantly goes up by 28%. Community college enrollment rate also goes up by 13% (Column(4) of Table 4), but we cannot reject that it is statistically different from zero at the 10% significance level.

Interestingly, non-ELL students benefit more from bilingual education in terms of college enrollment. This confirms that there exists a positive spillover effect of bilingual education to non-ELL students. According to column (2) of Table 6, college enrollment on average goes up by 16% for non-ELL students, if they went to a school which provides bilingual education in their 1st grade. Again, the effect is driven by increase in the enrollment to a 4-year university, which is 30% higher when a non-ELL student goes to a school where ELL students were educated in separate classrooms in the 1st grade (column (2) of Table 6). The effect on community college enrollment is noisier, but the enrollment rate goes up

by 16% (column (3) of Table 6). For ELL students, the positive effect of bilingual education on college enrollment is more concentrated to the 4-year university enrollment rate. Column (3) of Table 5 reports almost 50% increase in the 4-year university enrollment rate of ELL students as a result of the exposure to bilingual education in their 1st grade. However, community college enrollment drops by 4%, although the estimate is highly noisy (column (4) of Table 5).

6 Mechanisms: Effectiveness of Bilingual Education

In this section, we discuss mechanism of why a bilingual education program would impact outcomes for ELL and non-ELL students both in the short-run and long-run. An obvious mechanism by which bilingual education impacts ELL students is that they receive instruction in their native language. This should improve the student's ability to understand classroom content, especially in subjects such as math and science. Conversely, bilingual education may slow an ELL student's English acquisition because the student interacts with the teacher in the student's native language rather than English.

Bilingual education teachers may improve ELL student outcomes because they have a similar race or background to the students compared to a standard teacher. Previous research has found that having a teacher with a similar background can improve student short-run outcomes such as test scores and behavior (Dee (2004), Fairlie et al. (2014), Egalite et al. (2015)). This may be because these teachers better understand how to educate student that have the same background as them or because teachers act as role models. Having a teacher with a similar background has also been shown to improve student's long-run outcomes such as high school graduation and college enrollment (Delhommer (2022), Gershenson et al. (2018)). Long-run outcomes could be improved by bilingual education teachers for ELL students if these teachers improves skills that are not well measured by contemporaneous academic achievement or because teachers increase student's expectations about the return to education or their own ability.

Furthermore, exposure to bilingual education vastly changes the classroom composition that students experience. Under and ESL program, ELL and non-ELL students are mixed in the same classroom. Under a bilingual education program, ELL students will be placed in a separate classroom to non-ELL students. ELL students on average come from a lower socioeconomic background and have lower test scores compared to their non-ELL peers, so bilingual education reduces the average quality of an ELL student's classmates while increasing the average quality of the non-ELL student's classmates. Previous research has explored how the quality of a student's peers can impact the students' outcomes finding mixed

results (Sacerdote (2011)). In addition to test-scores, interacting with English proficient peers may be an important channel by which ELL students learn English. When separated into the bilingual education classroom, ELL students will have no classmates who are fully proficient in English, which potentially hampers their acquisition of English.

The long-run impacts of bilingual education in elementary school may be different than in the short-run. The goal of bilingual education is for ELL students to return to the standard classroom. This return to the standard classroom in a substantially different learning environment, and this abrupt change may have negative impacts on ELL students in the long-run, even if bilingual education may raise ELL student test scores in the short-run.[8]. Bilingual education may also provide students with improved abilities or desire to learn that does not appear in contemporaneous test scores, but does appear in future test scores and future outcomes such as college enrollment. Previous research has shown that some early education intervention can have limited lasting impacts on student test scores but long-run improvements in students future outcomes (Bailey et al. (2021)).

7 Conclusion

In this paper, we have analyzed how school districts' provision of bilingual education vs. ESL in first grade impacts student outcomes both in the short-run (elementary school test scores) and long-run (middle school test scores, high school graduation, and college enrollment). Our empirical strategy leverages policy variation in Texas that requires school districts with 20 or more ELL students with the same home language in the same elementary grade to provide bilingual education; schools below this cut-off are able to choose bilingual education or ESL programs for their ELL students. Students above and below this cut-off should be similar except that students above the cut-off are more likely to have been at a school district that provided bilingual education. In our analysis, we find that bilingual education has no observable impacts on elementary school test scores but improves four-year college enrollment. These results suggest that bilingual education in elementary grades improves students' ability that is not well measured by elementary school tests but appears in high school and beyond. Additionally, positive effects are found for both ELL and non-ELL students, suggesting that peer effects within the classroom between ELL and non-ELL students are important mechanisms impacting student outcomes.

We are in the process of analyzing additional outcomes. One goal of both ESL and

[8]Previous research has found mixed results on changing school environments on student outcomes in the context of school closings (Brummet (2014)), grade promotion (Rockoff and Lockwood (2010)), and voluntary student moves Hanushek et al. (2004)).

bilingual education is ELL students to obtain English proficiency, and understanding under which learning environment do ELL students learn English faster is important for evaluating these programs. Additionally, while our results highlight that bilingual education increases college enrollment, we plan to look at whether this continues through to positive effects on college completion and future labor earnings.

This study and most previous studies have focused on the effectiveness of ESL compared to bilingual education, but schools may choose a variety of bilingual education programs to implement. Bilingual education programs differ in the speed in which ELL students are transitioned back to the standard classroom (early vs. late exit programs). They also differ in the extent to which they promote proficiency in the native language (transitional vs. dual language programs). Some bilingual education programs even combine native and non-native English students into the same classrom (two-way immersion programs). Understanding the effectiveness of these different modes of bilingual education is an important next step for future research, especially for large school districts with many ELL students.

References

Anantha, S. (2024). How to become a certified esl teacher in texas. https://texestest.org/texas-teacher-certification/texas-esl-teachers/. Accessed: 2024-02-16.

Bailey, M. J., Sun, S., and Timpe, B. (2021). Prep school for poor kids: The long-run impacts of head start on human capital and economic self-sufficiency. *American Economic Review*, 111(12):3963–4001.

Brummet, Q. (2014). The effect of school closings on student achievement. *Journal of Public Economics*, 119:108–124.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., and Yagan, D. (2011). How does your kindergarten classroom affect your earnings? evidence from project star. *The Quarterly journal of economics*, 126(4):1593–1660.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014). Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American economic review*, 104(9):2633–79.

Chin, A., Daysal, N. M., and Imberman, S. A. (2013). Impact of bilingual education programs on limited english proficient students and their peers: Regression discontinuity evidence from texas. *Journal of Public Economics*, 107:63–78.

Collier, V. P. and Thomas, W. P. (2017). Validating the power of bilingual schooling: Thirty-two years of large-scale, longitudinal research. *Annual review of applied linguistics*, 37:203–217.

Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *Review of economics and statistics*, 86(1):195–210.

Delhommer, S. (2022). High school role models and minority college achievement. *Economics of Education Review*, 87:102222.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from head start. *American Economic Journal: Applied Economics*, 1(3):111–34.

Egalite, A. J., Kisida, B., and Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, 45:44–52.

Fairlie, R. W., Hoffmann, F., and Oreopoulos, P. (2014). A community college instructor like me: Race and ethnicity interactions in the classroom. *American Economic Review*, 104(8):2567–91.

Gershenson, S., Hart, C. M., Hyman, J., Lindsay, C., and Papageorge, N. W. (2018). The long-run impacts of same-race teachers. Technical report, National Bureau of Economic Research.

Gray-Lobe, G., Pathak, P. A., and Walters, C. R. (2021). The long-term effects of universal preschool in boston. Technical report, National Bureau of Economic Research.

Hanushek, E. A., Kain, J. F., and Rivkin, S. G. (2004). Disruption versus tiebout improvement: The costs and benefits of switching schools. *Journal of public Economics*, 88(9-10):1721–1746.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, 142(2):615–635.

Jackson, C. K. (2018). What do test scores miss? the importance of teacher effects on non–test score outcomes. *Journal of Political Economy*, 126(5):2072–2107.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2):281–355.

Lopez, B. (2022). Texas schools are majority hispanic. there's been a shortage of bilingual teachers since 1990, and the pandemic made it worse. *The Texas Tribune*.

Matsudaira, J. D. (2005). Sinking or swimming? *Evaluating the impact of English immersion versus bilingual education. Berkeley: University of California*.

NCES (2018). The nation's report card. https://www.nationsreportcard.gov/math_2017/#/nation/achievement?grade=4 and https://www.nationsreportcard.gov/reading_2017/#/nation/achievement?grade=4. Accessed: 2024-01-30.

NCES (2023). *The Condition of Education 2023*, chapter English Learners in Public Schools. U.S. Department of Education, Institute of Education Sciences, https://nces.ed.gov/programs/coe/indicator/cgf.

Pages, R., Lukes, D. J., Bailey, D. H., and Duncan, G. J. (2020). Elusive longer-run impacts of head start: Replications within and across cohorts. *Educational Evaluation and Policy Analysis*, 42(4):471–492.

Petek, N. and Pope, N. (2016). The multidimensional impact of teachers on students. Technical report, University of Chicago Working Paper.

Pope, N. G. (2016). The marginal effect of k-12 english language development programs: Evidence from los angeles schools. *Economics of Education Review*, 53:311–328.

Rockoff, J. E. and Lockwood, B. B. (2010). Stuck in the middle: Impacts of grade configuration in public schools. *Journal of public economics*, 94(11-12):1051–1061.

Rolle, R. A. and Jimenez-Castellanos, O. (2014). An efficacy analysis of the texas school funding formula with particular attention to english language learners. *Journal of Education Finance*, 39(3):203–221.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In *Handbook of the Economics of Education*, volume 3, pages 249–277. Elsevier.

Texas Education Agency (2022a). English as a second language (esl) program implementation rubric. https://tea.texas.gov/academics/special-student-populations/english-learner-support/esl-program-implementation-rubric.pdf. Accessed: 2024-02-16.

Texas Education Agency (2022b). Transitional bilingual education (tbe) program implementation rubric. https://tea.texas.gov/academics/special-student-populations/english-learner-support/tbe-program-implementation-rubric.pdf. Accessed: 2024-02-16.

Texas Teachers of Tomorrow (TTOT) (2023). How to become a bilingual teacher in texas. https://www.teachersoftomorrow.org/texas/teacher-resources/high-need-areas/bilingual/. Accessed: 2024-02-16.

Williams, C. and Zabala, J. (2023). How to grow bilingual teacher pathways: Making the most of u.s. linguistic and cultural diversity. https://tcf.org/content/report/how-to-grow-bilingual-teacher-pathways-making-the-most-of-u-s-linguistic-and-cultural Accessed: 2024-02-16.

Wright, W. E. (2010). Foundations for teaching english language learners: Research, theory. *Policy, and Practice. Philadelphia: Caslon Publishing*, pages 70–76.

A Tables and Figures

Table 1: School District Information in 2005
by District Size and Number of Spanish ELL Students

	≤200		>200	
	8-39	>39	8-39	>39
Avg. Number of Students	98.028	122.455	322.468	1,562.324
S.D. Number of Students	48.833	38.836	113.399	2,123.043
Avg. Number of Campuses	1.099	1.091	2.899	15.894
S.D. Number of Campus	0.351	0.302	1.972	22.269
Pct. District Provides Bilingual Education	0.271	0.727	0.443	0.953
Pct. Spanish LEP Students	0.225	0.593	0.068	0.267
Pct. FRL Students	0.618	0.764	0.475	0.500
Pct. Female	0.477	0.459	0.481	0.481
Pct. Black	0.078	0.039	0.115	0.127
Pct. Hispanic	0.511	0.804	0.271	0.526
Pct. White	0.404	0.154	0.592	0.321
Pct. Community College Enroll	0.417	0.388	0.237	0.076
Pct. University Enroll	0.195	0.199	0.120	0.034
Pct. Any College	0.463	0.440	0.270	0.085
Avg. Grade 3 Reading Test Score	-0.069	-0.209	0.142	0.019
S.D. Grade 3 Reading Test Score	0.312	0.165	0.254	0.249
Avg. Grade 3 Math Test Score	-0.116	-0.133	0.093	0.018
S.D. Grade 3 Math Test Score	0.327	0.149	0.246	0.242
Districts	181	11	79	170
First Grader Obs.	17,743	1,347	25,475	265,595

Note:

Table 2: School District Information in 2005
by District Size and Number of Spanish ELL Students

	1994-1998	1999-2008	2009-2019
Avg. Number of Students	101.879	99.260	102.674
S.D. Number of Students	52.842	50.456	64.749
Avg. Number of Campuses	1.123	1.105	1.243
S.D. Number of Campus	0.486	0.425	0.860
Pct. District Provides Bilingual Education	0.213	0.262	0.267
Pct. Spanish LEP Students	0.200	0.214	0.219
Pct. FRL Students	0.639	0.610	0.623
Pct. Female	0.483	0.484	0.486
Pct. Black	0.073	0.084	0.093
Pct. Hispanic	0.506	0.497	0.500
Pct. White	0.416	0.410	0.376
Pct. Community College Enroll	0.443	0.438	NA
Pct. University Enroll	0.220	0.201	NA
Pct. Any College	0.483	0.485	NA
Avg. Grade 3 Reading Test Score	-0.022	-0.058	-0.032
S.D. Grade 3 Reading Test Score	0.248	0.279	0.465
Avg. Grade 3 Math Test Score	-0.021	-0.117	-0.072
S.D. Grade 3 Math Test Score	0.265	0.307	0.446
Districts	595	1,659	2,321
First Graders Obs.	57,678	154,771	229,473

Note:

Table 3: First Stage Coefficients

	1994-1998	1994-1998	1999-2008	1999-2008	2009-2019	2009-2019
Coef., No Controls	0.292	0.336	0.260	0.215	0.031	0.059
St.E., No Controls	0.086	0.091	0.056	0.062	0.064	0.056
Coef, Controls	0.260	0.263	0.254	0.215	0.031	0.040
St.E., Controls	0.080	0.079	0.049	0.055	0.053	0.050
Districts	595	595	1,659	1,659	2,321	2,321
Campuses	668	668	1,833	1,833	2,902	2,902
Students	60,614	60,614	164,671	164,671	240,998	240,998

Note:

Table 4: Effect of Bilingual Education on Long-Run Outcomes

	(1)	(2)	(3)	(4)
IV Coefficient	0.007	0.060	0.059*	0.058
IV Standard Error	0.036	0.040	0.033	0.045
9				
FS Coefficient	0.231***	0.231***	0.231***	0.231***
FS Standard Error	0.047	0.047	0.047	0.047
9.1				
RF Coefficient	0.002	0.014	0.014*	0.013
RF Standard Error	0.008	0.009	0.007	0.010
9.2				
Outcome Mean	0.751	0.505	0.212	0.461
Observations	823,660	823,660	823,660	823,660
9.3				
Outcome	HS Graduation	Any College Enroll.	4-year Univ. Enrollment	CC Enrollment
Sample	All	All	All	All

*Note: *=10% Significance Level, **=5% Significance Level, ***=1% Significance Level*

Table 5: Effect of Bilingual Education on Long-Run Outcomes

	(1)	(2)	(3)	(4)
IV Coefficient	-0.052	-0.012	0.064*	-0.013
IV Standard Error	0.058	0.059	0.038	0.062
FS Coefficient	0.263***	0.263***	0.263***	0.263***
FS Standard Error	0.044	0.044	0.044	0.044
RF Coefficient	-0.014	-0.003	0.017*	-0.003
RF Standard Error	0.015	0.016	0.009	0.016
Outcome Mean	0.742	0.391	0.130	0.360
Observations	127,013	127,013	127,013	127,013
Outcome	HS Graduation	Any College Enroll.	4-year Univ. Enrollment	CC Enrollment
Sample	Spanish, LEP	Spanish, LEP	Spanish, LEP	Spanish, LEP

*Note: *=10% Significance Level, **=5% Significance Level, ***=1% Significance Level*

Table 6: Effect of Bilingual Education on Long-Run Outcomes

	(1)	(2)	(3)	(4)
IV Coefficient	0.025	0.082*	0.068*	0.079
IV Standard Error	0.039	0.048	0.040	0.052
FS Coefficient	0.221***	0.221***	0.221***	0.221***
FS Standard Error	0.050	0.050	0.050	0.050
RF Coefficient	0.005	0.018*	0.015*	0.017
RF Standard Error	0.009	0.010	0.008	0.011
Outcome Mean	0.752	0.526	0.227	0.480
Observations	668,247	668,247	668,247	668,247
Outcome	HS Graduation	Any College Enroll.	4-year Univ. Enrollment	CC Enrollment
Sample	Non-Spanish, Non-LEP	Non-Spanish, Non-LEP	Non-Spanish, Non-LEP	Non-Spanish, Non-LEP

*Note: *=10% Significance Level, **=5% Significance Level, ***=1% Significance Level*

Table 7: Effect of Bilingual Education on Long-Run Outcomes

	(1)	(2)	(3)	(4)
IV Coefficient	-0.082	-0.021	0.051	0.052
IV Standard Error	0.080	0.075	0.095	0.068
9				
FS Coefficient	0.231***	0.231***	0.232***	0.232***
FS Standard Error	0.047	0.047	0.047	0.047
9.1				
RF Coefficient	-0.020	-0.005	0.012	0.012
RF Standard Error	0.019	0.018	0.022	0.016
9.2				
Outcome Mean	-0.060	-0.039	-0.048	-0.017
Observations	808,221	808,221	620,285	620,285
9.3				
Outcome	Elementary Math	Elementary Reading	Middle Math	Middle Reading
Sample	All	All	All	All

*Note: *=10% Significance Level, **=5% Significance Level, ***=1% Significance Level*

Table 8: Effect of Bilingual Education on Long-Run Outcomes

	(1)	(2)	(3)	(4)
IV Coefficient	-0.105	-0.130	0.299	0.261
IV Standard Error	0.181	0.181	0.254	0.288
FS Coefficient	0.262***	0.262***	0.256***	0.256***
FS Standard Error	0.046	0.046	0.051	0.051
RF Coefficient	-0.029	-0.035	0.081	0.070
RF Standard Error	0.049	0.048	0.068	0.077
Outcome Mean	-0.488	-0.550	-0.355	-0.469
Observations	95,783	95,783	35,552	35,552
Outcome	Elementary Math	Elementary Reading	Middle Math	Middle Reading
Sample	Spanish, LEP	Spanish, LEP	Spanish, LEP	Spanish, LEP

*Note: *=10% Significance Level, **=5% Significance Level, ***=1% Significance Level*

Table 9: Effect of Bilingual Education on Long-Run Outcomes

	(1)	(2)	(3)	(4)
IV Coefficient	-0.082	-0.019	0.058	0.049
IV Standard Error	0.084	0.078	0.101	0.073
FS Coefficient	0.222***	0.222***	0.225***	0.225***
FS Standard Error	0.049	0.049	0.049	0.049
RF Coefficient	-0.019	-0.005	0.013	0.011
RF Standard Error	0.020	0.018	0.023	0.016
Outcome Mean	-0.014	0.019	-0.024	0.018
Observations	665,756	665,756	523,281	523,281
Outcome	Elementary Math	Elementary Reading	Middle Math	Middle Reading
Sample	Non-Spanish, Non-LEP	Non-Spanish, Non-LEP	Non-Spanish, Non-LEP	Non-Spanish, Non-LEP

*Note: *=10% Significance Level, **=5% Significance Level, ***=1% Significance Level*

Figure 1: First Stage Over Time

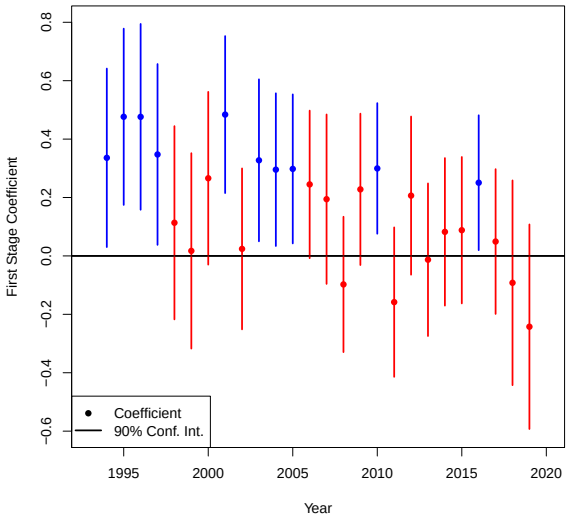

Figure 2: Discontinuous Change in Probability of Providing Bilingual Education at the Cut-off

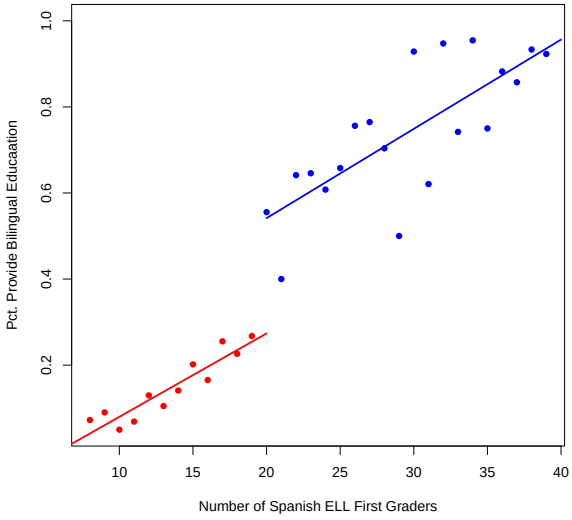

Figure 3: McCrary Figure