THE RELATIONSHIP BETWEEN CLASSROOM INSTRUCTIONAL LANGUAGE AND ACADEMIC ACHIEVEMENT AMONG ENGLISH LANGUAGE LEARNER STUDENTS

A Thesis
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the degree of
Master of Public Policy

By

Kristina Leticia Rodriguez, B.S.

Washington, D.C.
April 11, 2017
Nearly one-third of children living in the United States are English language learners (ELLs). These children’s academic achievement lags behind the achievement of their English-proficient peers. As more ELL students enter the public school system, educators require guidance regarding the best practices for closing this gap. A particularly important consideration in this regard relates to the question of whether students’ native languages are used in the classroom, or whether students are instead exposed to English-only instruction. This paper explores the relationship between instructional language and academic achievement among ELL students. Using data from the Early Childhood Longitudinal Study 2011, I find that there is a small, negative, and statistically significant relationship between non-English use in the classroom and reading test scores. I find no relationship between non-English use in the classroom and math test scores.
The research and writing of this thesis
is dedicated to everyone who helped along the way.

I would like to particularly thank:

Dr. Adam T. Thomas for his patience, thoughtfulness, and encouragement.

Emery Mathieson for his unwavering moral support and helpful edits.

Susan Hill for her careful proofreading.

Amaya Garcia for her initial suggestions and guidance regarding the topic.

Many thanks,
Kristina Leticia Rodriguez
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Background</td>
<td>3</td>
</tr>
<tr>
<td>Literature Review</td>
<td>5</td>
</tr>
<tr>
<td>Conceptual Framework</td>
<td>9</td>
</tr>
<tr>
<td>Data and Methods</td>
<td>12</td>
</tr>
<tr>
<td>Descriptive Statistics</td>
<td>16</td>
</tr>
<tr>
<td>Results</td>
<td>19</td>
</tr>
<tr>
<td>Discussion</td>
<td>26</td>
</tr>
<tr>
<td>Appendix: Full Models Including Motivation Variable</td>
<td>30</td>
</tr>
<tr>
<td>References</td>
<td>32</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. Factors Associated with Academic Achievement ........................................................... 9
LIST OF TABLES

Table 1. Variable Definitions................................................................. 14
Table 2. Descriptive Statistics............................................................... 17
Table 3. Descriptive Statistics, Disaggregated by Language of Instruction ......................................................... 18
Table 4. Estimates of the Relationship between Non-English Use and Reading Scores.......................... 21
Table 5. Estimates of the Relationship between Non-English Use and Math Scores......................... 24
Table 6. Full Models Including Motivation Variable................................................................................. 30
INTRODUCTION

The United States population has changed dramatically over the past several decades with the number of immigrants rising from nearly 10 million in 1970 to over 40 million in 2014 (Migration Policy Institute Data Hub, 2016). The debate over how policymakers should best respond to this demographic shift has placed immigration high on the country’s political agenda (Pew Research Center, 2016). Much of this attention has focused on broad issues surrounding labor, national security, and pathways to citizenship (White House, 2011). However, issues concerning young immigrants, children of immigrants, and the education of said groups are poised to gain attention.

Children living in immigrant families are the fastest growing demographic of children in the US (Hernandez, Denton, & Macartney, 2008). Most of these children are born into homes in which English is not the primary language spoken and enter school as English Language Learners (ELLs) (Morse, 2005). Currently, about one-third of children in the US live in a household in which the family speaks a language other than English (Child Trends Databank, 2014). As this population of children continues to grow, school districts around the country are pressed to provide greater ELL-focused resources and trainings to teachers so that they are able to appropriately instruct these students.

The present state of academic achievement among young ELLs is not promising. The 2015 Nation’s Report Card, which reports results from the National Assessment of Educational Progress, shows a 36-point difference on reading assessments and a 25-point difference on math assessments between ELLs and their English-speaking peers in fourth grade (US Office of English Language Acquisition, 2016). More generally, only about 19% of ELLs between the

---

---

1 Average scores for non-EL students are 225 out of 500 for reading and 243 out of 500 for math.
ages of five and seventeen are proficient in English (Migration Policy Institute Data Hub, 2014). This combination of challenges presents a unique challenge for school-aged ELLs, who must learn to communicate both in English and in their home languages, while also learning academic subject content (Espinosa, 2008).

Guidance from the US Department of Education suggests that ELLs’ academic achievement and English proficiency levels are best supported when teachers are properly trained in instructional approaches for ELLs, including the best methods for integrating students’ home languages into the classroom (US Department of Education, 2016; Escamilla, 2009). Using nationally representative data from The Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011), this paper examines the relationship between instructional language and academic performance among young ELL students.

The remainder of my paper is arranged as follows. In the background, I discuss the policy context surrounding how ELLs students are served by the US public education system and provide information regarding instructional approaches for ELL students. In the literature review, I summarize the literature on the relationship between instructional language and academic achievement and ELLs. An overview of my study’s framework, hypothesis, and analytical model follow the review. Lastly, I present my empirical findings, describe my study’s limitations, and discuss the policy implications of my results.
BACKGROUND

Under the US Civil Rights Act of 1964 and the Equal Educational Opportunities Act of 1974, schools are obligated to provide “meaningful and equal” educational opportunities for ELLs (US Department of Education, 2015a). In accordance with these regulations, the Department of Education and the Department of Justice provide joint guidance to state education agencies, districts, and even individual schools on their legal obligations to ELL students. This guidance includes requirements for schools to ensure that ELL students have access to language assistance, qualified teachers, and additional support as needed (US Department of Education, 2015b).

In addition to supporting the civil rights of ELLs, the Department of Education has adapted federal education policy to address the challenges that school districts face in serving the growing population of ELLs and their families. On December 10, 2015, the Department reauthorized the No Child Left Behind Act (NCLB), the country’s national education law, as Every Student Succeeds Act (ESSA) (US Department of Education, 2015c). Previously, the original passage of NCLB included requirements for schools to ensure English proficiency among their students. Schools were also required to implement and report the results of student assessments, among other accountability measures (US Department of Education, 2002).

The ESSA reauthorization included several revisions to the original NCLB requirements that schools and educators deemed too difficult to implement (US Department of Education, 2015c). Most notably for the purposes of this study, the ESSA aims to increase state accountability for ensuring that educators and school leaders can properly identify ELLs and provide them with appropriate instruction (US Department of Education, 2015c). However, the Working Group on ELL Policy (2016) argues that the law does not provide sufficiently strong guidance in terms of how teachers should instruct ELL students and how progress should be
measured. The Working Group recommends that school districts consider a growing evidence base when selecting instructional approaches for their ELL students.

Currently, schools adopt one or more of several ELL instructional approaches. These approaches, which are collectively referred to as Language Instruction Education Programs (LIEPs), have one of two main foci: developing literacy in two languages, or developing English-only literacy (National Clearinghouse for English Language Acquisition, 2013). Approaches that adopt the former focus include two-way immersion/dual language programs and transitional bilingual education programs (US Department of Education, 2015a). As the name implies, dual language programs aim to develop students’ academic knowledge in two languages. These programs provide classroom environments that cultivate proficiency in both English and a student’s native language (Moughamian, Rivera, & Francis, 2009). Transitional bilingual education programs provide a greater amount of instruction in students’ native languages, scaffolding the use of English language instruction (beginning with native language instruction) as their English language skills increase (Genesee, 1999).

Programs with an English-only focus include English as a Second Language (ESL) and Structured English Immersion (US Department of Education, 2015c). In ESL programs, also known as sheltered instruction programs, students learn in an all-English setting, and teachers often use visual aides or other language supports (Moughamian, Rivera, & Francis, 2009). Structured English Immersion programs are similar to ESL programs, but are most commonly used in classrooms with only ELL students, while ESL programs are used in mixed classrooms (National Clearinghouse for English Language Acquisition, 2013).
LITERATURE REVIEW

To date, there is no clear scientific consensus regarding which of the instructional approaches outlined above is best for ELL students. Over the past 20 years, some research suggests that academic outcomes among ELLs are positively associated with the use of their native languages in the classroom, other research suggests a negative association, and still, other studies find that there is no relationship at all. Much of this research focuses on whether a given approach helps students: (a) achieve English language proficiency; and (b) meet academic standards with respect to reading, math, and science (August & Shanahan, 2006; Lee, 2005). This research is motivated in large part by ELLs’ low performance on standardized tests within these subject areas (United States Office of English Language Acquisition, 2016).

The Positive Relationship Between Bi/Multilingual Instruction and Achievement

In a synthesis of the relevant research, August and Shanahan (2006) find that most of the studies included in their review favor the use of native language instruction, particularly with respect to students’ reading achievement. Syntheses by Cummings (2000) and Rolstad et. al. (2005) also suggest that native language use in the classroom is associated with improved academic outcomes for ELLs.

Some individual studies focus specifically on the relationship between instructional language and language development for young children. In an experimental study, Barnett et. al. (2007) use random assignment to place preschool students in bilingual classrooms versus English-only classrooms. Their results suggest a positive relationship between instruction in the bilingual classroom and progress in both English and Spanish language development for ELLs (Barnett et. al., 2007). Using a similar research design, Duran et. al. (2010) find that a bilingual approach is associated with growth in language development for ELLs, specifically when they
are assessed in Spanish. They see no significant difference when the assessments are conducted in English (Duran et. al., 2010).

Other studies examine outcomes for older students. Using data collected nationwide between 1991 and 2001, Thomas and Collier (2002) find that ELLs receiving English-only instruction scored lower on standardized reading tests by fifth grade than students receiving bilingual instruction. Lindholm-Leary and Hernandez (2011) find that, among Spanish-speaking, Latino students, participation in dual language programs—in which students are taught both in English and in their native language—is associated with higher reading achievement. However, the authors note that their comparison group contained non-Latino students which may have biased their estimates and caused them to understate the effect. Regardless, this result is consistent with findings documented in other research. For example, Lindholm-Leary and Block (2009) find that Latino ELL students in bilingual programs perform as well as—if not better than—their peers in English-only programs on reading measures. It is important to note that the authors focus this study on students in low-SES and predominantly Hispanic schools, which may diminish the external validity of these findings. In a longitudinal study of Spanish and Chinese speaking ELL students, Valentino and Reardon (2014) examine academic outcomes associated with four different instructional approaches, including bilingual programs. Generally, they find that students in English-only settings perform more poorly on measures of reading and math than their peers.²

Outside of the Valentino and Reardon study, research on the relationship between the language of instruction and student achievement in math and science among ELLs in the US is very limited compared to research for reading achievement. Lindholm-Leary and Barsato (2005)

---
² With the exception of one math measure, where students in a specific type of bilingual program performed more poorly than students in English-only programs.
examine the math performance of Hispanic high schoolers who participated in bilingual programs during elementary school. The authors find that students for whom follow-up data were available scored seven points higher than the statewide average. However, the sample size for the authors’ follow-up group was 19 students, which suggests that these results should be interpreted with substantial caution. In addition, a qualitative study of Kenyan elementary school science instruction by Cleghorn (1992) concluded that students found the concepts they were learning to be more accessible when teachers sometimes used local languages as part of their instruction, rather than solely using English.

*The Negative or Negligible Relationship Between Bi/Multilingual Instruction and Achievement*

Relevant literature also includes studies where findings present the negative or negligible effects of participation in bilingual programs. Ramirez et. al. (1991) conduct a longitudinal study of ELLs where findings suggest students in English-only classrooms perform better than their peers.³ Based on their review of studies on bilingual education programs, Baker and de Kanter (1981) conclude that there is insufficient evidence to believe that a positive association exists between bilingual instruction and student achievement. A later synthesis by Rossell and Baker (1996) reaches a similar conclusion. More recently, Slavin, and Cheung (2005) also present a review of newer research where most of the included studies find little to no difference in outcomes based on instructional language. Matsudaira (2005) employs a regression discontinuity design to compare English-only and bilingual approaches. The author’s study takes advantage of the fact that students are eligible for ELL services if their score places them below the 40th percentile on an English proficiency test. He finds that there is no statistically significant difference between students in these two types of classrooms (Matsudaira, 2005). More

³ These findings hold true in the early grades only.
compellingly, in a recent multiyear randomized evaluation of a bilingual education program, Slavin et. al. (2011) find that students in the bilingual program performed more poorly on assessments administered in English than their English-only counterparts.

The Present Study

Currently, literature exploring the relationship between instructional language and achievement does not provide a consensus on which approach is best for ELLs. Additionally, many of the studies presented in the review use data for specific subgroups, or groups of students from particular areas in the country. The present study contributes to the literature above by using the most recent available data to assess the relationship between instructional language and student achievement in math and reading. My data are both more recent and more nationally representative than the data used in many previous studies (Lindholm-Leary & Barsato, 2005; Lindholm-Leary & Block, 2009). The study also contributes to the literature by considering ELLs’ math achievement, which is an understudied outcome in the existing literature.
CONCEPTUAL FRAMEWORK

In order to investigate the relationship between instructional language and student achievement among ELLs, my model accounts for teacher, school, and student/family factors that are plausibly related to both variables. These factors are shown in Figure 1 and are discussed in more detail below.

Figure 1. Factors Associated with Academic Achievement
Teacher and Classroom Characteristics

Given that teachers are integral to students’ classroom experiences, it is plausible that teacher characteristics could influence academic outcomes. Mainly, teacher quality is considered a predictor of student academic outcomes (Darling-Hammond, 2000). I operationalize teacher quality using measures of certification status and degree type. Additionally, years of teaching experience is positively associated with higher student academic outcomes, particularly in reading (Croninger et al., 2007). The quality of students’ relationships with their teachers is also potentially important. Specifically, positive student-teacher relationships have been found to be associated with better academic performance (Hamre & Pianta, 2001). The quality of these relationships is operationalized by including data from a five-point scale measuring closeness.

School Characteristics

There is variation among schools in terms of the resources available to students and teachers, which may in turn influence students’ academic achievement. More specifically, students tend to perform better in schools and districts with greater resources (Greenwald et al., 1996). Districts with high rates of poverty or that serve large minority populations may not be able to provide students with adequate resources.

Student and Family Characteristics

Factors specific to a student’s home environment may also influence academic achievement. Parent education is particularly relevant for ELLs, who are more academically successful when their parents have higher levels of education (Lindholm-Leary & Hernandez, 2011). Students living in single-parent and low-income households are likely to have lower academic achievement (Amato, 2001; Aud et al., 2010). Children who experience food insecurity or poor nutrition are also more likely to perform poorly (Alaimo, 2001).
Academic achievement may be a function of a student’s individual characteristics as well. In particular, academic outcomes for students can vary by disability status, race, and gender (National Center for Education Statistics, 2013; Thomas & Collier, 2002). It is also important to consider students’ native languages when examining their academic performance.\textsuperscript{4} I considered controlling for native language, but given an enormous overlap between Hispanic ethnicity and Spanish language (about 60 percent), I only control for Hispanic ethnicity.

\textsuperscript{4} A child’s native language may also influence performance in bi/multilingual classroom settings based on specific features of the language (Genesee et. al., 2008; Valentino & Reardon, 2015).
DATA AND METHODS

I analyze individual-level data from The Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011), which is a nationally representative sample of 18,174 students. Data were collected from students, families, teachers, schools and care providers regarding students’ cognitive, social-emotional, and physical development. The survey’s questions gather information on students’ home environments, school environments, classroom environments, classroom curricula, and teacher qualifications. I specifically use ECLS-K data for first grade students who speak a language other than English at home. Data for first graders were collected in the fall and spring of the academic year 2011-2012. My outcome variables are measured using reading and math assessment results. The reading assessment measures basic skills such as letter recognition, vocabulary, and constructing meaning from print. The math assessment tests understanding of early math skills including number literacy, pattern recognition, and measurement of basic quantities like length.

As outlined in the conceptual framework, my control variables include teacher factors, school factors, and student/family characteristics. These data were all collected via parent interviews, teacher surveys, or administrator surveys. I estimate separate regression models with

---

5 My analytic sample includes 1,405 students who were identified as speaking a language other than English at home.
6 I excluded observations for which I had missing or insufficient data for my dependent and key independent variables from my final analytic sample.
7 Due to language barriers experienced by ELL students, researchers administered an English proficiency screener to students prior to administering the subject assessments. Those who passed the screener were given English versions of the math and reading tests. Students who did not pass the screener were administered translated versions of the math and reading tests.
8 Nine out of my 15 covariates had missing data. Missing values were imputed using single imputation with the other covariates as independent variables. Overall, 9.29% of the data points in my analysis file were imputed. A series of t-tests reveals that there were statistically significant differences between observations with and without imputed values in terms of teacher factors, school factors, and student and family factors between the groups. For example, imputed observations were more likely to have Master’s degrees and IEPs, while being less likely to be food secure. Imputed observations also had lower district poverty rates.
math and reading scores as the dependent variables. The general model is specified as follows.

Table 1 provides definitions for all variables included in the regression.

\[ TestScore = \beta_0 + \beta_1 NonEnglishInstruction + \beta_2 YearsExperience + \beta_3 Masters + \beta_4 Closeness + \beta_5 DistrictPoverty + \beta_6 SchoolELL + \beta_7 ParentHighSchool + \beta_8 Disability + \beta_9 FoodSecure + \beta_{10} TwoParents + \beta_{11} Poverty + \beta_{12} Hispanic + \beta_{13} WhiteNotHisp + \beta_{14} BlackNotHisp + \beta_{15} Female + \mu \]
Table 1. Variable Definitions

<table>
<thead>
<tr>
<th>Variables</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variables (TestScore)</strong></td>
<td></td>
</tr>
<tr>
<td>Math</td>
<td>A continuous variable measuring a student’s early math skills as measured using the Test of Early Mathematics Ability – 3rd edition (TEMA-3). Scores range from 1 to 100.(^9)</td>
</tr>
<tr>
<td>Reading</td>
<td>A continuous variable measuring a student’s early reading skills as measured using the Test of Early Reading Ability – 3rd edition (TERA-3). Scores range from 1 to 100.(^10)</td>
</tr>
<tr>
<td><strong>Key Independent Variable</strong></td>
<td></td>
</tr>
<tr>
<td>NonEnglishUse</td>
<td>A dichotomous variable indicating whether a student’s teacher uses a language other than English at least some of the time in the classroom.</td>
</tr>
<tr>
<td><strong>Teacher ad Classroom Characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>YearsExperience</td>
<td>A continuous variable measuring a teacher’s years of experience.</td>
</tr>
<tr>
<td>Masters</td>
<td>A dichotomous variable measuring whether a teacher has a master’s degree.</td>
</tr>
<tr>
<td>Closeness</td>
<td>An ordinal variable measuring affection, warmth, and open communication between the teacher and student on a five-point scale.</td>
</tr>
<tr>
<td><strong>School Characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>DistrictPoverty</td>
<td>A continuous variable measuring the percentage of children in the student’s school district who are in poverty.</td>
</tr>
<tr>
<td>SchoolELL</td>
<td>A continuous variable measuring the percentage the student body at the respondent’s school who are English Language Learners.</td>
</tr>
<tr>
<td><strong>Student and Family Characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>ParentHighSchool</td>
<td>A dichotomous variable measuring whether the student has at least one parent with a high school education.</td>
</tr>
<tr>
<td>Disability</td>
<td>A dichotomous variable indicating whether the student has a disability.</td>
</tr>
<tr>
<td>FoodSecure</td>
<td>A dichotomous variable indicating whether the student’s household is food secure.(^11)</td>
</tr>
<tr>
<td>TwoParents</td>
<td>A dichotomous variable indicating whether the student lives with both of her or his parents.</td>
</tr>
<tr>
<td>Poverty</td>
<td>A dichotomous variable indicating whether the student’s household is below the poverty threshold. Researchers classified household poverty status according to the US Census Bureau Poverty Thresholds for 2011 (Tourangeau, et al., 2015).</td>
</tr>
<tr>
<td>Hispanic</td>
<td>A dichotomous variable indicating whether the student is Hispanic.</td>
</tr>
</tbody>
</table>

\(^9\) Raw test scores were rescaled by the survey’s administrators using Item Response Theory (IRT) procedures. Rescaled scores are generally preferred to raw scores for research in which scores are compared between individual children or between groups of children (Tourangeau, et al., 2015).

\(^10\) See previous footnote.

\(^11\) The food security measure is created by combining several different responses from the ECLS-K parent interview in order to determine the extent to which a child’s household is geographically proximate to food source and has resources to obtain food.
Table 1. (cont.)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>WhiteNotHisp</td>
<td>A dichotomous variable indicating whether the student is White non-</td>
</tr>
<tr>
<td></td>
<td>Hispanic.</td>
</tr>
<tr>
<td>BlackNotHisp</td>
<td>A dichotomous variable indicating whether the student is Black non-</td>
</tr>
<tr>
<td></td>
<td>Hispanic.</td>
</tr>
<tr>
<td>Female</td>
<td>A dichotomous variable indicating whether the student is female.</td>
</tr>
</tbody>
</table>
DESCRIPTIVE STATISTICS

Table 2 reports descriptive statistics for the dependent and key independent variables in addition to the teacher, school, and student/family control variables. Table 3 provides descriptive statistics for all variables disaggregated by language of instruction. All descriptive analyses are weighted using a sample weight provided by the survey’s administrators to adjust for nonresponse.

Among the members of my sample, 31 percent are in classrooms where the teacher uses a second language in addition to English. Students’ average reading assessment scale score is 66.93, and the average math scale score 59.93 (both are on a scale of 1 to 100). On average, students attended schools where about one-third (31 percent) of students are ELLs and the average poverty rate in students’ school districts is about 27 percent.

Notably, nearly half (47 percent) of students live in households with incomes below the federal poverty line, and 15 percent live in households that are food insecure. The sample is also predominantly Hispanic (69 percent), while 11 percent of children identified as White non-Hispanic, and 2 percent identified as Black non-Hispanic. The gender distribution within my sample is about evenly split: 48 percent of respondents are female.
Table 2. Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading Score</td>
<td>66.93</td>
<td>0.41</td>
<td>25.27</td>
<td>94.84</td>
</tr>
<tr>
<td>Math Score</td>
<td>59.94</td>
<td>0.40</td>
<td>16.46</td>
<td>93.99</td>
</tr>
<tr>
<td><strong>Key Independent Variable</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uses Language Other than English</td>
<td>0.31</td>
<td>0.01</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Teacher Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Master’s Degree</td>
<td>0.49</td>
<td>0.01</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Years of Experience</td>
<td>13.87</td>
<td>0.26</td>
<td>1</td>
<td>37</td>
</tr>
<tr>
<td>Closeness with Student</td>
<td>4.17</td>
<td>0.02</td>
<td>1.14</td>
<td>5</td>
</tr>
<tr>
<td><strong>School Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% District Poverty</td>
<td>0.27</td>
<td>0.46</td>
<td>0.02</td>
<td>0.54</td>
</tr>
<tr>
<td>% ELL Students</td>
<td>0.31</td>
<td>0.64</td>
<td>1</td>
<td>0.70</td>
</tr>
<tr>
<td><strong>Student and Family Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent HS Education</td>
<td>0.59</td>
<td>0.01</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Lives with both Parents</td>
<td>0.74</td>
<td>0.01</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Household Poverty</td>
<td>0.47</td>
<td>0.01</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Food Secure</td>
<td>0.85</td>
<td>0.01</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Motivation</td>
<td>3.64</td>
<td>0.02</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Disability</td>
<td>0.11</td>
<td>0.01</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.69</td>
<td>0.01</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>White Not Hisp</td>
<td>0.11</td>
<td>0.01</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Black Not Hisp</td>
<td>0.02</td>
<td>0.004</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Female</td>
<td>0.48</td>
<td>0.01</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Sample size = 1,405

Table 3 presents descriptive statistics disaggregated by classroom language use. Overall, students in English-only classrooms have higher scores on both the reading and math assessments than their peers in bi/multilingual classrooms. However, the English-only students are more advantaged than their peers across several measures. Students in bi/multilingual classrooms tend to have teachers with less education and experience. They also attend schools in districts with higher poverty rates and larger numbers of ELL students. Disparities extend to students’ demographic and family factors as well. Their parents have lower levels of education, and their households are more likely to be poor. This pattern of disadvantage may be driving the gap in students’ test scores. Using multiple regressions, which are discussed in the next section, I
measure the relationship between the language of instruction and test scores while holding these other factors constant.

**Table 3.** Descriptive Statistics, Disaggregated by Language of Instruction

<table>
<thead>
<tr>
<th></th>
<th>Other Language Use n=884</th>
<th>English-Only n=2,325</th>
<th>Difference</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading Score</td>
<td>62.36</td>
<td>69.57</td>
<td>-7.21***</td>
<td>0.85</td>
</tr>
<tr>
<td>Math Score</td>
<td>57.28</td>
<td>61.86</td>
<td>-4.66***</td>
<td>0.84</td>
</tr>
<tr>
<td><strong>Teacher and Classroom Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Master’s Degree</td>
<td>0.38</td>
<td>0.51</td>
<td>-0.13***</td>
<td>0.03</td>
</tr>
<tr>
<td>Years of Experience</td>
<td>12.03</td>
<td>14.68</td>
<td>-2.65***</td>
<td>0.54</td>
</tr>
<tr>
<td>Closeness with Student</td>
<td>4.19</td>
<td>4.19</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td><strong>School Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% District Poverty</td>
<td>0.35</td>
<td>0.24</td>
<td>0.11 ***</td>
<td>0.91</td>
</tr>
<tr>
<td>% ELL Students</td>
<td>0.42</td>
<td>0.26</td>
<td>0.16 ***</td>
<td>0.13</td>
</tr>
<tr>
<td><strong>Student and Family Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent HS Education</td>
<td>0.43</td>
<td>0.66</td>
<td>-0.23***</td>
<td>0.03</td>
</tr>
<tr>
<td>Lives with both Parents</td>
<td>0.75</td>
<td>0.75</td>
<td>0.00</td>
<td>0.03</td>
</tr>
<tr>
<td>Household Poverty</td>
<td>0.63</td>
<td>0.40</td>
<td>0.23 ***</td>
<td>0.03</td>
</tr>
<tr>
<td>Food Secure</td>
<td>0.84</td>
<td>0.87</td>
<td>-0.03 *</td>
<td>0.02</td>
</tr>
<tr>
<td>Motivation</td>
<td>3.53</td>
<td>3.70</td>
<td>-0.17 ***</td>
<td>0.06</td>
</tr>
<tr>
<td>Disability</td>
<td>0.13</td>
<td>0.09</td>
<td>0.04 **</td>
<td>0.02</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.75</td>
<td>0.62</td>
<td>0.24 ***</td>
<td>0.03</td>
</tr>
<tr>
<td>White Non Hisp</td>
<td>0.05</td>
<td>0.14</td>
<td>-0.09 ***</td>
<td>0.02</td>
</tr>
<tr>
<td>Black Non Hisp</td>
<td>0.00</td>
<td>0.02</td>
<td>-0.02 ***</td>
<td>0.01</td>
</tr>
<tr>
<td>Female</td>
<td>0.49</td>
<td>0.49</td>
<td>0.00</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Total sample size = 1,405; *** p<0.01, ** 0.01 < p < 0.05, * 0.05 < p < 0.10
RESULTS

My regression results are presented in Tables 4 and 5 below. Table 4 reports the results of models in which reading achievement is the outcome, and Table 5 reports the results of models in which math achievement is the outcome. Both tables contain five sets of regression results; each model adds a layer of control variables to demonstrate how my three groups of controls—teacher factors, school factors, and home and individual factors—influence the key relationship by reducing omitted variable bias. Model 1 is a naïve Ordinary Least Squares (OLS) regression without controls; Model 2 adds controls for teacher-related factors; Model 3 adds controls for school-related factors; and Model 4 adds control variables related to the child’s home environment and individual characteristics. Model (4) is fully specified, including all controls.\(^{12}\) Lastly, Models 5-7 interact my key independent variable with measures of Hispanic ethnicity, IEP, and gender (female), respectively. Each regression is weighted using sample weights provided by the researchers who created the data set. I report robust standard errors for all coefficients.

\(^{12}\) I also estimated regression models that included a measure of motivation. However, it is possible that my key independent variable is also associated with a student’s motivation level, which could in turn influence the relationship between non-English language use and achievement. More specifically, an ELL student’s motivation level could fluctuate simply as a function of being in a multilingual classroom. The magnitude of my key coefficient was not greatly impacted by the inclusion of this motivation measure in the regressions. Therefore, I do not control for motivation in my main specifications. See the appendix for the results of regression models in which motivation is included as a control.
The Relationship Between Non-English Language Use and Reading Scores

The estimates presented in Table 4 provide evidence of a negative relationship between non-English use in the classroom and reading test scores. The coefficient from Model 1 shows that there is a relatively large (the test scores are on a scale from 1-100), negative, and significant raw correlation between reading scores and non-English language use in the classroom. However, as controls are added in Models 2-4, the key coefficient becomes smaller in magnitude (less negative). After I add teacher-related controls, the coefficient drops 0.3 points in magnitude. The magnitude of this relationship then drops by nearly three points when school-related factors are added. The coefficient is further reduced by another 1.2 points when home and individual factors are included. Ultimately, the key coefficient in the full model (Model 4) still suggests a negative and significant relationship, but it is much smaller in magnitude than the raw correlation. The results of my full model suggest that being in a classroom where a non-English language is spoken is associated with a test score reduction of 2.8 points on average, compared to being in an English-only classroom.

In Models 5-7, I include interaction terms in order to allow me to perform subgroup analysis. I specifically explore whether my key relationship differs for Hispanic students, students with IEPs, and females. The results for Model 5 show that the key coefficient for non-Hispanic students is -3.2 and for Hispanic students it is -2.7 (-3.2+0.5). Model 6 produces a key coefficient of about -3.1 for non-IEP students and -0.9 (-3.1+2.2) for students with IEPs. Lastly, they key coefficient for males in Model 7 is -3.2 and for females is -2.5 (-3.2+0.7). Each of these relationships is statistically significant. In other words, there is a significant and negative relationship between non-English language use in the classroom and reading test scores for all subgroups: Hispanic, non-Hispanic, IEP, non-IEP, male and female. However, it is important to
note that there is variation in the magnitudes of these relationships. More specifically, the variation in results for IEP and non-IEP students is important to note.

**Table 4.** Estimates of the Relationship between Non-English Use and Reading Scores

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Raw Correlation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.845)</td>
<td>(0.833)</td>
<td>(0.862)</td>
<td>(0.815)</td>
<td>(1.371)</td>
<td>(0.847)</td>
<td>(1.118)</td>
</tr>
<tr>
<td><strong>Teacher Factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Master’s Degree</td>
<td>1.425*</td>
<td>-0.268</td>
<td>1.003</td>
<td>1.012</td>
<td>0.980</td>
<td>1.033</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.796)</td>
<td>(0.796)</td>
<td>(0.740)</td>
<td>(0.740)</td>
<td>(0.740)</td>
<td>(0.742)</td>
<td></td>
</tr>
<tr>
<td>Years of Experience</td>
<td>-0.329</td>
<td>-0.030</td>
<td>-0.0327</td>
<td>-0.0324</td>
<td>-0.0323</td>
<td>-0.0327</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td>(0.043)</td>
<td>(0.0395)</td>
<td>(0.0395)</td>
<td>(0.0395)</td>
<td>(0.0395)</td>
<td></td>
</tr>
<tr>
<td>Closeness</td>
<td>3.137***</td>
<td>2.828***</td>
<td>2.248***</td>
<td>2.256***</td>
<td>2.230***</td>
<td>2.260***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.526)</td>
<td>(0.514)</td>
<td>(0.502)</td>
<td>(0.505)</td>
<td>(0.504)</td>
<td>(0.503)</td>
<td></td>
</tr>
<tr>
<td><strong>School Factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% District Poverty</td>
<td>-0.115***</td>
<td>0.0157</td>
<td>0.0157</td>
<td>0.0149</td>
<td>0.0164</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.0353)</td>
<td>(0.0353)</td>
<td>(0.0353)</td>
<td>(0.0355)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% ELL in School</td>
<td>-0.118***</td>
<td>-0.0659***</td>
<td>-0.0659***</td>
<td>-0.0657***</td>
<td>-0.0664***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.0192)</td>
<td>(0.0192)</td>
<td>(0.0192)</td>
<td>(0.0192)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Home and Individual Factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent high school education</td>
<td>3.677***</td>
<td>3.694***</td>
<td>3.720***</td>
<td>3.658***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.871)</td>
<td>(0.871)</td>
<td>(0.867)</td>
<td>(0.875)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lives with both parents</td>
<td>1.161</td>
<td>1.158</td>
<td>1.134</td>
<td>1.163</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.838)</td>
<td>(0.838)</td>
<td>(0.839)</td>
<td>(0.838)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.862)</td>
<td>(0.862)</td>
<td>(0.863)</td>
<td>(0.862)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food Secure</td>
<td>1.542</td>
<td>1.537</td>
<td>1.521</td>
<td>1.548</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.003)</td>
<td>(1.004)</td>
<td>(1.006)</td>
<td>(1.003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.413)</td>
<td>(1.417)</td>
<td>(1.937)</td>
<td>(1.414)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>-5.556***</td>
<td>-5.654***</td>
<td>-5.509***</td>
<td>-5.557***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.035)</td>
<td>(1.115)</td>
<td>(1.033)</td>
<td>(1.036)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4. (cont.)

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1) Raw Correlation</th>
<th>(2) Teacher Factors</th>
<th>(3) School Factors</th>
<th>(4) Home and Individual Factors</th>
<th>(5) Interaction: Hispanic</th>
<th>(6) Interaction: IEP</th>
<th>(7) Interaction: Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black Non-Hispanic</td>
<td>-2.084</td>
<td>-2.137</td>
<td>-2.016</td>
<td>-2.119</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.698)</td>
<td>(2.712)</td>
<td>(2.697)</td>
<td>(2.705)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White Non-Hispanic</td>
<td>0.178</td>
<td>0.164</td>
<td>0.199</td>
<td>0.199</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.194)</td>
<td>(1.196)</td>
<td>(1.195)</td>
<td>(1.196)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.983</td>
<td>0.982</td>
<td>0.987</td>
<td>0.752</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.710)</td>
<td>(0.710)</td>
<td>(0.710)</td>
<td>(0.851)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Interaction Terms

| Non-English Use X Hispanic       | 0.518                 |                     |                   |                                |                          |                     |                        |
|                                  | (1.610)               |                     |                   |                                |                          |                     |                        |
| Non-English Use X IEP            | 2.160                 |                     |                   |                                |                          |                     |                        |
|                                  | (2.751)               |                     |                   |                                |                          |                     |                        |
| Non-English Use X Female         | 0.746                 |                     |                   |                                |                          |                     |                        |
|                                  | (1.500)               |                     |                   |                                |                          |                     |                        |
| Constant                         | 69.57***              | 56.18***            | 64.04***          | 61.89***                       | 61.91***                 | 62.02***            | 61.94***               |
|                                  | (0.483)               | (2.300)             | (2.366)           | (2.740)                        | (2.743)                  | (2.761)             | (2.749)                |

F-Statistics and P-Values

| Non-English Use X Hispanic       | 6.59                  |                     |                   |                                |                          |                     |                        |
|                                  | (0.0014)              |                     |                   |                                |                          |                     |                        |
| Non-English Use X IEP            | 6.52                  |                     |                   |                                |                          |                     |                        |
|                                  | (0.0015)              |                     |                   |                                |                          |                     |                        |
| Non-English Use X Female         | 6.01                  |                     |                   |                                |                          |                     |                        |
|                                  | (0.0025)              |                     |                   |                                |                          |                     |                        |

Observations                      | 1,405                 | 1,405               | 1,405             | 1,405                          | 1,405                    | 1,405               | 1,405                  |
R-squared                         | 0.061                 | 0.092               | 0.152             | 0.390                          | 0.291                    | 0.291               | 0.291                  |

Robust standard errors in parentheses, p-values in parenthesis for F-Tests

*** p<0.01, ** p<0.05, * p<0.1
The models presented in Table 5 suggest a somewhat different relationship between non-English language use in the classroom and math achievement. While the coefficients for non-English use in my models for reading are persistently significant, the same coefficients in my models for math lose significance when a full set of controls is included. As is the case with the results presented in Table 4, these results suggest a negative relationship between non-English use in the classroom and student achievement—although, in this case, that relationship is small in magnitude and insignificant.

The raw correlation presented in Model 1 is smaller (less negative) than for reading but is still statistically significant. Adding teacher factors in Model 2 very slightly reduces the magnitude of the coefficient on non-English use by 0.2 points, and the coefficient remains highly significant. The inclusion of school factors shrinks the coefficient on non-English use by 2.6 points and dampens its significance level. In Model 4, the key coefficient becomes much smaller and insignificant when home and individual factors are added. Ultimately, the key coefficient is insignificant and is reduced by nearly four points compared to the coefficient in Model 1. The results of these models suggest that, once a full set of controls is included in the regression, there is little to no association between whether an ELL student is in a multilingual classroom and his or her performance in math.

I also explore the relationship between non-English use in the classroom and math achievement among the same subgroups used in my reading achievement models. Model 5 shows that the key coefficient for non-Hispanic students is -0.7 and is -3.2 (-0.7+2.5) for Hispanic students. For non-IEP students, the key coefficient is -1.2 and for students with IEP it is about 1.7 (-1.2+3.0). Lastly, results from Model 7 show that the coefficient is -2.2 for males and...
0.5 (-2.2+2.7) for females. Unlike the relationships in my reading models, most of these are insignificant. The relationship for males is marginally significant and negative, although the magnitude of the coefficient is small.

Table 5. Estimates of the Relationship between Non-English Use and Math Scores

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1) Raw Correlation</th>
<th>(2) Teacher Factors</th>
<th>(3) School Factors</th>
<th>(4) Home and Individual Factors</th>
<th>(5) Interaction: Hispanic</th>
<th>(6) Interaction: IEP</th>
<th>(7) Interaction: Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-English Use</td>
<td>-4.676*** (0.844)</td>
<td>-4.448*** (0.832)</td>
<td>-1.857*** (0.849)</td>
<td>-0.916 (0.800)</td>
<td>-0.701 (1.616)</td>
<td>-1.241 (0.818)</td>
<td>-2.213** (1.119)</td>
</tr>
<tr>
<td>Master’s Degree</td>
<td>1.523* (0.796)</td>
<td>-0.159 (0.783)</td>
<td>1.254* (0.738)</td>
<td>1.291* (0.734)</td>
<td>1.223* (0.738)</td>
<td>1.363* (0.735)</td>
<td></td>
</tr>
<tr>
<td>Years of Experience</td>
<td>-0.048 (0.041)</td>
<td>-0.050 (0.040)</td>
<td>-0.0631* (0.0372)</td>
<td>-0.0641* (0.0373)</td>
<td>-0.0625* (0.0372)</td>
<td>-0.0631* (0.0373)</td>
<td></td>
</tr>
<tr>
<td>Closeness</td>
<td>2.407*** (0.517)</td>
<td>2.15*** (0.510)</td>
<td>2.008*** (0.488)</td>
<td>1.989*** (0.487)</td>
<td>1.983*** (0.490)</td>
<td>2.050*** (0.486)</td>
<td></td>
</tr>
<tr>
<td>% District Poverty</td>
<td>-0.133*** (0.034)</td>
<td>0.0127 (0.0341)</td>
<td>0.0138 (0.0341)</td>
<td>0.0116 (0.0342)</td>
<td>0.0152 (0.0341)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% ELL in School</td>
<td>-0.084*** (0.019)</td>
<td>-0.0279 (0.0189)</td>
<td>-0.0292 (0.0190)</td>
<td>-0.0276 (0.0190)</td>
<td>-0.0297 (0.0189)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent high school education</td>
<td>2.998*** (0.851)</td>
<td>2.908*** (0.850)</td>
<td>3.055*** (0.846)</td>
<td>2.927*** (0.854)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lives with both parents</td>
<td>1.858** (0.887)</td>
<td>1.840** (0.881)</td>
<td>1.821** (0.887)</td>
<td>1.867** (0.884)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household Poverty</td>
<td>-3.013*** (0.845)</td>
<td>-2.942*** (0.840)</td>
<td>-2.964*** (0.844)</td>
<td>-3.013*** (0.844)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food Secure</td>
<td>0.811 (1.013)</td>
<td>0.823 (1.012)</td>
<td>0.781 (1.019)</td>
<td>0.831 (1.007)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IEP</td>
<td>-8.297*** (1.473)</td>
<td>-9.621*** (1.829)</td>
<td>-9.473*** (1.835)</td>
<td>-8.280*** (1.471)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>-7.184*** (1.835)</td>
<td>-6.671*** (1.829)</td>
<td>-7.119*** (1.835)</td>
<td>-7.188*** (1.835)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VARIABLES</td>
<td>Raw Correlation</td>
<td>Teacher Factors</td>
<td>School Factors</td>
<td>Home and Individual Factors</td>
<td>Interaction: Hispanic</td>
<td>Interaction: IEP</td>
<td>Interaction: Female</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>-----------------------</td>
<td>----------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Black Non-Hispanic</td>
<td>-4.084 (2.666)</td>
<td>-3.867 (2.675)</td>
<td>-3.989 (2.669)</td>
<td>-4.209 (2.667)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White Non-Hispanic</td>
<td>0.143 (1.278)</td>
<td>0.317 (1.275)</td>
<td>0.173 (1.276)</td>
<td>0.220 (1.280)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-2.107*** (0.708)</td>
<td>-2.102*** (0.708)</td>
<td>-2.101*** (0.794)</td>
<td>-2.945*** (0.841)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Interaction Terms
- Non-English Use X Hispanic: -2.452 (1.609)
- Non-English Use X IEP: 2.967 (3.037)
- Non-English Use X Female: 2.707* (1.513)

Constant: 61.859*** (0.472)
51.699*** (2.278)
58.99*** (2.384)
57.53*** (2.779)
57.46*** (2.771)
57.71*** (2.794)
57.72*** (2.785)

F-Statistics and P-Values
- Non-English Use X Hispanic: 1.51 (0.2201)
- Non-English Use X IEP: 1.33 (0.2639)
- Non-English Use X Female: 2.11 (0.1213)

Observations: 1,405
R-squared: 0.028

Robust standard errors in parentheses, p-values in parenthesis for F-Tests
*** p<0.01, ** p<0.05, * p<0.1
DISCUSSION

My results suggest that the relationship between non-English language use in the classroom and academic achievement differs for reading and math outcomes. The relationship between non-English language use and reading test scores is negative and statistically significant, but small in magnitude. This finding is inconsistent with previous research, where student participation in a bilingual classroom has a positive relationship with language development for young children (Barnett et. al., 2007; Duran et. al., 2010). The relationship between non-English language use and reading achievement is also negative and statistically significant for several subgroups, including non-Hispanic, Hispanic, non-IEP, IEP, male, and female students. This relationship is strongest for non-Hispanic students and weakest for non-IEP students. These subgroup findings make interesting contributions to the literature, as previous studies do not explore these subgroups in depth.

The relationship between non-English language use and math test scores is negative and small, but is not statistically significant. This finding is aligned research by Matsudaira, whose results suggest a negligible relationship between bilingual instruction and math achievement (Matsudaira, 2005). The relationship between non-English language use and math achievement varies somewhat among subgroups but it is not statistically significant for any group except males. This finding also expands upon existing studies by contributing results for subgroups.

Despite including controls for teacher-related, school-related, and individual-and-home-related factors, my analysis is still subject to omitted variable bias due to the limitations of my data and potential measurement error. More data on teacher-related factors, especially an indicator for bilingual (or multilingual) certification, would help account for variability in teachers’ bilingual instruction ability. Presumably, teachers with such certification would be more likely to teach in classrooms where non-English languages are used. Moreover,
demonstrated capacity in knowledge of the linguistic and cultural factors related to language can help bilingual teachers be more effective the classrooms with ELL students (Fillmore & Snow, 2000). If possession of bilingual certification is positively related to the use of non-English languages in the classroom and is also positively related to a teacher’s efficacy, excluding it from my model may have exerted an upward bias on my estimates. In other words, the key coefficient is understated.

Data regarding students’ English language proficiency levels at the time of assessment would also address omitted variable bias in my regressions. ELL students are a diverse group of students in terms of their native languages and in terms of how well they speak and understand English. Generally, ELL students who have higher English proficiency also perform better on academic assessments than students with lower English proficiency (Haas et. al., 2016). Thus, proficiency level would be presumably negatively correlated with inclusion in a non-English use classroom, but positively related to test scores. Consequently, the omission of a measure of proficiency from my regressions may have exerted a downward bias on my key coefficients.

Additionally, measurement error may have affected my analysis. This is largely due to the methods used to collect the data included in my sample. For example, many data points collected via parent, teacher, and school administrator surveys and interviews may not be entirely accurate or may be incomplete. The same is likely true for students’ reading and math test score data. These errors are mostly likely attributable to transcription errors, which is to say that they are probably random in nature. Consequently, my key coefficients may be affected by

---

13 Many schools with ELL populations have testing requirements to monitor students’ language development and to determine when students attain English language proficiency (National Research Council, 2011).
attenuation bias, causing them to be closer to zero than they would be in the absence of measurement error.

Overall, my analyses provide modest guidance for policy decisions regarding ELL students. Given that my results suggest that there is a negligible relationship between non-English language use and math achievement, it may not be wise to overprescribe how school districts should approach ELLs’ math education via federal education policy. On a local level, resources may be better spent on other aspects of ELL students’ education, including tutors or translation services. While I find that there is a statistically significant relationship between non-English language use and reading achievement, the magnitude of this relationship is relatively small. Estimates suggest that the achievement gap between ELLs and their native English-speaking peers is between 25 and 36 points on standardized assessments (US Office of English Language Acquisition, 2016). My coefficient for the reading model is -3.2, which is eclipsed by the magnitude of this broader achievement gap. On the other hand, the variation in my estimates of the magnitude of this relationship across outcomes and across different subgroups of students highlights the complexity of ELL students’ classroom experiences. Thus, one-size-fits-all guidance may not be appropriate for ELL students.

Future research that accounts for these limitations of my study would help to further elucidate the relationship between non-English language use in the classroom and academic achievement. Additionally, since academic outcomes can change dramatically over time, longitudinal research that monitors the relationship between instructional practices and ELL student achievement would provide better guidance for policymakers. Research that explores the relationship between non-English language use and preservation of students’ cultural identities would also contribute to the debate surrounding instructional approaches for ELLs, as test scores
are only one dimension of a students’ experiences. The existing literature does not currently explore this relationship using rigorous statistical methods.

Given the diversity of languages spoken by new immigrant and refugee children, it is important that future studies account for the specific languages used in the classroom. My data do not indicate whether a student’s native language matches the language(s) used in the classroom. As previously discussed, the classroom is much easier to navigate for many ELL students if the teacher uses the student’s native language versus a language that the student is unfamiliar with (Allison & Rehm, 2007). Since 30 percent of my sample speaks a language other than Spanish—which is also the dominant bilingual instructional language nationwide—it is possible that many students were in classrooms where their native languages were not used (Boyle et. al, 2015). These students might have performed more poorly because no resources were offered in a familiar language or cultural context. Ideally, future work in this area should therefore account for whether or not the language spoken in the classroom is the same as the language that the student speaks at home. More broadly, future policy research on ELL students and academic achievement should be to seek to provide evidence regarding the most effective instruction and support possible for ELL students as they navigate multiple languages and cultures.
**Appendix: Full Models Including Motivation Variable**

Table 6. Full Models Including Motivation Variable

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key Independent Variable</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-English Use</td>
<td>-2.520***</td>
<td>-0.637</td>
</tr>
<tr>
<td></td>
<td>(0.759)</td>
<td>(0.744)</td>
</tr>
<tr>
<td><strong>Teacher Factors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Master’s Degree</td>
<td>0.334</td>
<td>0.624</td>
</tr>
<tr>
<td></td>
<td>(0.709)</td>
<td>(0.704)</td>
</tr>
<tr>
<td>Years of Experience</td>
<td>-0.00401</td>
<td>-0.0360</td>
</tr>
<tr>
<td></td>
<td>(0.0382)</td>
<td>(0.0353)</td>
</tr>
<tr>
<td>Closeness</td>
<td>1.327***</td>
<td>1.133**</td>
</tr>
<tr>
<td></td>
<td>(0.476)</td>
<td>(0.456)</td>
</tr>
<tr>
<td><strong>School Factors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% District Poverty</td>
<td>0.0211</td>
<td>0.0178</td>
</tr>
<tr>
<td></td>
<td>(0.0327)</td>
<td>(0.0322)</td>
</tr>
<tr>
<td>% School ELL</td>
<td>-0.0808***</td>
<td>-0.0420**</td>
</tr>
<tr>
<td></td>
<td>(0.0182)</td>
<td>(0.0183)</td>
</tr>
<tr>
<td><strong>Home and Individual Factors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent high school education</td>
<td>2.182***</td>
<td>1.587*</td>
</tr>
<tr>
<td></td>
<td>(0.822)</td>
<td>(0.809)</td>
</tr>
<tr>
<td>Lives with both parents</td>
<td>0.201</td>
<td>0.953</td>
</tr>
<tr>
<td></td>
<td>(0.801)</td>
<td>(0.848)</td>
</tr>
<tr>
<td>Household poverty</td>
<td>-2.804***</td>
<td>-2.608***</td>
</tr>
<tr>
<td></td>
<td>(0.795)</td>
<td>(0.786)</td>
</tr>
<tr>
<td>Food Secure</td>
<td>0.365</td>
<td>-0.301</td>
</tr>
<tr>
<td></td>
<td>(0.947)</td>
<td>(0.937)</td>
</tr>
<tr>
<td>Motivation</td>
<td>5.072***</td>
<td>-6.654***</td>
</tr>
<tr>
<td></td>
<td>(0.369)</td>
<td>(1.432)</td>
</tr>
<tr>
<td>Special Education</td>
<td>-8.195***</td>
<td>4.804***</td>
</tr>
<tr>
<td></td>
<td>(1.338)</td>
<td>(0.373)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-4.966***</td>
<td>-6.620***</td>
</tr>
<tr>
<td></td>
<td>(0.961)</td>
<td>(0.994)</td>
</tr>
<tr>
<td>Black Non-Hispanic</td>
<td>0.0642</td>
<td>-2.049</td>
</tr>
<tr>
<td></td>
<td>(2.645)</td>
<td>(2.390)</td>
</tr>
<tr>
<td>VARIABLES</td>
<td>(1) Reading</td>
<td>(2) Math</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>White Non-Hispanic</td>
<td>-0.156</td>
<td>-0.173</td>
</tr>
<tr>
<td></td>
<td>(1.136)</td>
<td>(1.206)</td>
</tr>
<tr>
<td>Female</td>
<td>0.711</td>
<td>-2.367***</td>
</tr>
<tr>
<td></td>
<td>(0.663)</td>
<td>(0.665)</td>
</tr>
<tr>
<td>Constant</td>
<td>49.30***</td>
<td>45.62***</td>
</tr>
<tr>
<td></td>
<td>(2.756)</td>
<td>(2.781)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,405</td>
<td>1,405</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.390</td>
<td>0.347</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses, p-values in parenthesis for F-Tests

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


