Dual Language Education and Student Achievement

Andrew Bibler*

January 19, 2016

Abstract

Dual language classrooms provide English Language Learners (ELLs) an opportunity to receive instruction in their native language as they transition to English fluency. This might allow ELLs to build a stronger foundation in core subjects and lead to better academic outcomes. Dual language education has also grown substantially in popularity among English speaking families across the U.S., as it presents an option to learn content in, and presumably become fluent in, a second language. Despite the spike in practice, there is little causal evidence on what effect attending a dual language school has on student achievement. I examine dual language education and student achievement using school choice lotteries from Charlotte-Mecklenburg School District, finding local average treatment effects on math and reading exam scores of more than 0.06 standard deviations per year for participants who were eligible for English second language (ESL) services or designated limited English proficient (LEP). There is also some evidence that attending a dual language school led to a lower probability of having limited English proficient status starting in third grade. For applicants who were not eligible for ESL services or designated as LEP, attending a dual language school has resulted in higher end of grade exam scores of about 0.09 and 0.05 standard deviations per year in math and reading, respectively.

*I would like to thank Todd Elder for his guidance and suggestions throughout this project. I would also like to thank Stacy Dickert-Conlin, Scott Imberman, Madeline Mavrogordato, and Kelly Vosters for their thoughtful comments, as well as seminar participants at Michigan State University, the 2015 Annual Conference for the Association for Education Finance and Policy, and the 2015 Annual Conference for the Midwest Economics Association. I am also grateful to the North Carolina Education Research Data Center (NCERDC) and Charlotte-Mecklenburg School District (CMS) for providing the data. This research is supported by a Pre-Doctoral Training Grant from the IES, U.S. Department of Education (Award #R305B090011) to Michigan State University. The opinions expressed here are those of the author and do not represent the views of the U.S. Department of Education, NCERDC, or CMS.

1 Introduction

Dual language classrooms use a non-English language of instruction for a significant amount of the curriculum. They are primarily used to provide instruction to English language learners (ELLs) in their first language and to promote bilingualism and biculturalism among native English speakers. There are two types of dual language classrooms. *Two-way* classrooms typically enroll students from two different language backgrounds and teach curriculum in both languages. There were only about ten such programs in the U.S. in 1980, but that number was almost 250 by 2000 [Howard and Sugarman, 2001]. In contrast, most students in a *one-way* (*immersion*) classroom share a similar language background, but receive instruction in a second language. The number of one-way classrooms registered with the Center for Applied Lingustics increased from fewer than 50 to almost 450 over the last few decades (Center for Applied Linguistics, 2011). Recent expansions in several states have driven these numbers even higher.[1] Despite the growth, there is little causal evidence on the effect of dual language education on student achievement. In this paper, I use school choice lotteries from Charlotte-Mecklenburg School District (CMS) to estimate the effect of attending a dual language school on student achievement.

Districts target dual language education to two types of students: English language learners (ELLs) who might benefit from receiving instruction in their home language, and other students who want to have instruction in a second language. For ELLs, dual language education might allow for an easier transition to full English instruction, providing a potential route for improving outcomes of the growing and struggling ELL population. The alternative is often placement in an English-only classroom coupled with English second language (ESL) services, which might mean missing important classroom instruction time and disruption to the student and his or her peers, and ultimately making students more likely to fall behind. On the other hand, placement in an English-only classroom might expedite the development of English skills, leading to faster re-classification out of ELL status and higher scores on standardized exams that are written in English.

[1]Utah passed legislation for funding of dual language programs in 2008, and since then has implemented programs in 118 schools in 22 districts. New York City added or expanded 40 programs in 2015.

Districts also target dual language education to English speaking students who desire to learn in a second language, and the influx of dual language programs seems to be driven in large part by demand from English speaking families [Watanabe, 2011].[2] The primary goal for districts in offering dual language education to English speakers is to provide an option that allows them to become bilingual, biliterate, and bicultural. In addition, districts can use dual language schools, as well as other specialized programs, to offer a more diverse set of schooling options and compete with charter and private schools to retain students residing in the district boundaries. Dual language programs are often promoted using high test performance of participants as evidence of increased cognitive development [Maxwell, 2012, Maxwell, 2014]. However, lack of formal training in English could slow progress as measured by scores on standardized exams. It is unclear whether dual language programs would increase or decrease test scores for this group of students.

Whether dual language education has any effect on student achievement and the direction of those effects are empirical questions, but there is very little causal evidence due to endogeneity from self selection into the programs. In this paper, I estimate the causal effect of attending a dual language school on achievement on standardized math and reading exams by exploiting quasi-random assignment from oversubscribed admissions lotteries. I focus on students who applied through the Charlotte-Mecklenburg school choice lottery for their kindergarten year, and specified a dual language school as their first choice. I use quasi-random assignment to a dual language school through the lottery as an instrument for dual language school attendance to identify the local average treatment effect of dual language schooling on achievement. The treatment differs by whether or not the student uses English as a home language. For a native English speaker, the treatment is to receive instruction in a second language and the alternative to attending a dual language school is receiving instruction in their home language. For ELLs, the typical alternative to attending a dual language school is to receive instruction in English (not their home language) accompanied by other ESL services. Because of this divide in treatment, I estimate effects separately for

[2] About 70 percent of the estimation sample in this study are students who were never identified as English language learners or limited English proficienct in the data.

two subgroups using a proxy for whether or not the student was proficient in English when they entered school.

The first group is made up of students who were eligible for English second language (ESL) services or were designated as limited English proficient (LEP)[3] at any point in the data. I will refer to this group of students as the "ESL/LEP" sample. In the ESL/LEP sample, I find that attending a dual language school leads to increased scores on math and reading exams of more than 0.06 standard deviations per year of participation. Furthermore, I find some evidence that attending a dual language school has led to a lower probability of being designated as limited English proficient in grades three through six for students in the ESL/LEP sample. The second subgroup is made up of students who were never eligible for ESL services or designated LEP. I will refer to this group as the "English" or the "non-ESL/LEP" sample. Among this group, I estimate that attending a dual language school leads to 0.09 standard deviations higher achievement in math per year of participation, and 0.05 standard deviations higher achievement in reading per year. The estimates are statistically significant and generally robust to a number of alternate specifications.

1.1 Literature

Bilingual education broadly refers to educational programs that are targeted toward ELLs and include some amount of home language instruction. Drawing conclusions from previous literature is complicated by the fact that bilingual education can take several forms, and the degree to which home language instruction is used varies within and across program types. Two-way dual language classrooms tend to use the non-English language for a large proportion of instruction (50% or more) throughout elementary school. Instruction is not generally based on current English ability of ELL students, as two-way classrooms enroll dominant speakers of both languages and provide instruction in both languages. Other

[3]North Carolina uses the term limited English proficient (LEP) to refer to students who do not use English as a primary language in their home, and score below a specified cutoff on an English skills test. There is some variation among researchers and school districts in how they refer to this group of students. The term English language learner (ELL) is often used in place of, or interchangeably with LEP. I use the term LEP when referring to the designation given to students in North Carolina because that is the term the state uses, but use the terms ELL and LEP interchangeably when referring to this group of students generally.

forms of bilingual education, such as transitional bilingual education and structured English immersion, are more focused on expediting English fluency and do not necessarily group ELLs and non-ELLs in the same classroom.[4]

Some prior research focuses on the achievement gap for ELL students in dual language (DL) programs, showing that ELLs participating in DL programs have higher test scores than ELLs in non-dual language classrooms. Two detailed reports on six districts in North Carolina, including CMS, find that ELLs in two-way programs score higher than students in English-only classrooms on end-of-grade exams [Thomas and Collier, 2009, Thomas et al., 2010]. Collier and Thomas [2004] summarize 18 years of results on one- and two-way programs from 23 different school districts. Students in both program types close at least 70% of the ELL test score gap by the end of fifth grade, but this could be driven by self selection.

One technique sometimes employed in an attempt to overcome the self selection issue is matching on pretest scores or other observable characterisitcs. Cazabon, Lambert, and Hall [1999] studied a two-way program in Cambridge, MA. They used a pretest to match each DL student with a control, and show that ELLs assigned to a dual language classroom outperformed the control group on English based math and reading exams. Cobb, Vega, and Kronauge [2009] also match students on observable characteristics to consider the impact of DL education on achievement and find positive effects in writing and math for native Spanish speakers, with the effects being more pronounced one year after completion of the program.

Several attempts have been made by researchers to summarize estimates from the most methodologically rigorous studies. In a meta-analysis on transitional bilingual education programs, Rossell and Baker [1996] deemed only 25 percent of the studies they considered to be methodologically acceptable. A transitional bilingual education model teaches reading in the native language in early grades, but moves to complete English instruction as early as second grade. They determined that a relatively small percentage of the most rigorous studies estimate positive effects of transitional bilingual programs. While other meta-analyses agree that much of the prior research is flawed, they suggest that among acceptable stud-

[4] See Valentino and Reardon [2015] for a good description of some of the differences between dual language classrooms and other forms of bilingual education.

ies there are positive effects of bilingual programs across subjects and in different types of programs [Greene, 1998, Slavin and Cheung, 2005, Willig, 1985]. Most of the literature they examined does not account for selection bias, but there are a handful of studies that used random assignment in an attempt to estimate causal effects. However, these studies were generally based on small samples (e.g. less than 175 students) and from nearly 30 years ago [Greene, 1998].

More recently, Valentino and Reardon [2015] use data on student preferences from a large urban district to compare the test scores of students in bilingual education and English-only programs. They study student performance on exams across program types conditional on the type of program that the student preferred. The assignment is quasi-random, but they do not use knowledge of the assignment mechanism to completely exploit the randomness in the assignment. They find that dual language students progress faster in math and English language arts performance after second grade, leading to better long-run performance than any of the other three programs (including English immersion) [Valentino and Reardon, 2015]. Similar to this study, Steele et al. [2015] exploit random assignment from oversubscribed admissions lotteries into dual language programs in Portland, Oregon. They report mostly positive, insignificant effects for dual language students on reading and math exam scores. However, there are two important differences in this study. First, they pool two subgroups of students - ELLs and non-ELLs - together, despite the fact that the treatment effects for these two groups could be of different signs and magnitudes, which would have important policy implications.[5] Another important difference between the Steele et al. study and this paper is that the dual language programs in Portland Public Schools (PPS) are strand programs, meaning that they only make up a portion of the school. All of the dual language programs in CMS are housed in three schools, where every classroom in the school is a dual language classroom.

Other recent studies have examined causal effects of bilingual education programs, but the classrooms in these studies are not necessarily two-way dual language programs. Slavin

[5] About nine percent of students are ELLs at the time of application and fifteen percent have a non-English home language [Steele et al., 2015].

et al. [2011] use random assignment in kindergarten to either an English immersion or transitional bilingual classroom to study differences in English and Spanish reading scores for several years following the assignment. They found that students assigned to a transitional classroom scored lower on English reading exams in early grades, but there were no statistically significant differences by fourth grade. Guo and Koretz [2013] use a difference-in-differences framework to study the effect of a Massachusetts policy that shifted the early elementary education for ELLs from a several year transitional bilingual model to a one-year sheltered (or structured) English immersion model. Structured models target instruction to the current English ability of the students while expediting English fluency relative to transitional models, so this represents a clear shift away from instruction in the home language of the students. Similar to the findings of Slavin et al. [2011], Guo and Koretz find that the policy had no effect (or a small positive effect) on fourth grade English reading scores. Chin et al. [2013] use district level variation in the number of LEP students in Texas to study whether having a bilingual education option improves achievement for LEPs and their non-LEP peers. They identify the treatment effects using the discontinuity generated by a Texas rule that districts with at least twenty LEP students who share a common language in a specific grade must offer a bilingual education option to those students. They do not find significant increases in the test scores of LEP students from districts that offer bilingual education, but do find an increase in the scores of non-LEPs in districts that offer bilingual programs. Their findings suggest that offering bilingual education resulted in positive spillover effects to non-LEP students. Estimating peer effects directly has rarely been done in this setting and with mixed evidence [Cho, 2012, Geay et al., 2013].

While test scores are one outcome of interest, districts may also care about the duration of LEP classification of language minority students. When a student enrolls in a district in North Carolina, their parent takes a survey that asks about the languages the student uses at home. The district uses that survey, and possibly interviews with the parents and/or student, to determine the home language of the student. If the home language of the student is not English, then the student must take a test that determines LEP status and eligibility for ESL services. When a student is identified as LEP based on the score of the placement test, they are required to continue testing annually until they are re-classified out of LEP status. LEP classification is important for several reasons. It is another measure of student progress that differs from the math and reading exams. Second, students with LEP status may be eligible for testing accommodations. Lastly, offering ESL services is costly, so districts benefit from programs that expedite reclassification, all else equal. Umansky and Reardon [2014] show that dual language participants in a large urban district are reclassified out of LEP status at a slower rate in early grades, but have higher total reclassification and English proficiency than students from English immersion classrooms by the end of high school. Similarly, Steele et al. [2015] report slower reclassification out of ELL status for dual language participants throughout elementary and middle school. When estimating treatment effects though, they find that attending a dual language classroom led to a higher probability of exiting ELL status starting in fifth grade. In addition to estimating the effect of dual language schooling on achievement, I estimate the effect of dual language schooling on LEP classification among students who were ever identified as LEP or eligible for ESL services.

For native English speakers, there is concern that attending dual language schools may promote bilingualism at the expense of achievement as measured by standardized tests, which are written in English. The question for many parents is whether their child can attend a dual language school and become bilingual without falling behind in other subjects. Learning in a second language could create confusion or frustration that would negatively impact achievement, especially in the short-run. On the other hand, the mental juggling involved with thinking in two languages might promote cognitive development. This theoretical connection has previously been made in research related to working memory, which is used to store and process information and execute related tasks [Baddeley and Hitch, 1974, Baddeley, 2003, Alloway, 2010]. Working memory can be considered a measure of ability to learn. It is strongly correlated with academic outcomes and much of the growth in working memory capacity takes place before adolescence [Alloway, 2010]. Working memory is closely associated with second language acquisition [Baddeley, 2003], native language vocabulary [Dufva et al., 2001], and listening and reading comprehension [Chrysochoou et al., 2011, Dufva et al., 2001], but empirical evidence directly supporting a causal link between second language acquisition and cognitive development through working memory is sparse. In this specific setting, students apply for entry into a DL school for their kindergarten year but don't take their first high stakes exam until third grade, so students and teachers have some time to overcome any initial difficulties in adjusting to the new language. The gap in time between school assignment and testing allows teachers and administrators to commit to teaching in the second language in the first few years of school when there are no high stakes exams looming. If working in two languages can boost cognitive development, then one might expect it to show up in this environment.

Some prior literature has pointed out the positive achievement gap for English dominant students in dual language programs. English speaking participants of two-way programs in North Carolina score higher than their peers on end-of-grade exams and have better attendance [Thomas and Collier, 2009, Thomas et al., 2010]. Other research uses matching on pretests and observables, finding that learning a portion of curriculum in a second language does not hinder progress and might be associated with positive effects on achievement in reading [Cobb et al., 2009, Cazabon et al., 1999]. Again, Steele et al. [2015] find that dual language instruction led to positive, although often insignificant, effects on math and reading scores. Their effects are not directly comparable because they pool all students, but their sample is comprised of mostly native English speakers.

The most important contribution of this study is to provide causal estimates of the effect of attending a dual language school on achievement using school choice lotteries in Charlotte-Mecklenburg School District and, in particular, estimating separate treatment effects for two groups regarded very differently for policy purposes. This is a valuable addition to the existing research on bilingual education, because most prior literature does not credibly identify any treatment effect and does not disentangle effects for English language learners compared to native English speakers. The next section provides details on the lottery used in CMS. Section 3 discusses the data and some descriptives. In section 4 the empirical strategy used for the main results is discussed. Section 5 presents the empirical results and section 6 concludes.

2 Lottery

Every student enrolled in Charlotte-Mecklenburg School District is assigned to a neighborhood school based on geographic zones. The district uses a school choice lottery to allocate seats for students who wish to opt out of their neighborhood school. The empirical strategy used in this paper makes use of exogenous variation created from oversubscribed lotteries, so it is useful to describe how the lottery operates and why it facilitates the identification of treatment effects. This section provides details on the lottery.

2.1 Magnet Programs and Priority Groups

All CMS students can submit up to three programs in order of preference through a centralized lottery. All students with an older sibling in a school are guaranteed a seat in that school by making it their first choice.[6] Then non-guaranteed seats are assigned in three rounds. In the first round, only first choices are considered. If there are fewer applicants than seats available to a given program, then all of the applicants to that program will be assigned to their first choice. Identification comes from comparing winners and losers from the same lottery, so estimates are driven by oversubscribed lotteries. When the number of applicants is greater than the number of available seats (the choice is oversubscribed), seats are awarded quasi-randomly. Seat assignment is not completely random, because the probability of winning for a particular student depends on the priority group that the student is assigned to. Priority groups refer to sets of students that meet (or do not meet) some pre-specified criteria. In CMS, over the sample period they are based on geographic location and whether the student's neighborhood school is a Title I choice school.

With that in mind, the district gives priority with a couple of apparent goals. First,

[6] Students who meet admission criteria and have a twin or older sibling assigned to a magnet program receive guaranteed admission to that program. The applying student must specifiy it as their first choice in order to be guaranteed admission through sibling preference. The sibling guarantee requires that the students have the same residence and at least one common parent or guardian.

they care about transportation costs and allowing students to attend schools that are close to home. Students who live within close proximity to a full magnet school are given priority. In addition, the district is split into four geographic zones. Magnet schools offer transportation to at least one, and up to four of the zones, leading the zones to be referred to as transportation zones. Students who live in a zone served by a magnet are given priority for admission to that school over students who live in a zone that is not served by that school. Students outside of the zone can still apply, but living outside of the school's transportation zone means they have a lower probability of winning, all else equal. They are also required to provide their own transportation. The district also cares about equity. They show this by offering priority to students who are assigned to Title I *choice* schools. Title I schools are those with a high percentage of students eligible for free and reduced price lunch (FRPL). A Title I school becomes a Title I *choice* school if they fail to meet adequate yearly progress in the same subject for two consecutive years. No Child Left Behind (NCLB) requires that the district allow students assigned to Title I choice schools the opportunity to attend a non-Title I choice school, but it does not require the district to allow students to choose the school they are offered. In fact, they could be offered a school that they did not apply for in the lottery.

Assigning students to priority groups alters the probabilities of winning, and means that assignment is not unconditionally random. I use lottery (program of application by year by priority group) fixed effects to exploit the fact that winners should be randomly chosen within these groups. In addition to priority groups, all applicants are ordered based on randomly assigned numbers. When a choice is oversubscribed, the combination of priority groups and randomly assigned numbers determine who wins the lottery. The next subsection discusses the priority groups, and gives more detail on how lottery winners are determined during and after the first round.

2.1.1 Priority Groups

Seats are allocated based on priority group and lottery number. The top priority for applicants to full magnet schools in CMS is given to students who live within one-third mile of the school, but only twenty percent of seats can be assigned through that priority.[7] For example, if there are ten seats available to a specific full magnet school and more than two applicants live within one-third mile of the school, then the students with the first two numbers win under that priority. Then they move to the second priority group, students with Title I choice neighborhood schools.[8] Only ten percent of available seats can be assigned through this priority. Continuing with the example, the student with the first number who meets the second priority is assigned a seat, but the rest of the students assigned to Title I choice schools remain unassigned. Finally, they move to the third priority, all students who live in transportation zones served by the magnet school.[9] There is no limit on the number of seats assigned through this priority, so in this example, students with the next seven numbers who live in the transportation zone are admitted. The last priority is for students from transportation zones not served by the magnet school.[10] In this example, if more than two students meet priority one, then that priority group is oversubscribed. The identification strategy relies on comparing students who met a specific priority and won with students who met that priority and did not win. Similarly, if more than one student meets the second priority, then that lottery is oversubscribed as well. Finally, if more than seven students meet priority three, then that lottery is oversubscribed. In such a case, students from all three of those priority groups contribute to the estimates. In contrast, consider what happens to the students in the last priority group, those from outside of the transportation zone. Since no students in the last priority group won a seat, those students do not directly

[7]Students with this priority are still subject to the lottery if demand from this priority exceeds 20% of seats available. The area can be extended beyond 1/3 of a mile by the superintendent if the number of students enrolled meeting this criteria for a specific grade is less than 15.

[8]The first seats awarded are for students who qualify for FRPL and those below their grade level in reading. For kindergarten students, below grade level in reading is defined as having a personalized education plan.

[9]This priority first limits the number of seats from any particular neighborhood school assignment zone to be proportional to the potential number of applicants to the school. Then priority opens up to all students in the transportation zone. This restriction does not seem to be practically important.

[10]Applicants from outside of the transportation zone must provide their own transportation.

contribute to the estimates.

After going through all first choices, second choices are considered. If a student's second choice is already full from the first round of assignments, then they remain unassigned in the second round. Then third choices are considered. All students are assigned to a default neighborhood school based on pre-determined geographic zones if not otherwise assigned in the lottery. Since the lottery considers student choices in order, students are most likely to win a choice by picking it first, and more seats are awarded in the first round than in the second or third. In the following analysis I restrict to students who made a dual language school their first choice. The treatment assignment variable is a dummy variable for winning their first choice, which should be random within lottery.

2.1.2 Creating Lottery Fixed Effects

Although lottery fixed effects are not explicitly given in the data, I use available information to construct fixed effects. The data contain up to three choices for every student in order of preference, as well as sibling placement, Title I choice placement, FRPL status, and transportation zone.[11] I start with the sample of all applicants without a guaranteed seat and proceed in the following way to generate lottery fixed effects.

1. Proxy Title I choice school using whether or not any student from their neighborhood school was placed under the Title I choice option that year.
2. Generate priority groups using FRPL, transportation zone, and Title I choice proxy.
3. Lottery fixed effects are priority-year-program of application combinations.

Since the lottery fixed effects are generated, they are a proxy to the true lottery fixed effects. The assignment, conditional on lottery, provides the exogenous variation used to estimate causal effects.

[11]CMS stopped reporting FRPL after 2010. For 2011 I proxy for FRPL at the time of application using FRPL from the NCERDC data.

3 Data

There are three dual language schools in CMS. All three are full magnet schools, meaning that admission requires a lottery application and every student in the school participates in the dual language program. Collinswood Language Academy and Oaklawn Language Academy offer two-way English-Spanish classrooms, and Waddell Language Academy (formerly Smith Language Academy) offers full immersion strands in Mandarin, French, German and Japanese. Collinswood started in 1997 and now houses grades K-8. In kindergarten, 90% of instructional time is in Spanish [Thomas and Collier, 2009]. In grades one through five, half of the content is taught in each language. Oaklawn is a newer program, started in 2004, but follows a similar model to that of Collinswood. The curriculum is taught 90% in Spanish in kindergarten, 75% in first grade, and 50% in grades two through five.[12] Spanish is by far the most common non-English language among students in CMS, and the two-way programs are targeted toward native speakers of both languages. The German and French one-way immersion classes offered at Waddell have complete foreign language instruction in grades K-2, whereas the Mandarin and Japanese classes teach one hour in English per day in grades K-2 [Thomas and Collier, 2009]. All four programs at Waddell target 90% of instructional time in the non-English language in grades 3 – 5. The one-way programs primarily target English speaking students, but they admit ELLs who speak the partner language or another language altogether.

CMS and the North Carolina Education Research Data Center (NCERDC) provided the data for this study. CMS provided eight years (2006 – 2013) of lottery results with assignment into the three dual language schools. NCERDC merged the lottery data from CMS with statewide data. The following analysis will focus on end-of-grade exam scores in math and reading, which begin in third grade. Linking the lottery data with statewide data provides information on end-of-grade exam scores, and allows for the tracking of students who leave the district but stay in the North Carolina public school system. Since lottery results could impact school attendance decisions, this helps to mitigate attrition issues [Steele et al., 2015,

[12] Both schools use team teaching, divided by language of instruction.

Rouse, 1998].

My analysis sample includes students entering kindergarten from the 2006-2007 through 2010-2011 school years who submitted an application for a dual language school in the CMS school choice lottery, and were linked from the CMS data to the NCERDC data.[13] I start with the 2006-2007 school year because of changes implemented that year to the lottery system, including how the priority groups were determined. End-of-grade exams start in third grade, so the last year of entry used when estimating effects on exam scores is the 2010-2011 school year.[14]

Since estimation relies on applicants with non-guaranteed seats in oversubscribed lotteries, there are a couple of things worth noting. From the first row of Table 1, between 20 and 30 percent of the seats in each school were awarded to students with sibling guarantee. Those students are dropped from the estimation sample. The second row of Table 1 shows the percentage of applicants to each school that won their first choice. Only 56 percent of applicants who listed Collinswood as their first choice won their first choice, and 78 percent of first choice applicants to Waddell won their first choice in the lottery.[15] The CMS assignment mechanism only considers first choices in the first round of seat allocation, so if a school fills up in the first round, then second and third choice applicants to that school will not win a seat. Table 2 shows application numbers for students who chose one of the dual language schools as their second or third choice, but not as their first choice. From column 1, 49 percent of students who made a dual language school their second choice, won their first choice to a non-dual language school. About 14 percent of students who chose a dual language school with their second choice and not their first, won that choice, but only 10 percent attended a dual language school.

The CMS data also contain the neighborhood school that each student is assigned

[13]NCERDC was able to link between 93% and 97% of all observations from the CMS data in each year. Among observations of rising kindergarteners in the CMS data who chose a dual language school first in the lottery over the sample period, 93.5% were matched with the NCERDC data.

[14]The latest exam scores are from the 2013 - 2014 school year. This will be updated as NCERDC releases new exam scores each year.

[15]These percentages include students with guaranteed seats, so they are overestimates of the percentage of winners among those with non-guaranteed seats. About 43 percent of non-guaranteed applicants to Collinswood won and 69 percent of non-guaranteed applicants to Waddell won.

to, which helps to describe the outside options that students are foregoing to enter a dual language school. Characteristics of the neighborhood schools of the applicants are informative for thinking about the counterfactual. Language of instruction is not the only thing that changes for the student when they opt out of their neighborhood school and into a dual language school. Specifically, there could be changes in peer quality and composition of the student body. Mean characteristics of the schools that applicants are opting out of are displayed in Table 1. Applicants to Oaklawn come from schools that have a relatively high proportion of minorities (12 percent white), 76 percent of students on free and reduced price lunch, and score 0.3 standard deviations below the state average on end-of-grade math and reading exams. They come from neighborhood schools that score worse than the average for all applicants. On the other hand, applicants to Waddell and Collinswood come from neighborhood schools with a smaller percentage of FRPL students (57 percent and 65 percent, respectively), but still score below the state average on end-of-grade math and reading exams.

The three dual language schools are generally higher performing than the other schools in their respective neighborhoods. Table 3 shows that over 75 percent of students at Oaklawn are at grade level in reading, but only 50 percent of the students at schools in the area near Oaklawn are at grade level in reading.[16] Self selection and peer effects could play a significant role in the high performance of DL schools, but there are other features that might hurt their performance relative to neighborhood schools. Specifically, DL schools experience higher teacher turnover and begin with larger classes in kindergarten. Dual language classrooms need teachers who are fluent in the language of instruction, so the schools in CMS often recruit teachers from abroad. The teachers are permitted to work in the U.S. for a limited amount of time, leading to higher turnover. This is particularly true in Collinswood and Oaklawn. Table 3 shows that over 50 percent of the teachers in each of those schools has zero to 3 years of experience, compared to about 30 percent in the neighboring schools and other magnets. Not all teachers and staff members in dual language schools come from abroad,

[16]I refer to schools in their area as the neighborhood school zone that the dual language school is in as well as all of the school zones contiguous to that zone. Since all of the dual language schools are full magnet schools, no students are automatically assigned to them. Instead each student has a neighborhood school assignment that they attend unless they opt out through the lottery, change address, or enroll in a charter or private school.

nor are they necessarily fluent in a second language. Since they often implement team teaching, in most grades there is at least one English speaking teacher. Table 3 shows that the dual language schools do have highly experienced teachers, although they have a smaller proportion than the neighboring schools and other magnets. From column seven, 37 percent of teachers at other magnets have 11 or more years of experience, but that number is only 25 percent at Collinswood (column 4) and 27 percent at Waddell (column 2). Since students can not enroll in a dual language school after kindergarten (or first grade) without meeting a minimum language requirement, the schools start with larger class sizes, anticipating some attrition throughout elementary school. The average kindergarten class has 21.4 students at Collinswood and 22.3 at Waddell, as seen in columns 2 and 4 of Table 3. That is 3 more students than the other schools in their respective areas. From column 7 of Table 3, other magnet schools have 18.7 students in a kindergarten class on average. Although there are several differences between dual language schools and the typical neighborhood school, greater teacher turnover and larger early elementary school class size are two characterstics of DL schools that could lead to lower achievement.

Figures 1-3 provide descriptive comparisons of average standardized math and reading scores by LEP and DL status. Figures 1-2 compare average standardized math and reading scores for dual language and non-dual language students.[17] These are descriptive comparisions of DL and non-DL students similar to what has generally been examined in prior studies. They represent a good starting point, but ignore useful information on lottery fixed effects and sibling placement. Figure 1 graphs the comparison for non-LEP students. Non-LEP, dual language students score well above the state average in reading in every grade, and there is a divergence between dual language students and the rest of the district from grades three through eight. In seventh and eighth grade the dual language students score more than 0.3 standard deviations above the state average in reading. Non-LEP, dual language students

[17] Figures 1-2 graph average standardized (by year and grade across the state) residualized scores. They are residuals from linear regressions of standardized exam scores on grade dummies, year dummies, sex, FRPL status, and exceptionality. For the purposes of Figures 1-2, dual language students are all students who attended a dual language school in any grade, 3 - 8. Figure 1 uses all students who were not identified as limited English proficient in any grade, 3 - 8, and Figure 2 uses all students who were identified as limited English proficient in at least one of those grades.

also score well above the state average in math, but the gap between DL students and the rest of the district displays a slight downward trend with grade. The dual language, non-LEP students score about 0.3 standard deviations above the mean in grades 4 and 5, but about 0.2 standard deviations above the mean in eighth grade. The counterfactuals in Figure 1, lines for the non-DL students, include all non-DL students in the district, most of whom had no interest in attending a dual language school. Since there are likely systematic differences between DL applicants and non-applicants, estimates generated from this sort of analysis should not be considered causal. Figure 2 displays the analogous comparison for students who were identified as LEP in at least one grade, three through eight. Non-dual language, LEP students score below the state average in math and reading in every grade. On the other hand, LEP students who attend dual language schools score about 0.2 standard deviations above the state average in math in third grade, and more than 0.2 above the average in every grade after that. They also score at the state average in reading in third grade, and above it in every grade after third. Once again, Figure 2 provides evidence that LEP, DL students score above their non-DL peers in math and reading on average, but the differences should not be interpreted as causal effects.

While these graphs provide useful descriptions of the gaps in test scores for DL students, they do not provide causal evidence on the differences in scores. For causal evidence, I turn to the randomization created by the oversubscribed lotteries. Only lottery applicants with non-guaranteed seats are used to estimate causal effects because the estimation strategy relies on comparing winners and losers of the same lottery. Tables 4 and 5 describe the lottery winners and losers in the estimation samples, which is restricted to applicants who remained in the sample long enough to have test scores available. Average math and reading scores on their first exam are displayed for lottery winners in column 1 and those who lost the lottery in column 2. The differences in these scores give the raw test score gaps after restricting to the estimation sample. From Table 4, among the non-ESL/LEP students, lottery winners scored about 0.24 standard deviations (0.52 - 0.28) higher than lottery losers on their first end-of-grade math exam.[18] Lottery winners in that sample scored about 0.2 standard deviations

[18]The scores included are from the first exam score available for each student, which is typically the third

higher on their first end-of-grade reading exam. These differences still do not warrant a causal interpretation, because they ignore lottery fixed effects. The analogous differences between winners and losers in the ESL/LEP sample are shown in Table 5. Lottery winners scored 0.08 standard deviations above the state mean on their first math exam, and lottery losers scored about 0.2 standard deviations below the state mean. That is a difference of more than 0.28 standard deviations in favor of lottery winners on their first end-of-grade math exam. Similarly, lottery winners in the ESL/LEP sample scored about 0.24 standard deviations higher than lottery losers on their first end-of-grade reading exam, although both groups scored below the state average.

If students perfectly complied with the lottery assignment, then assignment would be synonomous with attendance and the causal effect could be estimated using OLS regressions of achievement on assignment/attendance. However, students do not perfectly comply with initial assignment from the lottery. From Table 4, 89.9 percent of first choice lottery winners and 37.5 percent of first choice lottery losers from the non-ESL/LEP subsample attend a dual language school, meaning that there is non-compliance among winners and losers. The first row of Table 5 gives the analogous figures for the ESL/LEP subsample. Ninety-three percent of lottery winners from the ESL/LEP sample attend a dual language school and 29 percent of lottery losers attend a dual language school.

Winners are not bound to attend the school they won the lottery for, and lottery losers can end up in a dual language school despite losing the initial lottery. There are several ways this can happen. First, they could win a seat to a different dual language program with their second or third choice in the lottery. This is somewhat unlikely since they are typically filled up by students making them their first choice, but it does happen. From Table 4, about 31 (21) percent of lottery losers in the non-ESL/LEP sample chose a dual language school with their second (third) choice. More than 11 percent of the lottery losers in that sample won a seat in a dual language program with their second or third choice. Table 5 shows that lottery losers from the ESL/LEP sample were less likely to choose a dual language school for their second (third) choice, as only 23 (12) percent did, and only 2 percent of them won a seat in

grade score.

a dual language school. Second, students who do not win their first choice are placed on a waiting list for that school, which is accessed if seats become available. If a lottery winner chooses not to take the seat offered to them, the seat is offered to the next student on the waiting list. This is likely a major source of non-compliance from lottery losers. From the non-ESL/LEP sample in Table 4, 10.1 percent of winners do not end up attending a dual language school. From Table 5, 7.1 percent of lottery winners in the ESL/LEP sample do not attend a dual language school. Even if a winning student enrolls in the dual language school and attends that school, but eventually exits, that seat can be offered to another student. The waiting list can be accessed all the way through the first academic quarter of the school year. Lastly, students can reapply in the school choice lottery for the subsequent year and win a seat.[19] As discussed in further detail in the next section, non-compliance does not invalidate the empirical strategy used in this paper.

For causal inference, assignment must be a significant predictor of attendance and must be exogenous conditional on lottery fixed effects. I first examine whether assignment is a significant predictor of attending a DL program. I test for differences in DL attendance between winners and losers, conditional on lottery fixed effects, by regressing the dummy variable for attending a DL school on a dummy for winning and lottery fixed effects. The first row of Tables 4 and 5 displays the estimated coefficients on the dummy for winning the lottery, which indicate whether winning the lottery actually predicts DL attendance. Rejecting the null hypothesis of no effect indicates that winning is correlated with attendance. From column 3 of Table 4, conditional on lottery fixed effects, lottery winners in the non-ESL/LEP sample are 55 percent more likely to attend a dual language school than lottery losers. Column 3 of Table 5 shows that in the ESL/LEP sample winners are 64 percent more likely to attend a DL school. Both estimates are statistically significant, suggesting that winning the lottery is a good predictor for attending a DL school, which is necessary to implement the identification strategy used in this paper.

The remaining tests, found in column three of Tables 4 and 5, give some indication

[19]There is also a second lottery mainly for students who enrolled in CMS after the deadline for the first lottery, but second lottery applicants are placed at the end of the waitlist for oversubscribed programs.

whether the lottery results are truly random. Assignment is random within lottery groups, so the generated lottery fixed effects are included in each test. Since lottery groups depend on geographic location and free and reduced price lunch status, we shouldn't expect fixed characteristics of applicants to be unrelated to winning the lottery unconditionally. A rejection of the null hypothesis suggests that winning the lottery might be related to that characteristic in some non-random way and generally gives cause for concern about the identification strategy proposed below. The only rejection in column 3 is on the coefficient in the regression of a dummy variable for black on winning the lottery in the non-ESL/LEP sample, found in Table 4. There are at least two reasons why the test might reject even if assignment is random. The first could be from non-random attirition from the sample. Since I am estimating effects on math and reading scores, students who do not remain in the sample long enough to observe test outcomes must be dropped for estimation. Even though assignment is random at the time of application, it is not necessarily random when restricting to the applicants that remain in the sample through third grade. Staying in the district could be related to winning the lottery and the resulting attrition would lead to selection bias [Steele et al., 2015, Rouse, 1998].[20] Tables 4 and 5 only include students in the estimation sample. Table A1 in the appendix shows the summary statistics and balance in the non-ESL/LEP sample of applicants. The balance test for black in this sample is still marginally significant, but much smaller in absolute value, suggesting that non-random attrition could be important. However, another possible explanation is that assigment is actually random, and the rejection of the null hypothesis is an artifact of measurement error in the constructed proxies used for lottery fixed effects. Since priority groups depend on free and reduced price lunch status, I control for this flexibly by including free and reduced lunch by cohort dummy variables in addition to the lottery fixed effects. The results for tests that include additional controls for FRPL are shown in column four of Tables A1 and A2 for the full sample of applicants. With the additional FRPL controls, all tests in the non-ESL/LEP subsample fail to reject the null. To further alleviate concerns of endogeneity and non-random attrition, I

[20] Attrition is likely higher because of the lag between application and testing and the focus on students entering kindergarten. I have at least one set of exam scores for about eighty-five percent of applicants (see Tables A1 and A2 in the appendix).

include a number of robustness checks including using weights based on the probability of remaining in the sample.

4 Empirical Strategy

The effect of attending a dual language school on achievement can be estimated directly using OLS to estimate equation 1.

$$Y_{i,j,g} = \gamma \cdot 1[DualLanguage]_{i,j} + \beta \cdot X_{i,j,g} + \Omega_j + \varepsilon_{i,j,g} \quad (1)$$

Where $Y_{i,j,g}$ represents an end-of-grade math or reading exam score of student i in grade g who applied to lottery j. The variable of interest, $1[DualLanguage]_{i,j}$, is a dummy variable that is equal to one if the student attended a dual language school.[21] Lottery fixed effects, Ω_j, are included because winning the lottery is not unconditionally random, but students are drawn randomly within lottery. Covariates, $X_{i,j,g}$, are also included. Grades are pooled for estimation, so grade of exam dummy variables are included in $X_{i,j,g}$. One concern with this approach is that, although the assignment is random conditional on lottery fixed effects, compliance with initial assignment may not be random, leading to a biased and inconsistent estimator for the average treatment effect. Compliance might be non-random for a couple of reasons. In particular, over 30 percent of lottery losers end up attending a dual language school. Students who attend a dual language school, despite losing the lottery for their first choice, might be systematically different from the students who lost and did not end up attending a dual language school. For example, students who chose a dual language program with their second and/or third choice are more likely to attend a dual language school relative to those who did not specify a dual language school with their second and/or third choice.

[21] Using enrollment in the year of the exam is one way to measure participation. That leaves a lot of time between application and when enrollment is measured. One might worry that this could bias estimates since students have time to apply to other schools or simply withdraw from the dual language program, both of which are likely non-random. For this reason, I prefer using enrollment in kindergarten as the participation measure. In one year of the data (2007), the school of attendance in kindergarten is missing for a non-trivial portion of applicants, many of whom show up in a dual language school in first grade. For this reason, I actually measure attendance as showing up in a dual language school in either kindergarten or first grade.

Non-compliance could represent strength of preferences for dual language schooling or for their neighborhood school, or the ability of parents to maneuver their way into their first choice school. Since OLS estimators are biased and inconsistent if attending a dual language school is non-random, I focus on estimating the intention-to-treat and local average treatment effects which are consistent when assignment is random conditional on lottery fixed effects.

I follow a standard approach for estimating treatment effects using applicants for over-subscribed lotteries [Deming et al., 2014, Rouse, 1998]. The intention-to-treat effect is estimated by regressing end-of-grade math and reading scores on a dummy for winning the lottery and a set of covariates in the sample of lottery applicants, as shown in equation 2.

$$Y_{i,j,g} = \gamma^{ITT} \cdot 1[LotteryWinner]_{i,j} + \beta^{ITT} \cdot X_{i,j,g} + \Omega_j^{ITT} + \varepsilon_{i,j,g}^{ITT} \tag{2}$$

Where $1[LotteryWinner]_{i,j}$ indicates whether student i was a winner of lottery j. In an alternate specification, neighborhood school fixed effects are included.[22] The estimated coefficient of interest, $\hat{\gamma}^{ITT}$, is an estimate of the intention-to-treat [Imbens and Angrist, 1994]. The difference between equations 1 and 2 is that equation 2 replaces the variable of interest, $1[DualLanguage]_{i,j}$, with the assignment variable, $1[LotteryWinner]_{i,j}$. The estimators from equations 1 and 2 are not estimating the same parameter, but $\hat{\gamma}^{ITT}$ is consistent under the assumption that assignment is random. Whereas, consistency of $\hat{\gamma}$ requires the less plausible assumption that attending a dual language school is random. Both the intention-to-treat and local average treatment effect estimators share this advantage over the OLS estimator from equation 1.

Equations 3 and 4 describe a two-stage estimation strategy using the dummy for winning the lottery as an instrument for attending a dual language school. Now $\hat{\gamma}^{LATE}$ is an estimate of the local average treatment effect, the effect for those who are induced to participate by winning the lottery [Imbens and Angrist, 1994]. In the main specification, the effects are actually estimated by pooling grades and interacting the treatment dummy with years

[22] Neighborhood school refers to the school that the student was assigned to at the time of the lottery. This is the school that the student would be assigned to attend in kindergarten unless the student either opts out during the lottery, enrolls in a charter or private school, or changes address.

of treatment (grade of exam plus one). Dummy variables are included for grade of exam, leading to a per-year of participation intepretation.

$$1[DualLanguage]_{i,j} = \gamma^{DL} \cdot 1[LotteryWinner]_{i,j} + \beta^{DL} \cdot X_{i,j,g} + \Omega_j^{DL} + \varepsilon_{i,j,g}^{DL} \qquad (3)$$

$$Y_{i,j,g} = \gamma^{LATE} \cdot \hat{1}[DualLanguage]_{i,j} + \beta^{LATE} \cdot X_{i,j,g} + \Omega_j^{LATE} + \varepsilon_{i,j,g}^{LATE} \qquad (4)$$

I perform specification checks to alleviate any concerns from exogeneity of the treatment or non-random attrition. Specifically, I estimate effects using weights based on the estimated probability of remaining in the sample.[23] Weighting the regressions adjusts for non-random attrition related to observable characteristics. I also include a second set of all estimates that include neighborhood school fixed effects. Neighborhood school is defined as the school that the student would have been assigned to if they did not win any seat in the lottery, change address, or enroll in a charter or private school. Having the same neighborhood school means that the students live in the same geographic area and have the same outside schooling option.

In addition to estimating the effect of attendance on acheivement, I estimate the effect of attending a dual language school on limited English proficiency status. I interact the treatment and attendance variables with each grade (three through six), and estimate the effect on having limited English proficient status in each grade on the ESL/LEP sample. Prior research suggests that dual language participants re-classify at a slower rate in early grades, but eventually surpass their non-dual-language-schooled peers [Umansky and Reardon, 2014]. This is a good point of reference, although we should not necessarily expect these results to be the same. These are two different contexts, and I focus on estimating effects in a select subsample, unlike the district wide analysis by Umansky and Reardon [2014].

[23] Remaining in the sample means that the student has valid end-of-grade exam scores for at least one grade. Weights are based on estimated probabilities from logit regressions of an indicator for staying in the sample on race, gender, FRPL, and a dummy for winning the lottery.

5 Results

I begin by providing first stage ($\hat{\gamma}^{DL}$ from equation 3) and treatment effect ($\hat{\gamma}^{ITT}$ from equation 2 and $\hat{\gamma}^{LATE}$ from equation 4) estimates from the main specification in Table 6. Panel A of Table 6 shows estimates for the non-ESL/LEP sample of applicants, students who were never identified as eligible for English second language services or as limited English proficient. The estimated effect on math scores in column 5 suggests that among compliers, attending a dual language school led to an increase in math scores of 0.086 standard deviations. This can be interpreted as a per-year gain in achievement. Column 7 shows that the effect on reading scores for this sample is about 0.06 standard deviations per year. Both estimates are statistically significant. These estimates are promising for the growing practice of dual language education for native English speakers. At least in CMS, the dual language schools have been successful in delivering instruction in a second language, and increasing math and reading exam scores for English proficient students. Although these estimates do not separate out the mechanisms through which achievement gains are operating, they show that it is possible to successfully promote bilingualism and increase academic achievement.

Panel B of Table 6 shows estimated treatment effects for the ESL/LEP sample. The estimated effect of attending a dual language school on math scores in column 5 is 0.063. From column 7, the estimated effect on reading scores in this sample is 0.069. Both estimates are statistically significant. While these estimates are large, they are in line with the fact that treatment is muti-year and begins at a young age. The estimates suggest that dual language education can be an effective teaching method for ELLs and help to reduce achievement gaps in math and reading. Consider the achievement gaps between LEP and non-LEP students, which are displayed in Figure 3. The district average math and reading scores for LEP students are below the state averages in every grade, and below the district non-LEP averages in every grade. The largest gap in math scores is about 0.2 standard deviations, so the estimate of 0.063 standard deviations per year is large enough to more than close that gap by third grade. The largest disparity in reading scores is a little more than 0.4 standard deviations. The estimated effect on reading scores of 0.069 standard deviations is enough to close the gap in test scores by the end of elementary school.

As shown in Table 7, estimates are generally robust to the inclusion of neighborhood school fixed effects. The estimate on reading scores for the non-ESL/LEP sample, shown in column 7 of Table 7, decreases to 0.053 but is still statistically significant at the 10 percent level. The estimated impact on math scores in that sample, reported in column 5, increases to 0.089, and is statistically significant at the 10 percent level. Estimated effects in the ESL/LEP sample are reported in Panel B of Table 7. The estimated effect on math scores is reported in column 5. It is higher than in the main specification, now 0.078, and statistically significant at the five percent level. The estimated effect on reading scores, reported in column 7, decreases from 0.069 to 0.064, but is still statistically significant at the five percent level. After adding neighborhood school fixed effects, all four of the estimates remain positive, in the same range as the initial estimates, and statistically significant.

Since non-random attrition would lead to inconsistent estimators and test scores are missing for a non-trivial portion of applicants, I include estimates that are weighted by the inverse of the estimated probability of having test scores in the data. I estimate the probability of remaining in the sample long enough to have valid test scores using logit regressions on dummy variables for race/ethnicity, gender, FRPL, and winning the lottery, then use the inverse of the estimated probabilities as weights in the estimation. Weighted estimates are reported in Table 8.[24] Panel A shows the inverse probability weighted estimates for the non-ESL/LEP sample. From colmun 5 of Panel A, the estimated treatment effect for math scores is now 0.086, the same as the initial estimate, and significant at the five percent level. The weighted estimate for reading in that sample, from column 7, is 0.057, and statistically significant. The weighted estimates on math and reading scores in the non-ESL/LEP sample are the same as the estimates from the initial specification, suggesting that non-random attrition is not likely to be a significant factor. Estimates for the ESL/LEP sample are shown in Panel B of Table 8. The estimated average treatment effects on math and reading scores are only 0.001 smaller than the initial estimates, and both are still statistically significant at

[24]Weighted estimates that include neighborhood school fixed effects are in Table A3 of the appendix. The estimates are not sensitive to this specification.

the five percent level. Weighting based on the inverse of the estimated probability of attrition has almost no impact on the point estimates, which suggests that non-random attirition is probably not inflating the estimates much, if at all. Weighting leads to smaller estimated effects on math and reading scores in the ESL/LEP sample, but the difference is so small that it is not practically significant.

I report estimates of heterogeneous treatment effects in Table 9; these allow effects to differ by gender (columns 1-2), program type (columns 3-4), or race/ethnicity (columns 5-7).[25] Heterogeneous treatment effects are estimated by interacting dummy variables indicating mutually exclusive sets of students with the attendance variable, and using the same dummy variables interacted with the assignment variable as instruments. Estimates for the non-ESL/LEP sample are reported in Panel A. Columns 1 and 2 of Panel A show that effects on math scores for females, 0.114, are stronger than for males, 0.051. Similary, the estimated effect on reading for females is 0.083 and statistically significant at the five percent level, and the effect for males is a statistically insignificant 0.024. On the other hand, columns 1 and 2 in Panel B suggest that the effect is stronger for males in the ESL/LEP subsample, but the difference in estimates is much smaller than the differences in the non-ESL/LEP subsample. The difference in heterogeneous effects by gender between samples is somewhat striking, and may reflect the difference in treatments. A big part of the treatment for students in the ESL/LEP subsample is likely that they receive some instruction in their home language as opposed to English immersion coupled with ESL services. On the other hand, treatment in the non-ESL/LEP sample is typically receiving instruction in a second language as opposed to English immersion. The differences in heterogeneity could result from differing treatments and potentially different mechanisms facilitating the effects.

Effects for one-way and two-way programs are reported in columns 3 and 4. The difference comes down to which school the student applied to since Waddell contains all of the one-way programs and the other two schools, Collinswood and Oaklawn, house two-way programs only. The size of the estimated effects are similar by program type for the non-ESL/LEP sample, but the estimates on effects for one-way programs have much larger

[25]Estimates on heterogeneous effects that include neighborhood school fixed effects are shown in Table A4.

standard errors. For example, the estimated effect on math scores for one-way programs in that sample is 0.068, but the standard error is 0.129. The estimate on math for two-way programs is 0.087 with a standard error of 0.040. There is also a statistically significant estimated effect for two-way applicants of 0.058 on reading scores, but the estimated effect for one-way applicants is smaller and statistically insignificant. Panel B in Table 9 reports estimated treatment effects for students in the ESL/LEP sample for one-way and two-way programs. Similar to the non-ESL/LEP sample, estimates for one-way programs are very noisy. The estimate on math scores for one-way programs is 0.111 in this sample, but the standard error is 0.154. The estimated effects for two-way applicants in the ESL/LEP sample are 0.058 and 0.071 on math and reading scores, respectivley. Both estimates are statistically significant.

Finally, I estimate heterogeneous effects by race/ethnicity in the non-ESL/LEP sample in columns 5, 6, and 7 of Panel A in Table 9. The estimated impact on math scores is largest in the white subsample, but estimates for the black and Hispanic subsamples are also positive and the estimate for the Hispanic subsample is statistically significant. The estimated treatment effect on math scores for the white subsample is 0.205, which is large relative to most other estimated effects, and is significant at the five percent level. The estimated effect on the black subsample is 0.063 but it is statistically insignificant. The estimated treatment effect on math scores in the Hispanic subsample is 0.076 and significant at the five percent level. Estimated effects on reading scores are relatively similar across the white, black, and Hispanic subsamples. From Panel A of Table 9, the estimated effect on reading scores in the black subsample, 0.055, is only half the size of that estimate in the white subsample, 0.104, but both estimates are statistically insignificant. The only significant effect on reading scores is on the Hispanic subsample, 0.085, and it is significant at the one percent level.

I do not estimate effects for each race in the ESL/LEP sample, because 85% of the students in that sample are Hispanic. Any estimate for other races would be unreliable. However, restricting to the Hispanic subsample using dummy interactions shows that the main finding is robust in this subsample. Column 7 in Panel B of Table A3 shows the estimated effects on math and reading for the Hispanic students in the ESL/LEP sample. The estimated effect on math scores is 0.068, and is significant at the five percent level. The esimtated effect on reading is 0.065 and significant at the five percent level.

Table 10 shows estimates by grade.[26] These are estimated by interacting the DL attendance variable and/or the indicator for winning the lottery with each exam grade. The estimated impact on math exam scores for the non-ESL/LEP sample are shown in column 5 of Panel A. The estimates increase with grade. The estimated effect for math scores on the third grade interaction is 0.290, and significant at the ten percent level. The estimated effect on math scores for sixth grade is 0.748 and significant at the five percent level. This estimate is only identified from two of the cohorts, leading to a relatively large standard error of 0.353. The estimates for the effect on reading scores in this sample are shown in column 6 of Panel A in Table 10, and they also exhibit an increasing pattern with grade. The estimates on the third and fourth grade interactions are 0.168 and 0.127, respectively. Neither of them are statistically significant. The largest estimate is on the fifth grade term, 0.449, and it is significant at the five percent level.

Estimated effects are also reported by grade for the ESL/LEP sample in Table 10. The estimated effects are stronger in the ESL/LEP sample, but the estimated effects for math scores do not exhibit quite as strong of an increasing pattern with grade. The effect on math scores on the third grade interaction from column 5 in Panel B of Table 10 is 0.362 and significant at the five percent level. The estimated coefficient on the sixth grade interaction is 0.425 and is also significant at the five percent level. The largest of all of the estimated effects on math scores for the ESL/LEP sample is on the fourth grade interaction. That estimate is 0.633 and significant at the five percent level. The largest estimated effect on reading in the ESL/LEP sample is on the sixth grade interaction, an effect of 0.537 and significant at the one percent level.

In addition to estimating treatment effects on math and reading scores, I estimate the effect of attending a dual language school on LEP classification among the sample of students ever eligible for ESL services or considered LEP. I estimate the effects by regressing a dummy

[26]Estimates including neighborhood school fixed effects are shown Table A5 of the appendix.

variable for being considered LEP in a given year on DL attendance by grade interactions. I instrument for attendance by grade interactions using a dummy for winning the lottery interacted with each grade. OLS estimates by grade are shown in columns 1 and 2 of Table 11. Column 1 shows estimates without neighborhood school fixed effects. Every estimate in column 1 is negative, meaning that students who attend DL schools are less likely to be considered limited English proficient in each grade. The largest in absolute value is the -0.210 estimate on the sixth grade interaction and it is significant at the one percent level. The analogous treatment effects are shown in column 3. They are all negative, but only the estimate on the sixth grade interaction, -0.168, is statistically significant. These estimates are in line with the higher English reading scores, but seems to counter some results in the prior literature [Umansky and Reardon, 2014] yet agree with others [Steele et al., 2015]. These results are not necessarily comparable with prior literature on re-classification since estimates are specific to a set of students who applied for dual language schools in CMS. Furthermore, all of the estimates on LEP classification are noisy and most of them are not significantly different from zero. In general though, they suggest that movements forcing English immersion on ESL/LEP students might be misguided. In this setting, students attending DL schools not only score higher on math and reading exams, but they are also less likely to be considered LEP in grades 3-6.

6 Conclusion

Dual language magnet schools in Charlotte-Mecklenburg offer an alternative option for students to learn curriculum in a non-English language. I find that, conditional on some baseline characterstics, dual language students score higher than their peers on end-of-grade math and reading exams. One concern with this initial descriptive analysis and previous literature is that differences may be driven by self-selection, so I use random assignment from school choice lotteries to estimate causal effects of attending a dual language school on student achievement. In the main specification, I estimate local average treatment effects of more than 0.06 standard deviations per year on math and reading exam scores among students who were ever eligible for ESL services or considered LEP. The effects are robust to several alternative specifications, and large enough to close the LEP - non-LEP achievement gap in math and reading if applied to an average LEP student in CMS. I find further evidence that among students in this sample, those who attend a dual language school are less likely to be considered LEP in grades three through six, although the differences are generally statistically insignificant. The estimates on achievement and LEP classification suggest that dual language education has led to large benefits for students with limited English proficiency and appears an effective way to serve the population of ELLs in CMS.

Among English first language applicants, the estimated impact on math scores of about 0.09 standard deviations per year is robust to different specifications and represents a large increase in achievement. The effect on math scores in the non-ESL/LEP subsample is substantially stronger among females and white students. The estimated effect in reading for this sample is 0.057 standard deviations per year in the main specification. The size of the effect on reading scores is also robust across specifications. There is some evidence that effects on reading scores might be stronger among female students, but there is less evidence of heterogeneity by school type or race. For English first language students, it appears that the dual language schools in CMS provide a good opportunity for them to become bilingual and biliterate without sacrificing achievement in other areas. Not only are they not losing ground in math or reading, they are experiencing large gains in both math and reading achievement.

References

[Alloway, 2010] Alloway, T. P. (2010). *Improving working memory: Supporting students' learning*. Sage.

[Baddeley, 2003] Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature reviews neuroscience*, 4(10):829–839.

[Baddeley and Hitch, 1974] Baddeley, A. D. and Hitch, G. J. (1974). Working memory. *The psychology of learning and motivation*, 8:47–89.

[Cazabon et al., 1999] Cazabon, M., Lambert, W., and Hall, G. (1999). Two-way bilingual education: A report on the Amigos Program. *Washington, DC: Center for Applied Linguistics*.

[Cho, 2012] Cho, R. M. (2012). Are there peer effects associated with having english language learner (ell) classmates? evidence from the early childhood longitudinal study kindergarten cohort (ecls-k). *Economics of Education Review*, 31(5):629–643.

[Chrysochoou et al., 2011] Chrysochoou, E., Bablekou, Z., and Tsigilis, N. (2011). Working memory contributions to reading comprehension components in middle childhood children. *The American journal of psychology*, 124(3):275–289.

[Cobb et al., 2009] Cobb, B., Vega, D., and Kronauge, C. (2009). Effects of an elementary dual language immersion school program on junior high achievement. *Middle Grades Research Journal*, 1(1):27–48.

[Deming et al., 2014] Deming, B. D. J., Hastings, J. S., Kane, T. J., and Staiger, D. O. (2014). School Choice , School Quality. 104(3):991–1013.

[Dufva et al., 2001] Dufva, M., Niemi, P., and Voeten, M. J. (2001). The role of phonological memory, word recognition, and comprehension skills in reading development: From preschool to grade 2. *Reading and Writing*, 14(1-2):91–117.

[Geay et al., 2013] Geay, C., Mcnally, S., and Telhaj, S. (2013). Non-native speakers of English in the classroom: What are the effects on pupil performance? *Economic Journal*, 123(November 2010):281–307.

[Greene, 1998] Greene, J. (1998). A Meta-Analysis of the Effectiveness of Bilingual Education.

[Howard and Sugarman, 2001] Howard, E. and Sugarman, J. (2001). Two way immersion programs: Features and statistics. *Occassional Reports UC Berkeley*, (March):10–13.

[Imbens and Angrist, 1994] Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

[Maxwell, 2012] Maxwell, L. A. (2012). Dual classes see growth in popularity. *Education Week*, 31(26):16–17.

[Maxwell, 2014] Maxwell, L. A. (2014). Dual Language Programs Take Root in N.C. *Education Week*, 34(8):1.

[Rouse, 1998] Rouse, C. E. (1998). Private school vouchers and student achievement: An evaluation of the milwaukee parental choice program. *The Quarterly Journal of Economics*, 113(2):553–602.

[Slavin and Cheung, 2005] Slavin, R. E. and Cheung, a. (2005). A Synthesis of Research on Language of Reading Instruction for English Language Learners. *Review of Educational Research*, 75(2):247–284.

[Steele et al., 2015] Steele, J. L., Slater, R. O., Zamarro, G., Miller, T., Li, J., and Burkhauser, S. (2015). The Effect of Dual Language Immersion on Student Achievement: Evidence from Lottery Data. *unpublished preliminary draft.*

[Thomas et al., 2010] Thomas, W., Collier, V., and Collier, K. (2010). English learners in North Carolina, 2010. *Fairfax, VA: George Mason University. A Research Report Provided to the North Carolina Department of Public Instruction.*

[Thomas and Collier, 2009] Thomas, W. P. and Collier, V. (2009). English Learners in North Carolina, 2009 The North Carolina Context. *Fairfax VA: George Mason University. A Research Report Provided to the North Carolina Department of Public Instruction*.

[Umansky and Reardon, 2014] Umansky, I. M. and Reardon, S. F. (2014). Reclassification patterns among Latino English learner students in bilingual, dual immersion, and English immersion classrooms. *American Educational Research Journal*, 51(5):879–912.

[Valentino and Reardon, 2015] Valentino, R. A. and Reardon, S. F. (2015). Effectiveness of four instructional programs designed to serve english learners variation by ethnicity and initial english proficiency. *Educational Evaluation and Policy Analysis*.

[Watanabe, 2011] Watanabe, T. (2011). Dual Language Immersion Programs Growing in Popularity. *Los Angeles Times*.

[Willig, 1985] Willig, A. C. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of educational research*, 55(3):269–317.

Figure 1

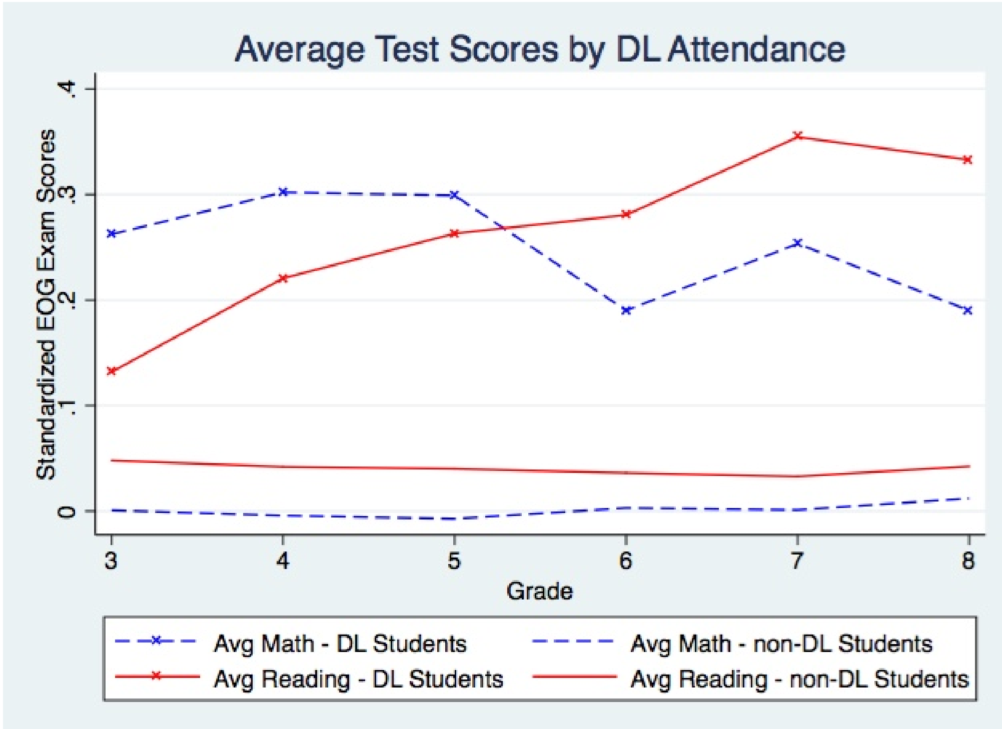

Figure 2

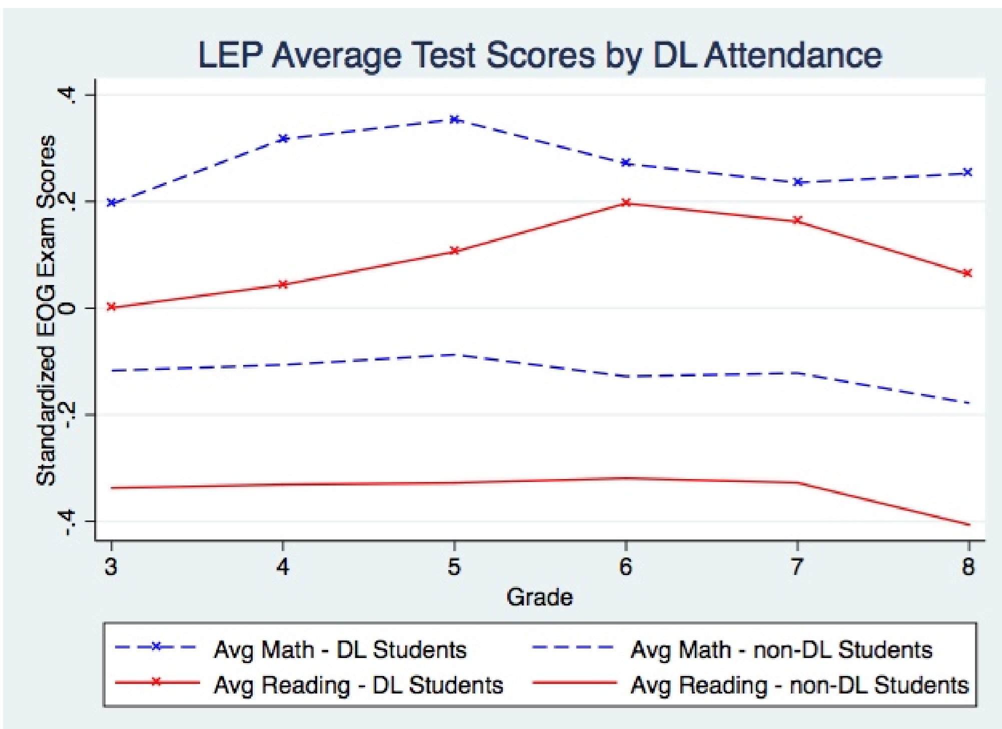

Figure 3

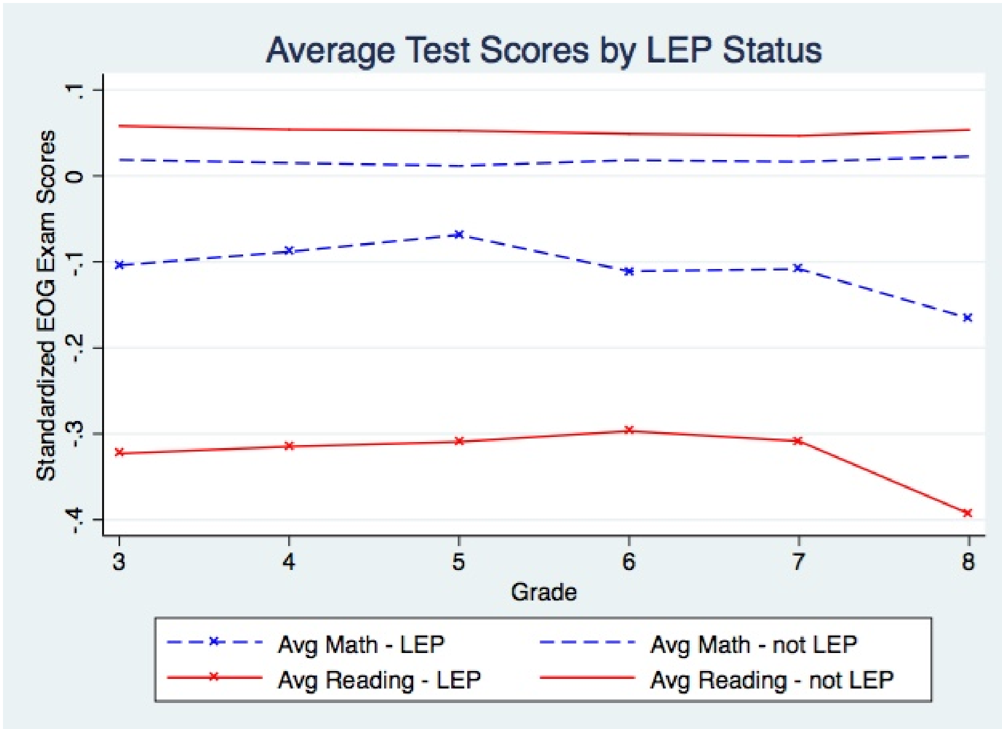

Figure 4

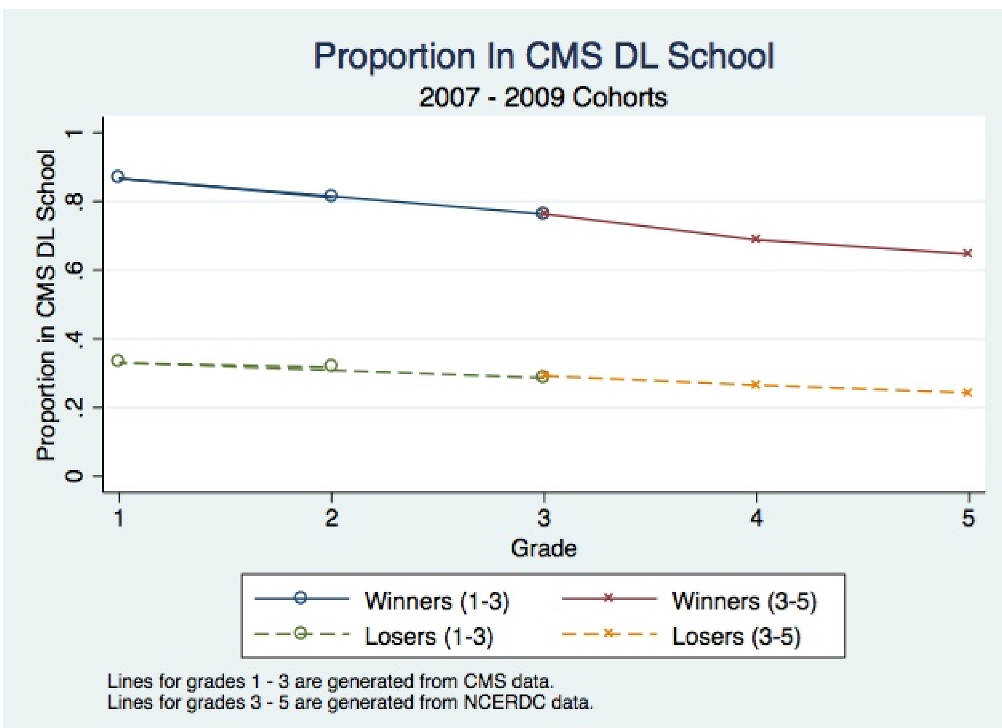

Table 1: Application Numbers and Neighborhood School Characteristics

	One-Way School	Two-Way Schools		
	Waddell	**Collinswood**	**Oaklawn**	**Other Applicants**
	[1]	[2]	[3]	[4]
Applicants				
Sibling Placement	0.296	0.232	0.218	0.230
Won First Choice	0.782	0.561	0.951	0.589
DL Applications	2.081	1.248	1.283	0.093
Neighborhood School				
White	0.315	0.250	0.119	0.197
Black	0.383	0.389	0.607	0.501
Hispanic	0.216	0.281	0.194	0.220
FRPL	0.574	0.652	0.760	0.688
LEP	0.180	0.232	0.175	0.194
EOG Exam Scores				
Math	-0.010	-0.080	-0.318	-0.186
Reading	-0.032	-0.136	-0.357	-0.224
N	1,147	1,112	533	13,071

*Notes: The first three rows display of the type and number of applications submitted for all those submitting applications in the CMS school choice lottery. The rest of the table shows mean characteristics of the neighborhood schools that applicants are assigned to weighted by the number of applicants from each school. Everything is based on first choice school.

Table 2: Second and Third Choices

	Second Choice DL [1]	**Third Choice DL** [2]
Attend DL School	0.105	0.087
Second Choice DL	1.000	0.219
Third Choice DL	0.203	1.000
Assignment		
Collinswood	0.000	0.000
Waddell	0.080	0.072
Oaklawn	0.077	0.075
Any DL Choice	0.157	0.147
Choice		
Collinswood	0.339	0.245
Smith	0.332	0.475
Oaklawn	0.329	0.279
Won		
First Choice	0.490	0.426
Second Choice	0.143	0.094
Third Choice	0.070	0.117
Any Choice	0.703	0.638
Observations	286	265

Table 3: Dual Language and Neighborhood School Characteristics

	Waddell		Collinswood		Oaklawn		
	Area	Waddell	Area	Collinswood	Area	Oaklawn	Other Magnets
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Teaching Experience							
0 - 3 Years	0.329	0.320	0.294	0.501	0.298	0.542	0.299
11+ Years	0.300	0.272	0.367	0.255	0.359	0.195	0.374
FRPL	0.806	0.334	0.759	0.573	0.900	0.688	0.568
AYP Targets	0.869	0.986	0.871	1.000	0.779	1.000	0.873
Pct at Grade Level							
Reading	0.612	0.822	0.612	0.880	0.496	0.758	0.718
Math	0.698	0.878	0.705	0.944	0.531	0.753	0.733
KG Class Size	18.057	21.400	19.049	22.333	18.200	19.333	18.736

*Note: Average characteristics at each dual language school, for the neighborhood schools with zones contiguous to each dual language school, and all other magnet schools.

Table 4: Summary and Balance - Estimation Sample - English

			Test Difference	
	Won	Lost	Lottery FE	FRPL
	[1]	[2]	[3]	[4]
Attend DL School	0.899	0.375	0.550***	0.543***
			(0.049)	(0.050)
Female	0.533	0.508	0.024	0.024
			(0.042)	(0.043)
Black	0.303	0.387	-0.112**	-0.096**
			(0.044)	(0.046)
White	0.373	0.262	0.076	0.030
			(0.051)	(0.051)
Hispanic	0.111	0.206	-0.036	-0.016
			(0.035)	(0.035)
FRPL	0.174	0.351		
EOG Math Score	0.524	0.281		
EOG Reading Score	0.452	0.249		
Won Any Choice	1.000	0.395		
Won Any DL Choice	1.000	0.117		
Second Choice DL	0.408	0.315		
Third Choice DL	0.247	0.210		
Lottery FE			X	X
FRPL Dummies				X
Observations	287	248	535	535
Number of Clusters			44	44

Table 5: Summary and Balance - Estimation Sample - ESL/LEP

	Won	Lost	Test Difference Lottery FE	FRPL
	[1]	[2]	[3]	[4]
Attend DL School	0.929	0.290	0.647***	0.645***
			(0.062)	(0.062)
Female	0.442	0.455	-0.004	-0.014
			(0.069)	(0.073)
Black	0.062	0.055	0.023	0.024
			(0.027)	(0.026)
White	0.044	0.034	0.007	0.015
			(0.025)	(0.024)
Hispanic	0.823	0.890	-0.061	-0.073*
			(0.039)	(0.038)
FRPL	0.690	0.724		
EOG Math Score	0.082	-0.199		
EOG Reading Score	-0.074	-0.313		
Ever LEP	0.841	0.821		
Ever ESL	0.929	0.972		
Won Any Choice	1.000	0.297		
Won Any DL Choice	1.000	0.021		
Second Choice DL	0.195	0.234		
Third Choice DL	0.142	0.117		
Lottery FE			X	X
FRPL Dummies				X
Observations	113	145	258	258
Number of Clusters			36	36

Table 6: Impact of Attending a Dual Language School on Achievement

Panel A: English Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.471***	0.040*		0.027*	
			(0.081)	(0.020)		(0.014)	
Attend DL School	0.011	-0.010			0.086**		0.057*
	(0.017)	(0.014)			(0.043)		(0.032)
Observations	1,472	1,472	1,472	1,472	1,472	1,472	1,472
Number of Clusters	44	44	44	44	44	44	44

Panel B: ESL/LEP Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.635***	0.040**		0.044**	
			(0.084)	(0.019)		(0.020)	
Attend DL School	0.039**	0.065***			0.063**		0.069**
	(0.019)	(0.023)			(0.031)		(0.030)
Observations	809	809	809	809	809	809	809
Number of Clusters	36	36	36	36	36	36	36

Robust standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1
*Notes: Each regression includes lottery fixed effects (priority-year-program) as well as controls for female, race, frpl-year, exceptionality, grade of exam, and year of exam. Attendance is measured by whether the student attended a DL school in kindergarten or first grade and interacted with years of treatment (grade plus one). Standard errors are clustered by lottery.

Table 7: Impact of Attending a Dual Language School on Achievement

Panel A: English Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.464***	0.042*		0.024	
			(0.064)	(0.022)		(0.015)	
Attend DL School	-0.004	-0.022*			0.089*		0.053*
	(0.018)	(0.011)			(0.047)		(0.032)
Neighborhood School FE	X	X	X	X	X	X	X
Observations	1,472	1,472	1,472	1,472	1,472	1,472	1,472
Number of Clusters	44	44	44	44	44	44	44

Panel B: ESL/LEP Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.667***	0.052**		0.042*	
			(0.069)	(0.024)		(0.024)	
Attend DL School	0.052***	0.065***			0.078**		0.064**
	(0.019)	(0.019)			(0.034)		(0.032)
Neighborhood School FE	X	X	X	X	X	X	X
Observations	809	809	809	809	809	809	809
Number of Clusters	36	36	36	36	36	36	36

Robust standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1
*Notes: Each regression includes lottery fixed effects (priority-year-program) as well as controls for female, race, frpl-year, exceptionality, grade of exam, year of exam, and neighborhood school fixed effects. Attendance is measured by whether the student attended a DL school in kindergarten or first grade and interacted with years of treatment (grade plus one). Standard errors are clustered by lottery.

Table 8: Impact of Attending a Dual Language School on Achievement - Weighted

Panel A: English Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.473***	0.041*		0.027*	
			(0.082)	(0.020)		(0.015)	
Attend DL School	0.011	-0.010			0.086**		0.057*
	(0.017)	(0.014)			(0.042)		(0.031)
Observations	1,472	1,472	1,472	1,472	1,472	1,472	1,472
Number of Clusters	44	44	44	44	44	44	44

Panel B: ESL/LEP Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.637***	0.040*		0.043**	
			(0.086)	(0.020)		(0.020)	
Attend DL School	0.039**	0.065***			0.062**		0.068**
	(0.019)	(0.023)			(0.031)		(0.030)
Observations	809	809	809	809	809	809	809
Number of Clusters	36	36	36	36	36	36	36

Robust standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1
*Notes: Each regression includes lottery fixed effects (priority-year-program) as well as controls for female, race, frpl-year, exceptionality, grade of exam, and year of exam. Attendance is measured by whether the student attended a DL school in kindergarten or first grade and interacted with years of treatment (grade plus one). Regressions are weighted by the inverse probability of having test scores available in the data. Weights were generated from logit regressions. Standard errors are clustered by lottery.

Table 9: Heterogeneous Effects

Panel A: English Sample

	Gender		School Type		Race/Ethnicity		
	Female	Male	One-Way	Two-Way	White	Black	Hispanic
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Math	0.114**	0.051	0.068	0.087**	0.205**	0.063	0.076**
	(0.048)	(0.042)	(0.129)	(0.040)	(0.103)	(0.054)	(0.036)
Reading	0.083**	0.024	0.043	0.058**	0.104	0.055	0.085***
	(0.041)	(0.042)	(0.122)	(0.027)	(0.080)	(0.043)	(0.031)
Observations	1,471		1,471		1,471		
Number of Clusters	44		44		44		

Panel B: ESL/LEP Sample

	Gender		School Type		Race/Ethnicity		
	Female	Male	One-Way	Two-Way	White	Black	Hispanic
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Math	0.047	0.070**	0.111	0.058**	-	-	0.068**
	(0.048)	(0.035)	(0.154)	(0.028)	-	-	(0.032)
Reading	0.058	0.074**	0.052	0.071***	-	-	0.065**
	(0.036)	(0.035)	(0.153)	(0.025)	-	-	(0.032)
Observations	809		809		809		
Number of Clusters	36		36		36		

Robust standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1
*Notes: Each regression includes lottery fixed effects (priority-year-program) as well as controls for female, race, frpl-year, exceptionality, grade of exam, and year of exam. Attendance is measured by whether the student attended a DL school in kindergarten or first grade and interacted with years of treatment (grade plus one). Heterogeneous effects are estimated using interactions with the assignment variable and instrumenting for interactions with the attendance variable. Standard errors are clustered by lottery.

Table 10: Effects by Grade

Panel A: English Sample

	OLS		ITT		LATE	
	Math	Reading	Math	Reading	Math	Reading
	[1]	[2]	[3]	[4]	[5]	[6]
Grade						
Third	0.042	-0.103	0.162*	0.095	0.290*	0.168
	(0.093)	(0.072)	(0.089)	(0.081)	(0.160)	(0.147)
Fourth	0.115	-0.033	0.200	0.067	0.398	0.127
	(0.098)	(0.075)	(0.131)	(0.086)	(0.256)	(0.173)
Fifth	0.090	-0.037	0.135	0.229**	0.250	0.449**
	(0.139)	(0.108)	(0.133)	(0.087)	(0.249)	(0.193)
Sixth	0.035	-0.004	0.356**	0.178	0.748**	0.365
	(0.131)	(0.120)	(0.146)	(0.147)	(0.353)	(0.328)
Observations	1,471	1,471	1,471	1,471	1,471	1,471
Number of Clusters	44	44	44	44	44	44

Panel B: ESL/LEP Sample

	OLS		ITT		LATE	
	Math	Reading	Math	Reading	Math	Reading
	[1]	[2]	[3]	[4]	[5]	[6]
Grade						
Third	0.031	0.356***	0.217***	0.288***	0.362**	0.478***
	(0.107)	(0.109)	(0.079)	(0.102)	(0.144)	(0.166)
Fourth	0.355**	0.305*	0.391***	0.189	0.633**	0.316
	(0.135)	(0.161)	(0.141)	(0.151)	(0.246)	(0.248)
Fifth	0.389***	0.403***	0.283**	0.224**	0.422**	0.343**
	(0.139)	(0.140)	(0.130)	(0.104)	(0.197)	(0.157)
Sixth	0.300***	0.360*	0.277*	0.351**	0.425**	0.537***
	(0.104)	(0.185)	(0.142)	(0.148)	(0.209)	(0.206)
Observations	809	809	809	809	809	809
Number of Clusters	36	36	36	36	36	36

Robust standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1
*Notes: Each regression includes lottery fixed effects and controls for female, race, frpl-year, exceptionality, grade of exam, and year of exam. Attendance is measured by whether the student attended a DL school in kindergarten or first grade and interacted with each exam grade. Instruments are interactions between grade of exam and the indicator for winning the lottery. Standard errors are clustered by lottery.

Table 11: Impact of Dual Language Schooling on LEP Status

	Limited English Proficient			
	OLS		LATE	
	[1]	[2]	[3]	[4]
Attend DL School				
Grade 3	-0.025	0.021	-0.032	0.045
	(0.047)	(0.055)	(0.084)	(0.101)
Grade 4	-0.148**	-0.120	-0.159	-0.107
	(0.071)	(0.080)	(0.137)	(0.160)
Grade 5	-0.192**	-0.176*	-0.071	-0.051
	(0.079)	(0.093)	(0.118)	(0.144)
Grade 6	-0.210***	-0.196**	-0.168*	-0.141
	(0.073)	(0.076)	(0.086)	(0.111)
Neighborhood School FE		X		X
Observations	809	809	809	809
Number of Clusters	36	36	36	36

Robust standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1
*Notes: Each regression includes lottery fixed effects (priority-year-program) as well as controls for female, race, frpl-year, exceptionality, grade of observation, and year of observation. Attendance is measured by whether the student attended a DL school in kindergarten or first grade and interacted with grade dummy variables. Estimates are from OLS and 2SLS interaction terms. Standard errors are clustered by lottery.

Appendix

Table A1: Summary and Balance - Application Sample - English

			Test Difference	
	Won	Lost	Lottery FE	FRPL
	[1]	[2]	[3]	[4]
Attend DL School	0.870	0.358	0.542***	0.529***
			(0.047)	(0.048)
Female	0.546	0.513	0.034	0.031
			(0.045)	(0.045)
Black	0.304	0.354	-0.080*	-0.071
			(0.043)	(0.046)
White	0.372	0.268	0.084*	0.036
			(0.046)	(0.045)
Hispanic	0.112	0.215	-0.055	-0.024
			(0.038)	(0.035)
FRPL	0.181	0.373		
Ever ESL	0.000	0.000		
Non-missing Test Scores	0.847	0.821		
Won Any Choice	1.000	0.364		
Won Any DL Choice	1.000	0.113		
Second Choice DL	0.407	0.328		
Third Choice DL	0.245	0.212		
Lottery FE			X	X
FRPL Dummies				X
Observations	339	302	641	641
Number of Clusters			44	44

Table A2: Summary and Balance - Application Sample - ESL/LEP

	Won	Lost	Test Difference: Lottery FE	Test Difference: FRPL
	[1]	[2]	[3]	[4]
Attend DL School	0.891	0.275	0.615***	0.620***
			(0.065)	(0.066)
Female	0.457	0.444	0.026	0.017
			(0.050)	(0.052)
Black	0.054	0.053	0.013	0.013
			(0.024)	(0.025)
White	0.054	0.029	0.017	0.024
			(0.022)	(0.022)
Hispanic	0.829	0.895	-0.058	-0.069*
			(0.035)	(0.036)
FRPL	0.679	0.726		
Ever ESL	0.938	0.977		
Non-missing Test Scores	0.876	0.848		
Won Any Choice	1.000	0.298		
Won Any DL Choice	1.000	0.023		
Second Choice DL	0.202	0.205		
Third Choice DL	0.147	0.111		
Lottery FE			X	X
FRPL Dummies				X
Observations	129	171	300	300
Number of Clusters			39	39

Table A3: Impact of Attending a Dual Language School on Achievement - Weighted

Panel A: English Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.468***	0.042*		0.024	
			(0.065)	(0.022)		(0.015)	
Attend DL School	-0.004	-0.022*			0.089*		0.052*
	(0.018)	(0.012)			(0.046)		(0.031)
Neighborhood School FE	X	X	X	X	X	X	X
Observations	1,472	1,472	1,472	1,472	1,472	1,472	1,472
Number of Clusters	44	44	44	44	44	44	44

Panel B: ESL/LEP Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.673***	0.052**		0.043*	
			(0.071)	(0.025)		(0.025)	
Attend DL School	0.053***	0.066***			0.078**		0.064**
	(0.019)	(0.020)			(0.033)		(0.032)
Neighborhood School FE	X	X	X	X	X	X	X
Observations	809	809	809	809	809	809	809
Number of Clusters	36	36	36	36	36	36	36

Robust standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1
*Notes: Each regression includes lottery fixed effects (priority-year-program) as well as controls for female, race, frpl-year, exceptionality, grade of exam, year of exam, and neighborhood school fixed effects. Attendance is measured by whether the student attended a DL school in kindergarten or first grade and interacted with years of treatment (grade plus one). Regression are weighted by the inverse probability of having test scores available in the data. Weights were generated from logit regressions. Standard errors are clustered by lottery.

Table A4: Heterogeneous Effects

Panel A: English Sample

	Gender		School Type		Race/Ethnicity		
	Female	Male	One-Way	Two-Way	White	Black	Hispanic
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Math	0.106**	0.068	0.081	0.090*	0.190**	0.046	0.090*
	(0.054)	(0.046)	(0.127)	(0.046)	(0.076)	(0.053)	(0.048)
Reading	0.069**	0.030	0.032	0.054**	0.084	0.034	0.115***
	(0.034)	(0.043)	(0.104)	(0.028)	(0.055)	(0.034)	(0.037)
Neighborhood School FE	X		X		X		
Observations	1,472		1,472		1,472		
Number of Clusters	44		44		44		

Panel B: ESL/LEP Sample

	Gender		School Type		Race/Ethnicity		
	Female	Male	One-Way	Two-Way	White	Black	Hispanic
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Math	0.051	0.090**	-0.018	0.079**	-	-	0.083***
	(0.048)	(0.038)	(0.176)	(0.034)	-	-	(0.031)
Reading	0.011	0.087**	-0.055	0.065**	-	-	0.062**
	(0.045)	(0.039)	(0.157)	(0.031)	-	-	(0.031)
Neighborhood School FE	X		X		X		
Observations	809		809		809		
Number of Clusters	36		36		36		

Robust standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1
*Notes: Each regression includes lottery fixed effects (priority-year-program) as well as controls for female, race, frpl-year, exceptionality, grade of exam, year of exam, and neighborhood school fixed effects. Attendance is measured by whether the student attended a DL school in kindergarten or first grade and interacted with years of treatment (grade plus one). Heterogeneous effects are estimated using interactions with the assignment variable and instrumenting for interactions with the attendance variable. Standard errors are clustered by lottery.

Table A5: Effects by Grade

Panel A: English Sample

	OLS		ITT		LATE	
	Math [1]	Reading [2]	Math [3]	Reading [4]	Math [5]	Reading [6]
Grade						
Third	-0.011	-0.152*	0.189**	0.108	0.374**	0.215
	(0.120)	(0.082)	(0.092)	(0.079)	(0.171)	(0.147)
Fourth	0.042	-0.086	0.211	0.070	0.446	0.151
	(0.112)	(0.078)	(0.143)	(0.093)	(0.284)	(0.188)
Fifth	-0.009	-0.115	0.139	0.216**	0.264	0.431**
	(0.138)	(0.094)	(0.150)	(0.096)	(0.276)	(0.205)
Sixth	-0.107	-0.107	0.327**	0.136	0.721**	0.298
	(0.121)	(0.091)	(0.141)	(0.135)	(0.346)	(0.295)
Neighborhood School FE	X	X	X	X	X	X
Observations	1,472	1,472	1,472	1,472	1,472	1,472
Number of Clusters	44	44	44	44	44	44

Panel B: ESL/LEP Sample

	OLS		ITT		LATE	
	Math [1]	Reading [2]	Math [3]	Reading [4]	Math [5]	Reading [6]
Grade						
Third	0.055	0.312***	0.244**	0.244*	0.393**	0.390**
	(0.097)	(0.101)	(0.109)	(0.128)	(0.174)	(0.187)
Fourth	0.374**	0.254*	0.413**	0.153	0.657**	0.238
	(0.138)	(0.134)	(0.171)	(0.167)	(0.274)	(0.248)
Fifth	0.397***	0.370***	0.320*	0.208	0.471**	0.302
	(0.131)	(0.100)	(0.163)	(0.137)	(0.229)	(0.188)
Sixth	0.417***	0.408***	0.367**	0.364**	0.542**	0.531***
	(0.134)	(0.145)	(0.171)	(0.140)	(0.231)	(0.183)
Neighborhood School FE	X	X	X	X	X	X
Observations	809	809	809	809	809	809
Number of Clusters	36	36	36	36	36	36

Robust standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1
*Notes: Each regression includes lottery fixed effects and controls for female, race, frpl-year, exceptionality, grade of exam, year of exam, and neighborhood school fixed effects. Attendance is measured by whether the student attended a DL school in kindergarten or first grade and intearacted with each exam grade. Instruments are interactions between grade of exam and the indicator for winning the lottery. Standard errors are clustered by lottery.

Table A6: Impact of Dual Language Education - Constant Effect

Panel A: English Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.498***	0.203*		0.130	
			(0.067)	(0.109)		(0.078)	
Attend DL School	0.060	-0.065			0.407*		0.261*
	(0.089)	(0.072)			(0.214)		(0.156)
Observations	1,472	1,472	1,472	1,472	1,472	1,472	1,472
Number of lotfe	44	44	44	44	44	44	44

Panel B: ESL/LEP Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.629***	0.232**		0.243**	
			(0.075)	(0.106)		(0.113)	
Attend DL School	0.207*	0.357***			0.369**		0.385**
	(0.107)	(0.126)			(0.174)		(0.175)
Observations	809	809	809	809	809	809	809
Number of lotfe	36	36	36	36	36	36	36

Robust standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1
*Notes: Each regression includes lottery fixed effects (priority-year-program) as well as controls for female, race, frpl-year, exceptionality, grade of exam, and year of exam. Attendance is measured by whether the student attended a DL school in kindergarten or first grade. The treatment and attendance variables are not interacted with years of treatment in this specification. Standard errors are clustered by lottery.

Table A7: Impact of Dual Language Education - 3rd Grade Attendance Measure

Panel A: English Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.394***	0.203*		0.130	
			(0.055)	(0.109)		(0.078)	
Attend DL School (3rd)	0.178*	0.042			0.514*		0.329*
	(0.097)	(0.084)			(0.277)		(0.198)
Observations	1,472	1,472	1,472	1,472	1,472	1,472	1,472
Number of lotfe	44	44	44	44	44	44	44

Panel B: ESL/LEP Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.576***	0.232**		0.243**	
			(0.064)	(0.106)		(0.113)	
Attend DL School (3rd)	0.233**	0.357**			0.403**		0.421**
	(0.097)	(0.138)			(0.202)		(0.193)
Observations	809	809	809	809	809	809	809
Number of lotfe	36	36	36	36	36	36	36

Robust standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1
*Notes: Each regression includes lottery fixed effects (priority-year-program) as well as controls for female, race, frpl-year, exceptionality, grade of exam, and year of exam. Attendance is measured by whether the student attended a DL school in third grade. The treatment and attendance variables are not interacted with years of treatment in this specification. Standard errors are clustered by lottery.

Table A8: Grades Three Through Five Only

Panel A: English Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.513***	0.032		0.023*	
			(0.063)	(0.021)		(0.013)	
Attend DL School	0.013	-0.012			0.063		0.044*
	(0.019)	(0.014)			(0.040)		(0.026)
Observations	1,172	1,172	1,172	1,172	1,172	1,172	1,172
Number of lotfe	44	44	44	44	44	44	44

Panel B: ESL/LEP Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.513***	0.032		0.023*	
			(0.063)	(0.021)		(0.013)	
Attend DL School	0.045*	0.067**			0.075**		0.060*
	(0.024)	(0.025)			(0.037)		(0.036)
Observations	623	623	623	623	623	623	623
Number of lotfe	36	36	36	36	36	36	36

Robust standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1
*Notes: Each regression includes lottery fixed effects (priority-year-program) as well as controls for female, race, frpl-year, exceptionality, grade of exam, and year of exam. Attendance is measured by whether the student attended a DL school in kindergarten or first grade and interacted with years of treatment (grade plus one). Standard errors are clustered by lottery.

Table A9: Grades Three Through Five Only

Panel A: English Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.503***	0.034		0.023*	
			(0.054)	(0.024)		(0.014)	
Attend DL School	-0.001	-0.023			0.068		0.046*
	(0.022)	(0.013)			(0.045)		(0.027)
Neighborhood School FE	X	X	X	X	X	X	X
Observations	1,172	1,172	1,172	1,172	1,172	1,172	1,172
Number of Clusters	44	44	44	44	44	44	44

Panel B: ESL/LEP Sample

	OLS			Math		Reading	
	Math	Reading	First Stage	ITT	LATE	ITT	LATE
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Won First Choice			0.652***	0.055*		0.028	
			(0.061)	(0.031)		(0.030)	
Attend DL School	0.049**	0.058**			0.084*		0.043
	(0.023)	(0.022)			(0.047)		(0.043)
Neighborhood School FE	X	X	X	X	X	X	X
Observations	623	623	623	623	623	623	623
Number of Clusters	36	36	36	36	36	36	36

Robust standard errors in parentheses

***p<0.01, **p<0.05, *p<0.1

*Notes: Each regression includes lottery fixed effects (priority-year-program) as well as controls for female, race, frpl-year, exceptionality, grade of exam, year of exam, and neighborhood school fixed effects. Attendance is measured by whether the student attended a DL school in kindergarten or first grade and interacted with years of treatment (grade plus one). Standard errors are clustered by lottery.

Table A10: Cohort Interactions

	English Sample		**ESL/LEP Sample**	
	Math	Reading	Math	Reading
	[1]	[2]	[3]	[4]
Attend DL School				
2007 Cohort	0.316*	0.209	0.088	0.145***
	(0.166)	(0.157)	(0.055)	(0.050)
2008 Cohort	0.016	0.006	0.032	-0.006
	(0.036)	(0.034)	(0.050)	(0.019)
2009 Cohort	-0.002	0.063***	0.064**	0.096
	(0.091)	(0.013)	(0.029)	(0.061)
2010 Cohort	0.122**	0.098***	0.284**	0.313***
	(0.054)	(0.033)	(0.136)	(0.117)
2011 Cohort	0.081**	0.009	0.024	0.043
	(0.037)	(0.029)	(0.081)	(0.073)
Observations	1,472	1,472	809	809
Number of Lottery FE	44	44	36	36

Robust standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1
*Notes: Each regression includes lottery fixed effects (priority-year-program) as well as controls for female, race, frpl-year, exceptionality, grade of exam, and year of exam. Reported coefficients are on interactions between the attendance variable and cohort. Attendance is measured by whether the student attended a DL school in kindergarten or first grade and interacted with years of treatment (grade plus one). Standard errors are clustered by lottery.

Table A11: Cohort Interactions

	English Sample		ESL/LEP Sample	
	Math	Reading	Math	Reading
	[1]	[2]	[3]	[4]
Attend DL School				
2007 Cohort	0.401**	0.200*	0.105***	0.133***
	(0.157)	(0.116)	(0.038)	(0.031)
2008 Cohort	-0.009	-0.014	0.048	-0.012
	(0.035)	(0.030)	(0.063)	(0.028)
2009 Cohort	0.074	0.132**	0.032	0.045
	(0.109)	(0.058)	(0.050)	(0.086)
2010 Cohort	0.151**	0.153***	0.286	0.239
	(0.069)	(0.040)	(0.197)	(0.166)
2011 Cohort	0.131***	0.029	0.067	0.061
	(0.044)	(0.034)	(0.086)	(0.096)
Neighborhood School FE	X	X	X	X
Observations	1,471	1,471	809	809
Number of Lottery FE	44	44	36	36

Robust standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1
*Notes: Each regression includes lottery fixed effects (priority-year-program) as well as controls for female, race, frpl-year, exceptionality, grade of exam, year of exam, and neighborhood school fixed effects. Reported coefficients are on interactions between the attendance variable and cohort. Attendance is measured by whether the student attended a DL school in kindergarten or first grade and interacted with years of treatment (grade plus one). Standard errors are clustered by lottery.