Summer 8-1-2020

Three Essays on the Academic Outcomes of Disadvantaged Students and their Peers

Maria Camila Morales Leon
Georgia State University

ABSTRACT

THREE ESSAYS ON THE ACADEMIC OUTCOMES OF DISADVANTAGED STUDENTS AND THEIR PEERS

BY

MARIA CAMILA MORALES LEON

August 2020
Committee Chair: Dr. Tim Sass
Major Department: Economics

Using longitudinal administrative data and quasi-experimental methods, this dissertation evaluates the impact of changing demographics in schools and education policies designed to enhance the academic achievement of English Learners.

Chapter 1 contributes to an emerging literature on the externalities of refugee integration by providing evidence on how this population affects the academic performance and behavior of incumbent students. Leveraging variation in the share of refugees within schools and across grades, I find that increasing the share of grade-level refugees by 1 percentage point leads to a 0.01 standard deviation increase in average math test scores. I find suggestive evidence that the positive spillovers in math achievement are driven by changes in classroom resources and access to academic support programs. While I find no effect on average English Language Arts (ELA) test scores, using nonlinear-in-means specifications I find evidence of negative spillovers in ELA

performance among low-achieving students and positive spillovers among high-achieving students.

Chapter 2 estimates the impact of a temporary intensive English program aimed at English Learners with very low English proficiency. Access to the program is based on a maximum score on a standardized English proficiency screening assessment and grade level at screening, which I leverage to employ difference-in-differences and regression discontinuity approaches in my analysis. I estimate the impact of program access and participation on ELA and math test scores in the short term, relative to receiving traditional English as a Second Language support. On average, students who are eligible for the program have lower ELA test scores one year after program eligibility. However, the impact is large and positive among the subsample of refugee students. Results on the impact on program enrollment also show that students who participate have lower ELA achievement. I also find lower math test scores among program participants, relative to English Learners who receive traditional English as a Second Language support.

Chapter 3 presents new evidence on the impact of Dual Language Immersion programs on student academic outcomes. Leveraging enrollment lotteries from five oversubscribed DLI schools, I estimate intent-to-treat and local-average-treatment effects of bilingual education on English Language Arts, reading, and math test scores. On average, I find no difference in reading and ELA achievement between students with access to DLIs and those enrolled in traditional public schools. However, I find weak evidence of lower math achievement among lottery winners. Results vary by students' initial EL status. Specifically, I find that native English speakers with access to DLI programs have higher reading and ELA achievement, relative to DLI lottery losers.

THREE ESSAYS ON THE ACADEMIC OUTCOMES OF DISADVANTAGED STUDENTS AND THEIR PEERS

BY

MARIA CAMILA MORALES LEON

A Dissertation Submitted in Partial Fulfilment
of the Requirements for the Degree
of
Doctor of Philosophy
in the
Andrew Young School of Policy Studies
of
Georgia State University

GEORGIA STATE UNIVERSITY
2020

ACCEPTANCE

This dissertation was prepared under the direction of Maria Camila Morales Leon's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Andrew Young School of Policy Studies of Georgia State University.

Dissertation Chair:	Dr. Tim Sass
Committee:	Dr. Daniel Kreisman Dr. Thomas Mroz Dr. Ross Rubenstein

Electronic Version Approved

Sally Wallace, Dean
Andrew Young School of Policy Studies
Georgia State University
August 2020

DEDICATION

First, I dedicate this dissertation to my family. To my parents, Julio Cesar and Ana Maria, whose invaluable support, example of hard work, dedication, and constant encouragement made this journey one worth fighting through. Los amo con toda mi alma, y todo mi esfuerzo es un agradecimiento constante por todos sus sacrificios. Esto es por ustedes, gracias por tanto. To my husband, Memusi, who has been my rock throughout these years, always believing in me, a constant presence of peace. You saw me through nights of tears and anxiety, and stood proudly by me through every milestone. This is for our family and for our future children. Lastly, I dedicate this dissertation to all immigrant and refugee children. To all students who struggle in school because of language barriers and feel isolated in a world that is different to the one they were used to. You are valued, you matter, and your lives are the future of this country.

ACKNOWLEDGEMENTS

I would like to thank all my mentors, advisors, colleagues, and friends who have invested in my growth as a young scholar and have supported my journey. To my advisor, Tim Sass, thank you for your invaluable guidance, generous support, and for helping me develop my skills as a researcher. From the beginning you nudged me to look at the refugee students in Georgia and that has opened a fruitful line of work, thank you. To Dan Kreisman, thank you for believing in me and pushing me to achieve my best. I sincerely appreciate your commitment to my success and your generous advice. To my committee members, Thomas Mroz, and Ross Rubenstein, thank you for your continued support and advice. I could not have asked for a better dissertation committee; the quality of my work reflects your guidance and constructive feedback.

I am also grateful for the financial support of the Georgia Policy Labs, and to the Metro Atlanta Policy Lab for Education for providing access to the data that I use in this dissertation. My appreciation also to all the school districts that agreed to participate in each of these projects. Thank you for seeing the actionable value in the research.

To everyone in the Economics Department at Georgia State, past and present. My schooling career began as an undergraduate at GSU, and many of you have taught me, inspired me, and ultimately shaped me into the scholar that I am today. Special thanks to Felix Rioja who encouraged me to apply to the AEA Summer Training Program, to Paul Kagundu who introduced me to economic research, and to Rachana Bhatt who guided my first research paper.

My gratitude also goes to my mentors Luisa Blanco, Catalina Amuedo-Dorantes, Marie Mora, Sarah Jacobson, Kalena Cortes, and many others. Thank you for helping me navigate through challenges and inspiring me to believe that women belong in the economics profession. I would also like to thank the constant support of everyone in the AEA Mentoring Program, the

AEA Summer Training Program, and the American Society of Hispanic Economists. I am also grateful for the financial support of the National Science Foundation.

Additionally, I would like to thank my classmates and colleagues at Georgia State and the AEA Summer Training Program at MSU. I would not have made it through this program without your support, encouragement, and friendship. To everyone in my PhD cohort who helped me get through comps, exams, and problem sets, thank you. Special thanks to Kalee Burns, Carycruz Bueno, Tareena Musaddiq, Alexa Prettyman, and Britni Wilcher. Thank you for proofreading my papers, talking through identification strategies, and giving me feedback in presentations. You are excellent researchers and even better friends.

Lastly, and most importantly, I would like to thank God – my source of being and all good things. You have orchestrated my every move and have seen me through the roughest of times, shining your light and wisdom through it all. I do all things through you and this is no exception.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

In recent years, heterogeneity in school demographics has expanded to new immigrant destination states. This dissertation studies the implications of changes in school demographics on the academic achievement of peers and evaluates education policies that affect English Learners.

Chapter 1 studies the impact of attending school with refugee children on the academic outcomes of nonrefugees. At a time when the refugee crisis has reached historic peaks, changes in US refugee resettlement policies have occurred in a vacuum of empirical research. There is a dearth of evidence on the impact of contemporary refugee resettlement on the integration outcomes of refugees and effects on local communities. This chapter contributes to the growing literature on the externalities of refugee resettlement and provides novel evidence within the context of public education.

I leverage temporal variation in the share of refugee students within schools and across grades to estimate the impact of an increase in refugee students on the English Language Arts (ELA) and math test scores of peers. I find that increasing the share of grade-level refugees by 1 percentage point leads to a 0.01 standard deviation increase in average math test scores. I find suggestive evidence that the positive spillovers in math achievement are driven by changes in classroom resources and access to academic support programs. While I find no effect on average ELA test scores, using nonlinear-in-means specifications I find evidence of negative spillovers in ELA performance among low-achieving students and positive spillovers among high-achieving students.

Chapter 2 evaluates the efficacy of an Intensive English program on the academic achievement of English Learners (ELs). ELs represent the fastest-growing student group in the

US and are among the lowest achieving. While there is significant research on the impact of education policies on EL achievement, little is known on how these policies affect ELs with the lowest level of English proficiency. I present new evidence on the impact of a temporary school intervention designed for ELs with very low initial English proficiency, relative to receiving traditional English as a Second Language instruction. I leverage program eligibility criteria to estimate intent-to-treat and treatment on the treated effects using difference-in-differences and regression discontinuity specifications. On average, students who are eligible for the program have lower ELA test scores one year after program eligibility. However, the impact is large and positive among the subsample of refugee students. Results on the impact on program enrollment also show that students who participate have lower ELA achievement. I also find lower math test scores among program participants, relative to English Learners who receive traditional English as a Second Language support.

Chapter 3 studies the impact of access and enrollment in a Dual Language Immersion (DLI) program. DLIs use two languages for instruction – English and a target language – and have seen rapid proliferation in recent years. Despite their growth, rigorous research on the impact of DLIs remains limited in comparison. I present new evidence on the impact of DLIs on student achievement by leveraging randomized access to five oversubscribed DLI programs.

I estimate intent-to-treat and local-average-treatment effects of bilingual education on English Language Arts, reading, and math test scores. On average, I find no difference in reading and ELA achievement between students with access to DLIs and those enrolled in traditional public schools. However, I find weak evidence of lower math achievement among lottery winners. Results vary by students' initial EL status. Specifically, I find that native English speakers with access to DLI programs have higher reading and ELA achievement, relative to students who lose the lottery and enroll in traditional public schools.

The remainder of this dissertation is organized by chapter. Each includes an introduction outlining the relevance of the research question, a literature review and summary of contributions, sections with background information and program details, a description of the data and methods used, and a discussion of the results and conclusion. Appendices are included for additional background information and subsample analyses.

Chapter 1: Do Refugee Students Affect the Academic Achievement of Peers? Evidence from a Large Urban School District

1.1 Introduction

The number of refugees and individuals displaced by conflict is at a historical record high. The office of the United Nations High Commissioner for Refugees (UNHCR) reports that, by 2018, over 70 million people had been forced to flee their country of origin, a notable increase from approximately 16 million in 2005.[1] Historically, the United States has taken in more refugees than any other country;[2] however, recent changes in the refugee admissions ceiling for 2017 led the United States to resettle fewer refugees than the rest of the world for the first time in the since the creation of the U.S. Refugee Resettlement Program in 1980.[3] Further restrictions to refugee resettlement continue to take place.[4] The policy debate on whether to change the number of refugee arrivals is largely driven by the perceived adverse effects on local communities, and costs that refugees may impose on all levels of government and the native population.[5] However, there is sparse credible evidence to ascertain the impact of refugees on local communities. To fill this void, I present evidence on the indirect costs of refugee integration within the context of public schools.

[1] See Figure A1.1 in the appendix for a time trend of the total population displaced by conflict from 1990 to 2018. Displaced persons include refugees, asylees, and internally displaced individuals.
[2] The US has resettled over 3.7 million people since the beginning of the Refugee Resettlement program in 1980. For reference, less than 1% of refugees are offered the resettlement option; it is known as the solution of last resort (UNHCR, 2019). I obtained data on total refugee admissions from the U.S. Department of State Refugee Admissions Program.
[3] See Figure A1.2 in the appendix for a comparison in refugee resettlement flows to the US vs. all other countries from 1990 to 2018.
[4] The Refugee Resettlement Program was suspended for 120 days in 2017 (Fix et al., 2017) and the proposed ceiling for FY2020, 18,000 individuals, is the lowest ever recorded.
[5] Concerns at the local level appear to be on the rise, as reflected in the recent Executive Order #13888 by which state and local governments must consent to resettlements before any refugees arrive in a city (Trump 2019).

In this paper, I estimate the causal effect of a change in the proportion of refugee students in a grade on the academic outcomes of non-refugees. I also present suggestive evidence of three potential mechanisms that may explain the estimated peer effect: teacher responses, changes in class size and classroom resources, and access to academic support programs. I utilize individual-level administrative data of students in grades 3 through 8 who were enrolled in public schools in the district with the highest inflow of refugees in Georgia between 2008 and 2017.[6] Access to these records allow me to observe student outcomes (test scores, disciplinary incidents, and attendance) and rich demographic information. In addition, I obtain data from the district's International Welcome Center with information on students' self-reported refugee status and date of entry to the United States. To my knowledge, this is the first paper to utilize school administrative data that also contain students' immigration-related information.

I contribute to the literature on the impact of refugee integration on local communities and research on the peer effects of immigrant students. I specifically address three gaps. First, prior research on the impact of refugee integration on local communities in the United States has focused on the adult refugee population and the impact on local crime (Amuedo-Dorantes, Bansak, and Pozo 2018; Masterson and Yasenov 2019). I extend this literature by exploring the spillovers of refugee resettlement of children in a context that has not been studied before, namely local public schools. As of 2017, children make up nearly 43 percent of all refugees resettled in the United States (Mossaad, 2019), and while the Federal government funds the majority of resettlement costs, state and local governments are the primary sources of public education funding. Therefore, education is a critical context in which to study the impact of refugee resettlement on host communities.

[6] Georgia ranks among the top ten resettlement states in the US. See Figure A1.3 in the appendix.

Second, most of the literature on refugee integration and the impacts of resettlement relies on a proxy method of identifying refugees from large data sets (Capps et al. 2015; Cortes 2004; Evans and Fitzgerald 2017). In contrast, I use data on students' self-reported refugee status and date of arrival to the United States, which allow me to assess the validity of the proxy method and explore refugee peer effects by length of stay.[7]

Third, while there is a growing literature on the impact of immigrants on the academic achievement of peers, studies find mixed evidence with some showing that an increase in the proportion of immigrant students has a positive or null effect on the educational outcomes of peers (Conger 2015; Figlio and Özek 2019; Hunt 2017; Schwartz and Stiefel 2011) and others finding negative spillovers (Brunello and Rocco 2013; Frattini and Meschi 2017; Gould, Lavy, and Daniele Paserman 2009; Jensen and Rasmussen 2011). Moreover, the results from this literature may not translate to the unique context of refugees (Cortes 2004; Dustmann et al. 2017).[8] I contribute to this literature by studying the peer effects of a unique group of foreign-born students using a contemporary, diverse, and representative sample of refugees.

Using variation in the within-school share of refugee students across grades, I find no evidence of negative spillovers in average academic achievement. Rather, results from my preferred specification show that increasing the grade-level share of refugees by one percentage point results in a 0.01 standard deviation increase in average math achievement. The magnitude of the spillover in math achievement corresponds to roughly one-tenth of the impact of a highly effective teacher (Aaronson et al., 2007; Rivkin et al., 2005; Rockoff, 2004), and it is comparable to the peer effects associated with having more girls as peers (Hoxby, 2000). Additional analyses

[7] See Appendix B1 for a comparison between the self-reported refugee data and the proxy approach.
[8] The UN defines refugees as persons who have been forced to flee their home country because of persecution, war, or violence. In contrast, immigrants are commonly considered to be persons who voluntarily leave their home country in search of better opportunities, primarily economic opportunities (Cortes 2004).

exploring differences by refugees' English Learner classification and length of stay in the country provide suggestive evidence that changes in classroom resources, in the form of additional teachers or smaller class sizes, and access to academic support programs available to refugees may be driving the positive spillovers in math performance.

While I find no effects on average ELA test scores, results from nonlinear specifications by non-refugees' initial achievement levels show negative spillovers in ELA performance among low-achieving students and positive spillovers among high-achieving students. These results provide suggestive evidence of possible competition over teacher time and resources between low-achieving peers and refugees.

The remainder of the paper is organized as follows. Section 1.2 presents background information on the institutional characteristics of the resettlement process and describes the refugee population in Georgia. Section 1.3 discusses the relevant literature and highlights the contributions. Section 1.4 describes the data and provides descriptive statistics. Section 1.5 outlines the empirical approach. Section 1.6 discusses the results. Section 1.7 concludes.

1.2 Background

1.2.1 The Refugee Resettlement Process and Integration Policies

As of 2018, over 25 million persons were registered as refugees, corresponding to over one-third of all individuals forcefully displaced (UNHCR 2019).[9] Resettlement is the least likely permanent solution available to refugees upon their displacement and verification of refugee status; less than one percent of refugees are offered the resettlement option.[10] Refugees who are

[9] Refugees are persons who have been forced to flee their home country due to persecution, war, or violence for reasons of race, religion, nationality, political opinion, or membership to a particular social group (UNHCR, 2019).
[10] Other solutions include repatriation to their home country once the fear of persecution has subsided – nearly 600,000 refugees returned to their home country in 2018 (UNHCR 2019) – and integration in the country to which the refugee first fled – a total of 62,600 refugee naturalizations were reported in 2018 (UNHCR 2019).

recommended for resettlement in the United States go through a lengthy screening process that can take 18-24 months and includes health and security checks (Capps and Fix 2015). Once the Department of Homeland Security (DHS) clears applicants for admission, one of nine national nonprofit resettlement agencies becomes responsible for their case, including the choice of final destination state, initial reception, and orientation.[11,12] Refugees are resettled throughout the country, with more than half concentrated in the top ten receiving states.[13]

Upon arrival to the United States, refugees have access to an array of services to facilitate their integration and economic self-sufficiency. They are eligible to enroll in public assistance programs such as Temporary Assistance for Needy Families (TANF), Medicaid, and the Supplemental Nutrition Assistance Program (SNAP).[14] In addition, they receive targeted services such as employment assistance and English as a Second Language instruction to help adult refugees overcome barriers to employment and social integration. Federal grants such as the Targeted Assistance Grant (TAG) and the Refugee Social Service Program cover most of the costs of these services. In 2016, each of these grants was estimated to cost 48 million and 1 million, respectively (Office of Refugee Resettlement 2018).

In light of the high direct costs of resettlement, recent work points to positive economic integration outcomes in the long-run. Bollinger and Hagstrom (2008) show that while the use of public assistance among refugees is high upon arrival, it declines rapidly overtime. Moreover,

---

[11] As of 2019, the nine agencies include: Church World Service, Ethiopian Community Development Council, Episcopal Migration Ministries, Hebrew Immigrant Aid Society, International Rescue Committee, US Committee for Refugees and Immigrants, Lutheran Immigration and Refugee Services, United States Conference of Catholic Bishops, and World Relief Corporation.

[12] Refugees with family ties in the country are commonly placed near their relatives. However, the destination decision for "free cases" is based on the characteristics of the refugee (e.g., age and health) and the availability of local resources (e.g., job and education opportunities) (Fix et al., 2017).

[13] See Figure A1.3 in the appendix for a chart of the top resettlement destination states in the US.

[14] Refugees initially qualify for most means-tested Federal public assistance programs up to 5 or 7 years. Thereafter, eligibility depends on state rules (Fix et al., 2017).

there is evidence that during the first twenty years of resettlement, adult refugees pay more in taxes than they receive in benefits (Evans and Fitzgerald 2017). Relative to economic migrants, refugees experience faster economic integration and human capital accumulation post-migration (Cortes, 2004).

The integration of refugee children is also an important goal of the refugee resettlement program.[15] Schools play an essential role, not only as a place to acquire academic and language skills, but also as a place to connect with the broader community, promote civic engagement, and integrate into the local culture. The Office of Refugee Resettlement (ORR) runs a separate grant specifically designed to assist regions with a high concentration of children in public schools. The Refugee School Impact Grant (RSIG) funds activities that promote the academic achievement of refugees and facilitate social integration. Some programs include after-school tutoring, summer activities, and interpreter services. States with more than 50 school-aged arrivals during the two years preceding a funding request qualify for the grant. In 2016, ORR awarded 38 grants totaling $17 million (Office of Refugee Resettlement 2018).

I provide evidence on the peer effects of refugee students in light of access to a diverse range of academic services. Thus, part of the peer effects can be driven by spillovers of a change in funding for support programs, which would not be available in the absence of refugee students. This is a possible mediating factor, and results may differ in the absence of such services. However, given that educational support services for refugee children is a common feature of traditional refugee-hosting communities – as evidenced by the number of RSIG awardees – I do not view this as a major threat to the external validity of my results.

[15] https://www.acf.hhs.gov/orr/programs/school-impact

1.2.2 Refugees in Georgia

I focus on the refugee population in Georgia, a state that has taken in over 37,000 refugees from 2002 to 2018 and ranks among the top ten resettlement states in the country.[16] There has been substantial variation in the annual arrival of refugees over time, and this trend closely mirrors the national pattern in arrivals. As seen in Figure 1.1, there was an overall steady increase in refugee arrivals from 2002 up to 2016, at which point there was a sharp decline in refugee resettlement flows. In fact, the total refugee arrivals to Georgia in 2018, roughly 900, was the lowest in over a decade. The overwhelming majority of refugees are resettled in counties within the Atlanta metropolitan area,[17] with the top resettlement county comprising over 80 percent of the total refugee population in the state.[18] For this analysis, I use data from the school district that serves students living in the top refugee-resettlement county in Georgia.

Generally, the refugee population in Georgia is comparable to that across the US in both age and countries of origin. For example, in 2014, refugees ages 18 and under made up roughly 35 percent of all arrivals to the US and 40 percent of arrivals to Georgia (Mossaad 2016). Similarly, as shown in Figure 1.2, over one-third of Georgia refugees are from Myanmar (Burma), followed by the Democratic Republic of Congo, and Bhutan. All of these countries are among the top five countries of origin for refugee arrivals to the U.S. in 2015 (Mossaad 2019)

In sum, Georgia has a high share of refugee arrivals over time, with heterogeneity in refugee composition by country of origin, and is closely representative of the refugee population

---

[16] See Figure A1.3 in the appendix for a chart of the top resettlement destination states in the US.
[17] There is also a growing refugee population in the Savannah-Metro area.
[18] The refugee arrival trends for the top resettlement county follows the state trend closely, with some exceptions in the latter years. As seen in Figure A1.4 in the appendix, resettlements to the City of Atlanta exceeded those in the top resettlement county starting from 2014.

in the United States. Thus, I conclude that the context in Georgia is generalizable, and results here may be useful for other traditional refugee resettlement states.

1.3 Relevant Literature and Contribution

This paper builds on two related strands of literature. First, studies on the impact of immigrant students on the educational outcomes of peers.[19] Generally, this literature finds mixed evidence on the effect of grade-level concentration of foreign-born students on native students' test scores, high school graduation, and college enrollment. Using US Census data, Betts (1998) finds that as states' immigrant population share increases, the likelihood of graduating high school decreases among native-born Hispanic and Black students. Gould, Lavy, and Paserman (2009) use variation from a large inflow of immigrants to Israel in the 1990s and find that an increase in the share of immigrant students in 5th grade leads to a decrease in the passing rate of the high school matriculation exam. Cross-country evidence using PISA test score data also suggests a negative effect of immigrant share on natives' academic performance at age 15 (Brunello and Rocco 2013; Jensen and Rasmussen 2011).

Studies show that negative impacts of immigrants on native student outcomes are nonlinear in peer ability, classroom composition, and immigrant age of arrival. Frattini and Meschi (2017) find that adverse effects in test scores are larger among lower-achieving students and concentrated in classrooms with a high share of foreign-born students, and Bossavie (2018) finds that an increase in recently arrived immigrants is associated with small negative effects on language test scores. Lastly, recent evidence suggests that linguistic distance and ethnic diversity

[19] There is a related strand of literature that studies the impact of immigrant concentration on native flight. Research finds that an increase in immigrants leads to an increase in private school enrollment among native-born students (Betts and Fairlie 2003; Cascio and Lewis 2012; Murray 2016; Tumen 2019).

among peers can explain part of these negative spillovers (Ballatore et al., 2018; Frattini and Meschi 2017).

Alternatively, other studies find positive or precise null immigrant peer effects. Using administrative data on New York City primary school students, Schwartz and Stiefel (2011) find no relationship between attending school with foreign-born students and test scores in reading and math. Figlio and Özek (2019) present the first evidence of the effect of a particularly vulnerable group of immigrants on the academic outcomes of peers. Using the 2010 earthquake in Haiti as an exogenous shock, the authors study whether the inflow of evacuees to Florida public schools affected the academic achievement of incumbent students. Results from their study show that variation in the proportion of grade-level evacuees had no adverse effect on the academic outcomes of peers, whether immediately after the earthquake or two years later.[20] These results align with previous evidence from Conger (2015), which finds no relationship between the share of foreign-born students in Florida and high school academic performance, irrespective of immigrant students' English Learner status. Lastly, using US Census data, studies find a net positive effect of immigration on high school graduation, especially among Black students (Hunt 2017), and an increase in college enrollment in states with a high number of unskilled immigrants relative to skilled immigrants (Jackson 2015).

[20] While Figlio and Özek (2019) closely relates to my paper, there are limitations to the external validity of their results. First, Haitian refugees in the aftermath of the earthquake were granted Temporary Protective Status (TPS), which expired in January 2018 (Schulz and Batalova 2017). Therefore, these individuals fail to represent the refugee population for whom it is important to understand permanent, long-term resettlement effects. Second, given the unexpected and short-term cause of their migration, Haitian refugee children are less likely to have interrupted education, or gaps in content knowledge to the same extent as traditional refugee students. Third, Florida is a traditional immigrant-hosting state with a sizable Haitian population even before 2010. Therefore, local integration of Haitian refugees is likely to have been easier than for other refugee ethnic groups resettled in non-traditional refugee-hosting states. In light of these conditions, my paper can provide new knowledge on the indirect costs of refugees on local education systems and serve as a test of the external validity of the findings of Figlio and Özek.

From an international context, some evidence suggests no effect of immigrants on native students' test scores (Geay, McNally, and Telhaj 2013; Bossavie 2018; Ohinata and Ours 2013), grade repetition (Schneeweis 2015; Pedraja-Chaparro et al., 2016), or likelihood to enroll in high-track schools after compulsory education (Schneeweis 2015). While some studies find no average negative effects, there is evidence of adverse impacts on performance among other immigrants, particularly those of the same native country (Schneeweis 2015; Pedraja-Chaparro et al., 2016). However, it remains inconclusive whether the concentration of foreign-born students matters for the academic achievement of immigrants.[21]

The substantial literature on immigrant peer effects may not translate to the unique context of refugees.[22] Thus, it remains unknown whether there are achievement spillovers from attending school with refugee children. Apart from differences in the nature of their migration (voluntary vs. forced), we can expect differences in immigrant and refugee peer effects for three specific reasons. First, due to their unplanned and traumatic displacement, refugee students are more prone to involuntary gaps in education and are likely to experience acute emotional and psychological distress as a result of their past circumstances (Fazel et al., 2012). Both of these characteristics have implications for the academic achievement and behavior or refugee students, which can also affect their peers. Second, in contrast to low-income voluntary migrants, refugee families have access to a host of services upon arrival to the US, which can ease their integration process. Third, public school districts that receive a large inflow of refugee students qualify to receive additional federal funding through the Refugee School Impact Grant (RSIG) to provide support for refugee students and families. Schools can either access these grants directly, using

[21] There is evidence that shows positive immigrant spillovers (Åslund et al. 2011), as well as no peer effect for immigrant students (Cortes 2006).
[22] For example, in a different context, Bollinger and Hagstrom (2008) show differences in food stamp policy responses between refugees and immigrants.

them to provide support programs for refugees, or partner with refugee-serving organizations that can also receive RSIG funds. Traditionally, these additional resources fund afterschool and parent-engagement activities. In light of these differences, it is conceivable that the peer effects associated with refugee students can differ from previous immigrant studies. Therefore, my paper contributes to the literature on immigrant peer effects by studying refugees as a unique group of foreign-born students.

This paper also contributes to the emerging literature on the impacts of refugee integration and resettlement. Generally, this literature focuses on both the effects on local communities and the integration outcomes of refugees. The latter has traditionally been constrained by data availability. For example, the Annual Survey of Refugees (ASR) is the only representative data set that focuses on refugees exclusively, and there is no independent identification of refugees apart from other foreign-born populations in large data sets such as the American Community Survey or the Current Population Survey.[23] To circumvent this issue, studies using publicly available data construct a proxy for refugee status based on an individual's country of birth and year of arrival matched to data on aggregate refugee arrivals by country of origin (Cortes 2004; Evans and Fitzgerald 2017; LoPalo 2019). Other studies rely on administrative data from refugee-serving organizations (Beaman 2012).

Research on refugee integration primarily focuses on the labor market outcomes of refugees (Mask 2018; Evans and Fitzgerald 2017), the role of networks on labor market integration (Dagnelie et al., 2019; Beaman 2012), and the impact of public assistance generosity on wages (LoPalo 2019). On the other hand, and to my knowledge, most studies on the impact of US refugee resettlement on local communities focus exclusively on the effect on crime.

[23] Although ASR data have been collected annually since 1980, administrative records from the 2016 survey became available to researchers for the first time in 2018

Amuedo-Dorantes, Bansak, and Pozo (2018) leverage geographic and temporal variation in the share of refugees across US counties to study its relationship to local crime and terrorist events. Results show no statistically significant effect. Similarly, Masterson and Yasenov (2019) leverage a sudden drop in refugee resettlement in January 2017 due to an Executive Order that halted arrival flows. While there was a decrease in arrivals of 65 percent, there was no change in local crime rates.

This paper expands the literature on the impacts of contemporary refugee resettlement on local communities by considering the effects within a context not previously studied, namely local public schools.[24] In addition, while most extant analyses focus on adult refugees, I investigate the effects of the resettlement of refugee children. The focus on education is essential to understand the total costs of refugee resettlement at the local level. In particular, I study the indirect cost associated with having refugee peers in school and whether this has an impact on the academic achievement of non-refugee students. Further, I investigate potential mechanisms that may drive peer effects. Lastly, this is the first study to use school administrative data to identify refugees, utilizing both a proxy and direct identification of refugee students.

1.4 Data and Descriptive Statistics

1.4.1 Data

I utilize individual-level administrative data on the universe of students in grades 3 through 8 who attended public schools between 2008 and 2017 in the district with the highest inflow of refugees in Georgia.[25] Access to administrative records allow me to observe test scores in the End-of-Grade exams for math and English Language Arts (ELA), which I use as my main

[24] In concurrent work, van der Werf (2019) studies the impact of Indochinese refugees from the 1970s on the school and labor market outcomes of natives. Her findings show zero to small positive effects.
[25] I denote years by the end of the Spring semester, such that 2008 refers to the school year 2007-2008. I access the data from the Metro Atlanta Policy Lab for Education (MAPLE).

outcomes of interest. [26] I also obtain information on student absenteeism, disciplinary incidents, and a host of demographic characteristics (e.g. race/ethnicity and gender) and program participation variables such as indicators for participation in Special Education and English as a Second Language (ESL) programs, and whether students are eligible to receive Free or Reduced-Price Lunch (FRL).

The primary variable of interest, the share of refugees, is the number of refugee students divided by the total number of students in each school-grade-year combination.[27] In order to generate this variable, it is obviously necessary to identify refugee students in the sample. As mentioned above, distinguishing refugees from other foreign-born individuals is a common roadblock to the study of refugee integration. To address this issue, I obtain school registration records from the district's International Welcome Center, which contain information on students' self-reported refugee status and date of arrival to the United States.[28] These unique records allow me to distinguish between refugees and other foreign-born students directly.

While the variable on refugee status is self-reported, there are three reasons that validate its credibility. First, given the legal immigration status of refugees and possible access to afterschool programs aimed at refugee students exclusively, families have little incentive to withhold information regarding their refugee status. Second, refugee status is informally verified using students' immigration documents as a form of identification at the time of registration.[29]

[26] I standardize all test score variables to have zero mean and unit variance with respect to the statewide subject-grade-year distribution.
[27] Students are assigned to the school with the longest enrollment. It may or may not be the same school that students attend at the time of the test.
[28] The sample district does not track students' immigration status. It only allows for self-reported identification of refugee students in order to target programs and services to this population.
[29] In conversations with education program coordinators from several refugee resettlement agencies, I learned that most refugees only have their I-94 as their identification document at the time of school registration. Therefore, while the school does not ask for this form during the registration process, refugee parents end up using it to enroll their kids in school and hence it is used to informally verify that students are in fact refugees.

Third, it is common for staff from refugee-serving organizations to accompany recently arrived refugees at the time of registration and encourage parents to provide this information.

1.4.2 Descriptive Statistics

Refugee students make up roughly 3 percent of all student-year observations in the sample. Out of 124 elementary and middle schools, 109 (88 percent) enrolled at least one refugee student during the school years 2008-2017, and in 21 schools (17 percent), refugees make up over one percent of students. Henceforth, the latter will be referred to as high refugee-concentration schools.[30] Among refugee-serving schools, there is an average of five refugee students per grade or 2 percent of total enrollment. In refugee-serving schools with at least one percent of their students identified as refugees, there is an average of 24 refugees per grade or 11 percent of total enrollment. As seen in Figure 1.3, most schools have significant variation in the share of refugee students across grades.[31] Lastly, students in refugee-serving schools are similar across several observable characteristics to students in schools that never enroll refugees. For example, while refugee-serving schools have a higher share of students that qualify for FRL, the difference compared to schools with no refugees is, on average, trivial.[32]

Table 1.1 reports summary statistics for the students in the sample, stratified by nativity status. Whether US-born, immigrant, or refugee, students in the district tend to score below the state average in both ELA and math. Refugee students, on average, score 1.25 standard deviations below the state mean in ELA, compared to 0.21 and 0.33 standard deviations for US-born and immigrant students, respectively. The pattern for math test scores across groups is

[30] Refugee school type is a time-invariant measure that classifies schools by their mean share of refugee students across all grades.
[31] Results are robust to dropping the school with the highest variation in the share of refugee students across grades.
[32] Table 1.2 shows the summary statistics of all non-refugee students across school types. Column (1) presents summary statistics for students in schools where no refugee student was ever enrolled. Columns (2) and (3) show summary statistics for non-refugee students enrolled in refugee-serving schools and the set of schools with a concentration of refugees above 1 percent.

similar, although group differences are smaller. Specifically, refugee students score 0.93 standard deviations below the state average, compared to 0.27 and 0.21 standard deviations for US-born and immigrant students, respectively.

Refugee students attend school fewer days per school year compared to both US-born and immigrant students – in part because refugees can be resettled in the United States at any point during the school year. However, there is no substantial difference in the share of days absent across groups. On average, students in the sample miss school 3 percent of the total days they are enrolled. In contrast, there are noticeable differences in disciplinary incidents across groups. Compared to both US-born and immigrant students, refugee students have fewer disciplinary incidents and are less likely to be involved in disciplinary infractions that lead to school suspensions.

Refugees make up a diverse student group and are likely to live in low-income households. On average, 62 percent of refugees are Asian and 30 percent are Black, 89 percent qualify for FRL, and 84 percent receive English as a Second Language (ESL) services. The peers of refugees are also likely to live in low-income households – 71 percent of US-born and 79 percent of immigrant students qualify for FRL. On average, 73 percent of US-born students in the district are Black and 9 percent receive Special Education services. On the other hand, 44 percent of immigrant students are Hispanic and 42 percent receive ESL services.

1.5 Empirical Strategy

There are three main issues concerning the empirical estimation of peer effects.[33] First, there is simultaneity of outcomes, known as the reflection problem, in which own performance impacts peer performance and simultaneously reflects on own achievement (Manski 1993).

[33] See Epple and Romano (2011) and Moffitt (2004) for thorough theoretical and empirical reviews of the peer effects literature.

Second, individuals in a group tend to be exposed to common inputs, such as sharing the same teacher, thereby impeding the causal identification of peer effects apart from unobservable correlated factors. Third, in the absence of randomization, group formation is endogenous. In light of these challenges, estimating a naïve peer effects regression would lead to biased estimates.

In the context of this paper, correlated inputs and endogenous group formation are of most concern. First, refugee students are not randomly assigned to schools. Therefore, peer selection is endogenous inasmuch as parents make school decisions based on the demographic or socioeconomic composition of the student body. Failure to control for this mechanism would confound any effects driven by school quality. Second, teachers and students are typically not randomly assigned to classrooms. To the extent that there is systematic teacher-student matching as a function of unobserved characteristics that are correlated with the outcomes of interest, failure to account for this confounding factor would not allow for the isolation of peer effects from the impact of teacher quality.

Following the literature on immigrant peer effects, I estimate refugee peer effects using cohort variation in the concentration of refugees within schools and across grades. That is, any confounding effects of school selection are eliminated by comparing students within the same school, and the consequences of endogenous teacher assignment are mitigated by measuring peers at the grade, not the classroom level.[34]

1.5.1 Reduced-form, Linear-in-Means Peer Effects

I specify a reduced-form regression to estimate the effect of a change in the proportion of refugee students at the school-grade-year level on a host of student-level outcomes. Equation 1

[34] Evidence suggests that classroom peer effects are stronger than grade-level effects (Burke and Sass 2013). Thus, by measuring peer effects at the grade-level, my estimates are likely biased toward zero.

represents the preferred regression specification:

$$A_{igst} = \beta_o + \beta_1 \left(\frac{Refugee_{gst}}{N_{gst}}\right) + \beta_2 A_{it-1} + X'_{it}\gamma + \lambda_{st} + \mu_g + \varepsilon_{igst} \qquad (1)$$

A_{igst} is an outcome measure for non-refugee student i in grade g at school s in year t; A_{it-1} is the same outcome in year $t-1$; X_{it} is a vector of individual time-varying characteristics; λ_{st} is a vector of school-by-year fixed effects to account for time-varying school characteristics that can drive changes in the share of refugees and student outcomes (e.g., changes in school leadership); μ_g is a vector of grade fixed effects controlling for grade-level differences in student outcomes; and ε_{igst} is an idiosyncratic error term.[35]

The main variable of interest, $\left(\frac{Refugee_{gst}}{N_{gst}}\right)$, is the share of refugee students in a particular grade, school, and year; where N_{gst} is the total number of students. This variable captures the concentration of refugees at the grade level to which a non-refugee student is exposed. The coefficient β_1 measures the refugee peer effect. The main source of identifying variation is intertemporal changes in the proportion of refugee students in a particular grade-school-year combination. Therefore, in order to interpret β_1 as the causal effect of a change in the grade-level proportion of refugees, it must be the case that the variation in the share of refugee students (across grades within a school in a given year) is orthogonal to any unobserved variables that may affect the change in non-refugee student test scores after controlling for past achievement, time-varying school characteristics, and grade-specific differences in achievement.

Given this preferred specification, any threat to identification would have to come from systematic variation in the share of grade-level refugees and changes in student outcomes in the

[35] I also run specifications where I control for grade-year fixed effects and school-grade fixed effects. Results are robust to these changes. Results are shown in table A1.4.

same school and year. This is arguably an unlikely event, especially due to the inherent uncertainty in the proportion of refugee students in each grade within a school. The within-school temporal variation in refugee students across grades primarily depends on the change in refugee arrivals and the age composition of refugees, which are a function of factors that are plausibly exogenous to other within-school variables affecting student performance across grades. First, annual refugee inflows are directly a function of the supply of refugees (determined by international war and conflict), annual refugee admissions ceilings determined at the federal level,[36] and service capacity of resettlement agencies in the state. Second, the share of refugees across grades depends on the age distribution of the incoming cohort of refugees, which is expected to vary independently of native students' outcomes.

I estimate equation (1) using ELA and math test scores as outcome variables, as well as non-cognitive outcomes like student absenteeism and disciplinary incidents.[37] In addition, I run separate regressions to disentangle the peer effect across different categories of non-refugee peers, namely US-born and immigrant students. I do this to investigate whether there are differential effects possibly driven by differences in mechanisms. For example, it is likely that immigrant and refugee students compete for the same classroom resources (e.g., language support from the teacher) such that an increase in refugee concentration can lead to negative spillover effects among immigrant students.

1.5.2 Reduced-form, Nonlinear-in-Means Peer Effects

Prior evidence suggests that peer effects are stronger nonlinearly with respect to initial student achievement levels (Burke and Sass 2013; Imberman, Kugler, and Sacerdote 2012). It

[36] The Refugee Act of 1980 makes explicit that the lawful entry of refugees into the country is a matter of federal, not state jurisdiction. However, recent changes by Executive Order #13888 gives state and local governments the right to consent to resettlements before any refugees are resettled in a locality (Trump, 2019).
[37] Lagged outcomes are omitted in the regressions for absenteeism and disciplinary incidents.

has also been shown that teachers adjust the level at which they teach in response to changes in classroom composition (Duflo et al., 2011; Lavy and Schlosser 2011). Given that achievement among refugee students is lower than that of non-refugees, it is possible that teachers allocate more time toward review of academic content or slow down the pace of instruction. If this is the case, an increase in the share of refugee students would result in a positive spillover effect among low-achieving peers who would benefit from additional review and reinforcement. On the other hand, teachers may face a tradeoff between spending more time on language support services to accommodate the needs of refugees and spending more time on reviewing core content instruction. If teachers spend more time on language support, this can lead to a decrease in the achievement of peers, especially low-achieving students. Results of the nonlinear-in-means estimations can provide evidence of whether these mechanisms are in effect.[38]

I relax the linear-in-means assumption from equation (1) to explore these possible heterogeneities by student baseline achievement. I estimate the following specification:

$$A_{igst} = \beta_o + \beta_1 \left(\frac{Refugee_{gst}}{N_{gst}}\right) \times low_i + \beta_2 \left(\frac{Refugee_{gst}}{N_{gst}}\right) \times middle_i + \beta_3 \left(\frac{Refugee_{gst}}{N_{gst}}\right) \times high_i + \beta_4 A_{it-1} + \lambda_{st} + \mu_g + \varepsilon_{igst} \quad (2)$$

where low_i, $middle_i$, and $high_i$ are indicators of time-invariant initial achievement; and all other variables are defined as in equation (1). In particular, following Burke and Sass (2013), I assign each student's initial test performance to a low, middle, or high "type" based on whether the student's first observed test score falls in the bottom quintile, between the 20th and 80th percentiles, or the top quintile of the grade-by-year state test score distribution. Estimates of β_1,

[38] There can also be competition over teacher resources between refugees and non-refugee language learners. I estimate heterogeneous refugee peer effects by non-refugee ESL classification to investigate this potential mechanism. Results are shown in Table A1.2 in the appendix.

β_2, and β_3 measure refugee peer effects across students of different underlying baseline achievement.

1.6 Results

1.6.1 Linear-in-Means Refugee Peer Effects

Table 1.3 reports the linear-in-means estimates of refugee peer effects on ELA test scores of non-refugee students, obtained by estimating equation (1). Column (1) shows results using the full sample of non-refugee students, while columns (2) and (3) present results using the subsamples of US-born and immigrant students, respectively. Panel A presents results for the full sample of refugee-serving schools, and Panel B reports results for high refugee-concentration schools. I present results by different school "types" in order to explore whether the peer effects depend on the concentration of refugees at the school level, and as a robustness check that results are driven by significant variation in the share of grade-level refugees, not idiosyncratic differences driven by a handful of students. In all specifications, standard errors are clustered at the school-grade-year level.[39]

Results in Table 1.3 suggest that, on average, there is no impact of refugees on non-refugee performance in ELA. The estimated coefficients are small in magnitude and none are statistically different from zero. Specifically, I estimate that increasing the share of grade-level refugees in high concentration schools by 1 percentage point is associated with a decrease in ELA test scores by 0.0003 standard deviations. Thus, I conclude that changes in the proportion of grade-level refuges, whose average ELA test scores are more than one standard deviation below non-refugee students, does not have a statistically significant impact on average non-refugee ELA test scores.

---

[39] Tables 11 and A4 show estimates checking for the robustness of the results clustering the errors at the school and school-year levels, respectively.

Table 1.4 presents the linear-in-means refugee peer effects for math test scores. Unlike the effects for ELA, all of the point estimates are positive, and I find statistically significant and meaningful impacts among the subset of high refugee-concentration schools. The coefficients are large in magnitude and increase with the school-level concentration of refugees. Specifically, I find that increasing the share of grade-level refugees by 1 percentage point in schools that serve a high proportion of refugees results in higher math scores for nonrefugee students by 0.01 standard deviations. The magnitude of the spillover in math achievement corresponds to roughly one-tenth of the impact of a highly effective teacher (Aaronson et al., 2007; Rivkin et al., 2005; Rockoff, 2004), and it is comparable to the peer effects associated with having more girls as peers (Hoxby, 2000).

A further breakdown of the peer group by nativity status shows that the positive refugee peer effect is higher among US-born students.[40] I also run specifications that allow for non-linearities in the share of refugee students at the grade level.[41] I find a small and insignificant negative spillovers in math achievement at small shares of grade-level refugees, and large positive spillovers at high shares of refugees. Specifically, for the high refugee-concentration schools, positive peer effects in math test scores are realized when refugees in a grade make up at least 1.2 percent of students.

1.6.2 Nonlinear-in-Means Refugee Peer Effects

The first set of results assumed that all non-refugee peers are equally impacted by the proportion of refugee students in their grade. However, extant evidence suggests that peer effects

---

[40] I present additional heterogeneous effects by grade levels in Table A1.3, showing that the positive spillovers in math achievement are higher among students in middle school.

[41] I run specifications that include the square of the share of refugees in a grade, and a separate model interacting the share of refugees in a grade with the level of refugees in a school. Tables A1.5 and A1.6 present results from these estimations using ELA and math test scores as the outcomes, respectively.

are stronger nonlinearly with respect to students' initial achievement (e.g. Burke and Sass, 2013). In addition, nonlinear effects can uncover mechanisms driven by changes in teacher behavior. Tables 5 and 6 present the results where I relax the linear-in-means assumption and estimate differential effects for non-refugee students who are initially low, middle, or high achieving.

Table 1.5 shows the nonlinear peer effects for ELA test scores. I find differential peer effects by initial ELA achievement that suggest possible competition over teacher resources between low-achieving students and refugees.[42] Specifically, a 1 percentage point increase in the share of refugee students leads to lower ELA scores by as much as 0.006 standard deviations for low-achieving students. On the other hand, I find positive spillover effects for high-achieving students by up to 0.01 standard deviations. Estimates are statistically significant for the pooled sample of non-refugees and the subsample of US-born students, and across all school types. I find small and insignificant effects for middling students.

Table 1.6 presents nonlinear-in-means results for math test scores of non-refugee students. I find positive spillovers in math scores for students across all levels of initial achievement, with large and statistically significant effects for high-achieving students, and middling students enrolled in high refugee-concentration schools. In sum, to the extent that teachers may adjust their class time allocation or core content focus, I find no evidence that refugee peers experience a decrease in math performance. This stands in contrast to the nonlinear results for ELA where I find evidence of potential competition between low-achieving students and refugees. The differences in nonlinear effects across subjects are reasonable given that

[42] Table A1.2 in the appendix shows results of a separate specification that explores whether there are competition effects between refugees and non-refugees who receive ESL services. If there is competition over teacher resources it should be stronger among students who have similar needs, in this case language support. I find a negative relationship between the share of refugees and ELA achievement among non-refugee students who receive ESL support across all school types, with significant effects for high refugee-concentration schools. I find no differential effect in math.

average test scores differences between refugees and non-refugees are wider for ELA, making it likely that refugees require relatively more teacher time and resources in this subject.

1.6.3 Additional Mechanisms

In addition to changes in teacher time allocation, discussed above, I explore two other mechanisms that can give rise to spillover effects driven by a change in the share of refugee peers. First, I explore changes in class size and classroom resources due to variation in the share of English Learners (ELs) in the classroom. Approximately 84 percent of refugee students in the sample are classified as ELs; therefore, an increase in the share of refugee students also increases the proportion of ELs in a grade.

While previous research shows that having more EL peers is associated with lower test scores for non-EL students (Cho 2012; Diette and Oyelere 2014; Ahn and Jepsen 2015), a related strand of literature suggests that effects can differ by the type of ESL service provided to ELs.[43] For example, Chin et al. (2013) find that providing bilingual education programs aimed at increasing achievement among ELs has positive spillover effects for non-EL students. Thus, it is plausible that an increase in EL students due to an increase in refugees has implications for non-EL students depending on the type of ESL instruction. If EL students are served by a “pull-out” model and thus are instructed in a separate classroom during a portion of the day, a higher share of refugee students implies a temporary reduction in class size for non-refugee peers. On the other hand, if EL students are served by a “push-in” model where a co-teacher is uniquely focused on assisting ELs in the classroom, a higher share of refugee peers leads to an increase in classroom resources which allow the principal teacher to re-allocate their time and instruction to exclusively service non-refugee students.

[43] Evidence on the impacts of bilingual education programs (e.g. Steele et al. 2017 and Bibler 2018) suggests possible positive EL peer effects in non-traditional school settings.

While I do not present results that directly disentangle these two mechanisms, I explore whether there are differential effects by refugee students' EL classification. Results are shown in Table 1.7. Most of the peer effects associated with EL refugees are positive, and all math effects are statistically significant. This suggests that the positive spillovers in math are possibly explained by a change in classroom resources tied to changes in the proportion of ELs in a grade. On the other hand, most of the coefficients associated with non-EL refugee peer effects are negative and all are statistically insignificant.

The second mechanism that I explore is changes in access to auxiliary services aimed at refugees' academic success and overall school integration. Schools with a significant number of refugees commonly partner with refugee-serving organization to provide academic support programs that focus on homework assistance and tutoring.[44] Although in principle these programs are intended to serve refugee students, some of these services are also made available to non-refugees. To the extent that there are spillovers to non-refugee students, this can impact peer test scores positively. It may also be the case that these additional resources loosen schools' budget constraints, thus allowing them to provide more services to nonrefugee students. Importantly, these services are targeted to specific grades with a relatively high proportion of refugees. Therefore, access to these programs can vary across grades within the same school and year. Given that these afterschool services would not be made available in the absence of refugee students, I interpret this as part of the total refugee peer effect.

While I do not have data on afterschool program provision or student participation, I explore this mechanism by exploiting the fact that funding for several of these programs is tied to refugee length of stay in the country. For example, afterschool programs funded by the Refugee

[44] Other services include summer programs and enrichment activities.

School Impact Grant (RSIG) focus exclusively on refugees who have been in the country five years or less.[45] Thus, I leverage individual-level information on students' date of entry into the US to explore whether peer effects differ between short-term refuges (five years or less) and long-term refugees (6 years or more).[46] If non-refugee students are exposed to a higher share of short-term refugees, they are more likely to have access to afterschool programs and other targeted educational support. Thus, estimates from this specification provide suggestive evidence of whether this mechanism is in effect.[47]

Table 1.8 presents results on the peer effects associated with short-term (up to 5 years) and long-term (6 years of more) refugees. I find that an increase in the grade-level share of short-term refugees results in higher math achievement. For example, a 1 percentage point increase in the proportion of short-term refugees in high refugee-concentration schools increases math test scores of nonrefugees by 0.013 standard deviations. I do not find statistically significant effects associated with the share of long-term refugees, or any impacts on average ELA performance.[48]

1.6.4 Nonacademic Outcomes

I also estimate several specifications of equation (1) using non-academic outcomes, namely student absenteeism and disciplinary incidents. I measure absenteeism using the share of school days absent, and an indicator of chronic absences equal to one if the student misses at least 10 percent of days enrolled. Table 1.9 presents results on student absenteeism. I find

[45] https://www.acf.hhs.gov/orr/programs/school-impact

[46] The choice to group refugees into these categories comes primarily from the differences in access to targeted educational services, but also the fact that refugees are eligible to apply for US citizenship after living in the country for five years, which can have implications on the educational achievement of refugee children (Felfe et al., 2019).

[47] I also estimate a separate specification where I identify refugees as "recently arrived" if they have been in the country for less than one year, and "settled" if they have lived in the US for one year or more to capture the immediate short-term effects of resettlement. Results are shown in Table A1.1 in the appendix.

[48] I also run specifications where I measure the share of refugees whose year of arrival is the same as the school year. Results are shown in Table A1.4 in the appendix.

evidence of differential impacts across non-refugee student groups. Specifically, I estimate an overall negative and statistically significant decrease in absenteeism for the full sample of non-refugees and the subsample US-born peers. For example, increasing the share of grade-level refugees in high concentration schools by 1 percentage point results in a 0.1 percentage point decrease in the likelihood of being chronically absent. In contrast, I find an increase in the likelihood that immigrant students are chronically absent.

Table 1.10 presents results on the refugee peer effects on disciplinary incidents of non-refugee students. I use three outcome variables: the number of incidents, the likelihood that students have infractions that lead to in-school or out-of-school suspensions, and the likelihood of a fighting incident. I explore changes in fighting incidents to consider infractions that involve peer interactions. Most of the estimates in Table 1.10, using the full sample of non-refugee and the subsample US-born students, are negative for all outcome variables except the likelihood of fighting incidents.[49] On the other hand, all the coefficients using the subsample of immigrant students are positive and some show a statistically significant increase in the likelihood of fighting incidents. For example, I find that increasing the share of grade-level refugees by 1 percentage point leads to a 0.2 percentage point increase in the likelihood of fighting incidents for immigrant students.

Together, unlike the average effects on test scores, I find differential refugee peer effects for non-academic outcomes by non-refugee student groups. In sum, results show that an increase in the share of refugees is results in lower absenteeism and disciplinary incidents for US-born students, while I find opposite results for immigrant students. This provides suggestive evidence

[49] In Table A1.7, I present results for disciplinary outcomes using the subsample of middle school students. I find no statistically significant effects.

of potential differences in peer interactions between refugees and US-born students, and refugees and immigrants.

1.6.5 Robustness Checks

Table 1.11 presents results for several robustness checks of the linear-in-means specification using ELA and math achievement as outcome variables. These are obtained from estimating variants of equation (1). In all previous specifications, I clustered standard errors by school-grade-year, given that this is the level of treatment. I first check whether results are robust to clustering at the school-level. Results remain unchanged for ELA; however, the standard errors for math increase substantially leading to statistically insignificant results.[50]

I conduct three robustness checks that address issues of potential measurement error in the outcome variables. First, in 2015 there was a change in the End-of-Grade exams and it was reported that, as part of the transition, some districts experienced technology-related issues that led to possibly unreliable test scores.[51] I re-run the preferred specification excluding the school year 2015 to check whether the results are sensitive to this potential source of measurement error in test scores. Second, I run gains models where I assume the coefficient on lagged achievement is equal to 1 and estimate the main specification using the change in test scores as the outcome variable in order to account for possible bias from the inclusion of a lagged dependent variable, which is measured with error (Koedel et al., 2015; Sass, Semykina, and Harris 2014). Third, I estimate models where I control for both lagged test scores to mitigate the effects of measurement error (Lockwood and McCaffrey 2014). While the coefficients estimated in these robustness checks are smaller in magnitude, they are all qualitatively similar to the main results.

---

[50] In Table A1.4 I present results where I cluster errors at the school-year level. All baseline results remain unchanged.

[51] See https://www.ajc.com/blog/get-schooled/testing-glitches-mean-milestones-will-not-count-for-retention/azvpotAK40vloy7ndmx5bL/

I also check whether my results are sensitive to definitions of refugee students and refugee-serving schools. Frist, some of the students who identify as refugees also report being born in the US. I exclude these students when I generate the share of grade-level refugee used in the main specification. In this section, I check whether results change when I count these students as “refugees”. Second, I classify schools as refugee-serving if the school-wide share of refugee students is nonzero at any point during the sample period, including all grades, even those that are not used in the estimation sample.[52] I check whether results are sensitive to restricting this variable to grades 3 through 8. Third, I exclude from the sample the school with the highest variation in the share of grade-level refugees. As shown in Table 1.11, the main linear-in-means effects are robust to redefining refugees and refugee-serving schools.

Lastly, I run specifications controlling for school-by-grade fixed effects and grade-by-year fixed effects to account for variation in grade-level characteristics within schools and grade-level characteristics over time, respectively. I also run a specification where I control for all two-way fixed effects so that the identifying variation comes from temporal changes in the proportion of refugees within the same grade and school. As shown in Table A1.4 in the appendix, even accounting for all these fixed effects, the conclusion remains that increasing the proportion of refugees at the grade level leads to positive spillovers in math achievement.

1.7 Conclusion

Over the past few years, the United States has made large cuts to the number of refugees who are allowed to resettle in the country. The FY2020 cap on arrivals, set at 18,000, is the lowest since the beginning of the refugee resettlement program in the 1980s. Much of the policy debate on whether to change the flow of refugee resettlement centers in part on the perceived

[52] For example, I use the K-5 concentration of refugees to designate whether an elementary school serves refugees.

costs that refugees may impose on local communities. While there is some research on the impact of resettlement, I present the first evidence on the spillovers associated with the resettlement of refugee children within the specific context of local public schools. Using individual-level data from the school district with the largest refugee resettlement population in Georgia, I estimate the effects on ELA and math achievement of non-refugee students resulting from increases in the share of refugees in their school and grade. I also estimate effects on non-academic outcomes such as student absenteeism and disciplinary incidents.

I find no evidence of widespread detrimental academic effects due to an increase in the share of refugee students at the grade level. Rather, results show that increasing the share of refugees is associated with higher math test scores for non-refugee students in schools that enroll a high proportion of refugees, and no average impact in ELA achievement. Specifically, I find that increasing the proportion of refugees by 1 percentage point (roughly half the current average share) in high refugee-concentration schools is associated with increases in math scores of non-refugee students by 0.01 standard deviations. Heterogeneous analyses by refugees' EL classification and length of stay in the country provide suggestive evidence that the positive peer effects in math may be driven by changes in classroom resources and access to academic support programs.

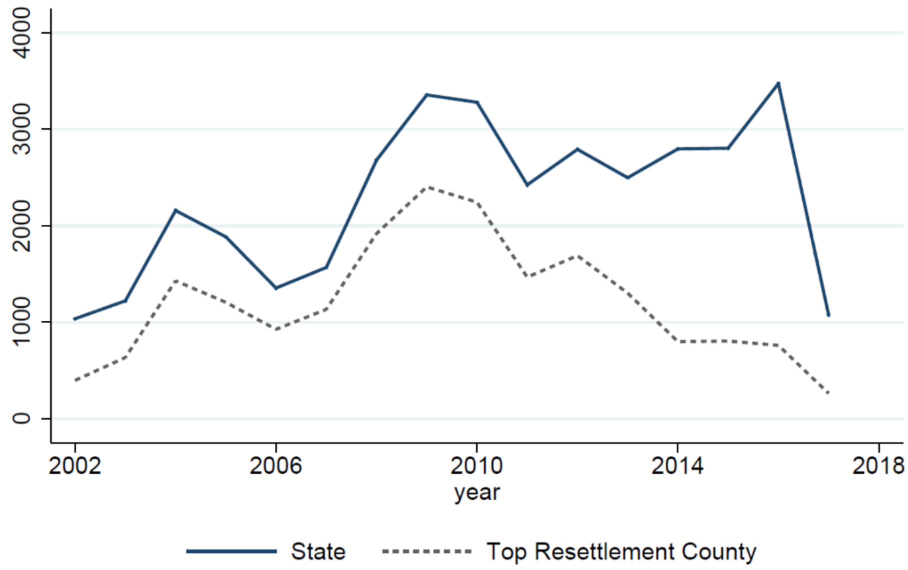
While I find no peer effects on average ELA achievement, using nonlinear-in-means estimations I find evidence that an increase in the share of refugees is associated with negative effects in ELA test scores for low-achieving students and positive effects for high-achieving students. These findings suggest possible competition over teacher and classroom resources between initially low-achieving students and refugees.

I also find evidence of differential effects on nonacademic outcomes across US-born and immigrant students. Results show positive spillovers in student attendance and no evidence of an increase in disciplinary incidents among US-born students. However, results show an increase in the likelihood of being chronically absent and the likelihood of fighting incidents among the immigrant student subsample.

This paper focuses on the indirect costs of an increase in the share of refugees in a grade and mostly finds no evidence of widespread detrimental effects, which is in line with extant evidence on the effect of vulnerable immigrants on the academic achievement of incumbent students (Figlio and Özek 2019). However, it is important to note that these results only provide evidence on the spillover impacts and do not reflect the complete cost of educating refugee children. In addition, I only consider the externalities of resettlement within the context local public schools. However, while the focus on education is a narrow one, research has shown null effects of resettlement in other contexts, such as local crime (Amuedo-Dorantes et al., 2018; Masterson and Yasenov 2019). Together, these findings point to small, if any, negative externalities associated with refugee resettlement.

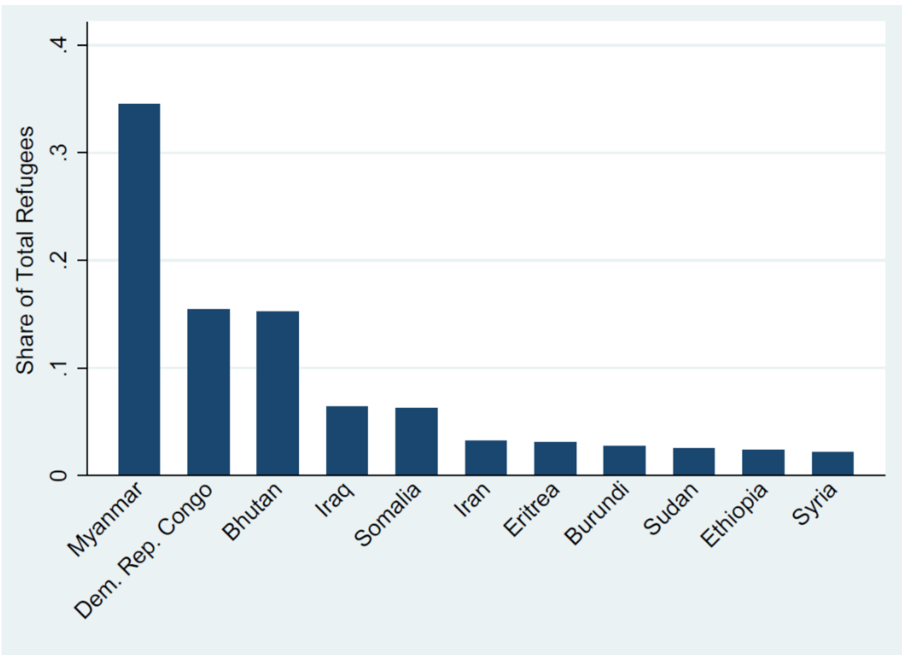
1.8 Figures and Tables

Figure 1.1: Georgia Annual Refugee Arrivals (2002-2017)



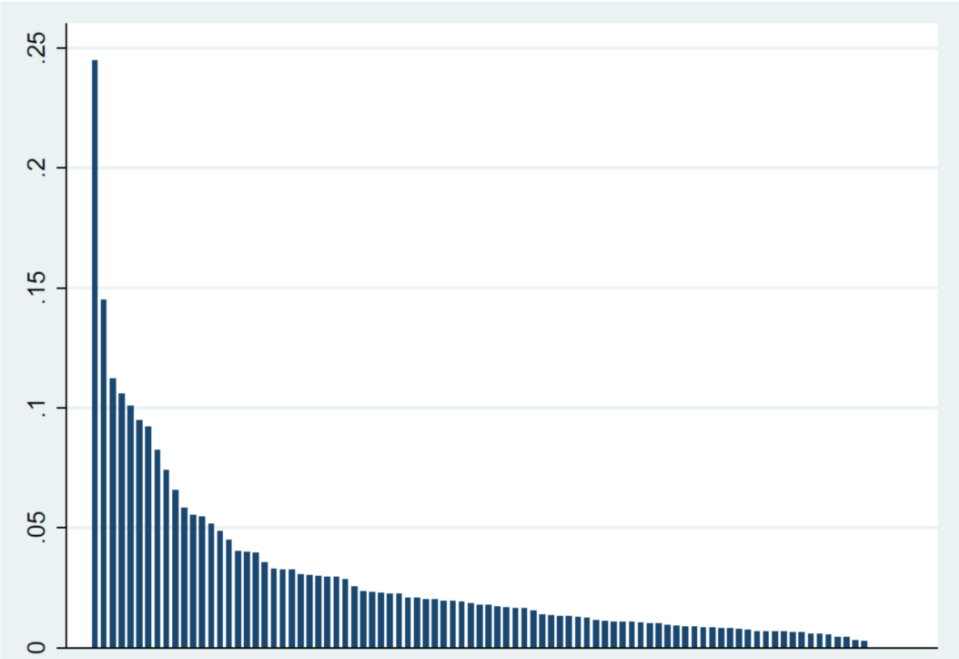
Source: U.S. Refugee Processing Center

Figure 1.2: Top Countries of Origin, Georgia Refugees (2015)



Source: U.S. Refugee Processing Center

Figure 1.3: Maximum Variation in the Share of Grade-Level Refugees, All Refugee-Serving Schools



Note: Each bar represents a school.

Table 1.1: Summary Statistics by Student Nativity Status, Grades 3-8 (2008-2017)

	U.S. Born		Foreign Born			
			Immigrant		Refugee	
	Mean	SD	Mean	SD	Mean	SD
Achievement						
Normalized ELA Score	-0.21	1.01	-0.33	1.12	-1.25	1.03
Normalized Math Score	-0.27	0.95	-0.21	1.06	-0.93	0.90
Absenteeism						
Days Enrolled	167.89	28.10	164.66	32.78	157.82	38.99
Share days Absent	0.03	0.04	0.03	0.04	0.03	0.05
Chronically Absent	0.07	0.26	0.05	0.22	0.07	0.26
Discipline						
Number of Disciplinary Incidents	0.49	1.57	0.29	1.12	0.20	0.97
Serious Disciplinary Incident	0.17	0.38	0.12	0.33	0.08	0.27
Fighting Incident	0.09	0.28	0.05	0.21	0.04	0.20
Demographics						
Female	0.49	0.50	0.48	0.50	0.48	0.50
Hispanic	0.12	0.32	0.44	0.50	0.03	0.16
Black	0.73	0.44	0.27	0.45	0.30	0.46
White	0.14	0.35	0.16	0.36	0.07	0.25
Asian	0.03	0.16	0.22	0.41	0.62	0.49
Special Ed	0.09	0.29	0.05	0.22	0.02	0.14
Gifted	0.17	0.37	0.11	0.32	0.01	0.12
FRL	0.71	0.46	0.79	0.41	0.89	0.31
Current EL	0.05	0.21	0.42	0.49	0.84	0.37
Ever EL	0.08	0.28	0.61	0.49	0.98	0.13
Observations	396,560		29,546		12,737	

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. “Immigrant” refers to all foreign-born students who do not self-report as refugees. Nativity status is based on the country of birth reported in the first year that students appear in the data.

Table 1.2: Summary Statistics of Non-Refugee Students by School Type, Grades 3-8 (2008-2017)

	No Refugee Serving Schools		All Refugee-Serving Schools		Refugee Serving Schools above 1 Percent	
	Mean	SD	Mean	SD	Mean	SD
Total Refugee Students by Grade			5.24	17.46	24.81	33.10
Share Refugee Students by Grade			0.02	0.07	0.11	0.13
Achievement						
Normalized ELA Score	-0.22	1.02	-0.22	1.01	-0.32	0.99
Normalized Math Score	-0.32	0.93	-0.26	0.96	-0.36	0.90
Absenteeism						
Days Enrolled	168.75	27.00	167.51	28.66	165.11	31.44
Share days Absent	0.03	0.04	0.03	0.04	0.04	0.04
Chronically Absent	0.06	0.25	0.07	0.26	0.08	0.28
Discipline						
Number of Disciplinary Incidents	0.36	1.38	0.49	1.56	0.59	1.72
Serious Disciplinary Incident	0.12	0.33	0.17	0.38	0.20	0.40
Fighting Incident	0.07	0.25	0.09	0.28	0.10	0.30
Demographics						
Female	0.49	0.50	0.49	0.50	0.48	0.50
Hispanic	0.05	0.22	0.15	0.36	0.12	0.33
Black	0.85	0.36	0.68	0.47	0.70	0.46
White	0.10	0.30	0.15	0.35	0.14	0.35
Asian	0.01	0.11	0.04	0.20	0.07	0.26
Special Ed	0.09	0.28	0.09	0.29	0.10	0.30
Gifted	0.16	0.37	0.16	0.37	0.11	0.32
FRL	0.68	0.47	0.72	0.45	0.80	0.40
Current EL	0.02	0.15	0.08	0.27	0.11	0.31
Ever EL	0.04	0.18	0.13	0.34	0.16	0.37
No. of Schools	15		109		21	
Observations	52,582		373,524		72,267	

Note: Sample of all non-refugee peers. I determined school type based on the mean proportion of refugee students over the sample period and across all grades. By construction, the total and share of refugees by grade in schools that never serve refugees is zero.

Table 1.3: ELA Achievement of Non-Refugee Students (Grades 4 – 8): Linear-in-Means Refugee Peer Effects

	(1)	(2)	(3)
	All	US Born	Immigrant
VARIABLES			
Panel A: All refugee-serving schools			
Share of grade-level refugees	0.102	0.126	-0.095
	(0.216)	(0.224)	(0.495)
Obs.	239,789	223,770	16,019
Panel B: Refugee-serving schools above 1 % concentration			
Share of grade-level refugees	-0.030	-0.026	-0.033
	(0.247)	(0.256)	(0.457)
Obs.	44,772	40,379	4,393
School-Year FE	X	X	X
Grade FE	X	X	X
Demographic controls	X	X	X
Lagged test score	X	X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. Column (1) reports estimates using the full sample of all non-refugee students. Columns (2) and (3) report estimates using the US Born and Immigrant sub-samples, respectively. ELA test scores are normalized to zero mean and unit variance with respect to the state-wide test score distribution by year-subject-grade. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table1.4: Math Achievement of Non-Refugee Students (Grades 4 – 8): Linear-in-Means Refugee Peer Effects

	(1)	(2)	(3)
	All	US Born	Immigrant
VARIABLES			
Panel A: All refugee-serving schools			
Share of grade-level refugees	0.603	0.662	0.311
	(0.401)	(0.429)	(0.428)
Obs.	238,920	222,035	16,885
Panel B: Refugee-serving schools above 1 % concentration			
Share of grade-level refugees	1.002**	1.041**	0.855*
	(0.412)	(0.443)	(0.470)
Obs.	44,800	40,056	4,744
School-Year FE	X	X	X
Grade FE	X	X	X
Demographic controls	X	X	X
Lagged test score	X	X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. Column (1) reports estimates using the full sample of all non-refugee students. Columns (2) and (3) report estimates using the US Born and Immigrant sub-samples, respectively. Math test scores are normalized to zero mean and unit variance with respect to the state-wide test score distribution by year-subject-grade. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.5: ELA Achievement of Non-Refugee Students (Grades 4 – 8): Nonlinear-in-Means Refugee Peer Effects

	(1)	(2)	(3)
	All	US Born	Immigrant
VARIABLES			
Panel A: All refugee-serving schools			
Share of grade-level refugees × low achievement	-0.405*	-0.422*	-0.382
	(0.225)	(0.235)	(0.502)
Share of grade-level refugees × mid achievement	0.070	0.081	0.036
	(0.220)	(0.228)	(0.506)
Share of grade-level refugees × high achievement	0.930***	0.990***	0.496
	(0.232)	(0.241)	(0.514)
Obs.	239,789	223,770	16,019
Panel B: Refugee-serving schools above 1 % concentration			
Share of grade-level refugees × low achievement	-0.576**	-0.615**	-0.303
	(0.256)	(0.266)	(0.461)
Share of grade-level refugees × mid achievement	-0.025	-0.041	0.103
	(0.249)	(0.259)	(0.459)
Share of grade-level refugees × high achievement	0.840***	0.883***	0.477
	(0.258)	(0.269)	(0.480)
Obs.	44,772	40,379	4,393
School-Year FE	X	X	X
Grade FE	X	X	X
Demographic controls	X	X	X
Lagged test score	X	X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. Column (1) reports estimates using the full sample of all non-refugee students. Columns (2) and (3) report estimates using the US Born and Immigrant sub-samples, respectively. ELA test scores are normalized to zero mean and unit variance with respect to the state-wide test score distribution by year-subject-grade. Students are classified into “low”, “high”, or “middle” initial achievement based on whether their first-observed ELA test score falls in the bottom quintile, highest quintile, or middle quintiles of the state-wide distribution. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.6: Math Achievement of Non-Refugee Students (Grades 4 – 8): Nonlinear-in-Means Refugee Peer Effects

	(1)	(2)	(3)
	All	US Born	Immigrant
VARIABLES			
Panel A: All refugee-serving schools			
Share of grade-level refugees × low achievement	0.288	0.283	0.143
	(0.411)	(0.438)	(0.439)
Share of grade-level refugees × mid achievement	0.513	0.584	0.224
	(0.401)	(0.428)	(0.434)
Share of grade-level refugees × high achievement	1.402***	1.509***	0.949**
	(0.404)	(0.431)	(0.454)
Obs.	238,920	222,035	16,885
Panel B: Refugee-serving schools above 1 % concentration			
Share of grade-level refugees × low achievement	0.632	0.624	0.663
	(0.417)	(0.446)	(0.474)
Share of grade-level refugees × mid achievement	0.954**	1.002**	0.785
	(0.411)	(0.441)	(0.478)
Share of grade-level refugees × high achievement	1.855***	1.928***	1.507***
	(0.420)	(0.450)	(0.506)
Obs.	44,800	40,056	4,744
School-Year FE	X	X	X
Grade FE	X	X	X
Demographic controls	X	X	X
Lagged test score	X	X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. Column (1) reports estimates using the full sample of all non-refugee students. Columns (2) and (3) report estimates using the US Born and Immigrant sub-samples, respectively. Math test scores are normalized to zero mean and unit variance with respect to the state-wide test score distribution by year-subject-grade. Students are classified into "low", "high", or "middle" initial achievement based on whether their first-observed math test score falls in the bottom quintile, highest quintile, or middle quintiles of the state-wide distribution. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 1.7: ELA and Math Achievement of Non-Refugee Students (Grades 4 – 8): Heterogeneous Effects by Refugee EL Status

VARIABLES	(1) ELA	(2) Math
Panel A: All refugee-serving schools		
Share of EL grade-level refugees	0.174	0.784*
	(0.238)	(0.424)
Share of non-EL grade-level refugees	-0.185	-0.132
	(0.357)	(0.681)
Obs.	239,789	238,920
Panel B: Refugee-serving schools above 1 % concentration		
Share of EL grade-level refugees	0.039	1.175***
	(0.261)	(0.433)
Share of non-EL grade-level refugees	-0.413	0.046
	(0.407)	(0.738)
Obs.	44,772	44,800
School-Year FE	X	X
Grade FE	X	X
Demographic controls	X	X
Lagged test score	X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. EL status of refugees is determined based on whether the student receives ESL services. Column (1) reports estimates using the full sample of all non-refugee students using ELA test scores as the outcome variable. Column (2) reports estimates using the full sample of all non-refugee students using math test scores as the outcome variable. ELA and math test scores are normalized to zero mean and unit variance with respect to the state-wide test score distribution by year-subject-grade. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses.

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

Table 1.8: ELA and Math Achievement of Non-Refugee Students (Grades 4 – 8): Heterogeneous Effects by Refugees Length of Stay, Up to 5 Years Post Arrival

VARIABLES	(1) ELA	(2) Math
Panel A: All refugee-serving schools		
Share of grade-level refugees, up to 5 years	0.100	0.753*
	(0.225)	(0.420)
Share of grade-level refugees, 6+ years	0.109	0.080
	(0.301)	(0.507)
Obs.	239,789	238,920
Panel B: Refugee-serving schools above 1 % concentration		
Share of grade-level refugees, up to 5 years	-0.002	1.331***
	(0.257)	(0.426)
Share of grade-level refugees, 6+ years	-0.141	-0.279
	(0.343)	(0.549)
Obs.	44,772	44,800
School-Year FE	X	X
Grade FE	X	X
Demographic controls	X	X
Lagged test score	X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. Column (1) reports estimates using the full sample of all non-refugee students using ELA test scores as the outcome variable. Column (2) reports estimates using the full sample of all non-refugee students using math test scores as the outcome variable. ELA and math test scores are normalized to zero mean and unit variance with respect to the state-wide test score distribution by year-subject-grade. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.9: Absenteeism of Non-Refugee Students (Grades 3 – 8)

	Outcome: Indicator for chronically absent			Outcome: Share of days absent		
	(1) All	(2) US Born	(3) Immigrant	(1) All	(2) US Born	(3) Immigrant
VARIABLES						
Panel A: All refugee-serving schools						
Share of grade-level refugees	-0.096*	-0.132**	0.141*	-0.021**	-0.026**	0.010
	(0.050)	(0.058)	(0.082)	(0.009)	(0.010)	(0.016)
Obs.	372,593	344,803	27,790	372,593	344,803	27,790
Panel B: Refugee-serving schools above 1 % concentration						
Share of grade-level refugees	-0.109**	-0.153**	0.133	-0.021**	-0.028**	0.005
	(0.054)	(0.063)	(0.086)	(0.010)	(0.011)	(0.016)
Obs.	71,336	63,333	8,003	71,336	63,333	8,003
School-Year FE	X	X	X	X	X	X
Grade FE	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. Students are identified as chronically absent if their share of days absent is 0.1 or higher. The share of days absent is computed across all enrollment spells. Column (1) reports estimates using the full sample of all non-refugee students. Columns (2) and (3) report estimates using the US Born and Immigrant sub-samples, respectively. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.10: Disciplinary Incidents of Non-Refugee Students (Grades 3 – 8)

	Outcome: Number of disciplinary incidents			Outcome: Indicator for serious disciplinary incidents			Outcome: Indicator for fighting incident		
	(1) All	(2) US Born	(3) Immigrant	(1) All	(2) US Born	(3) Immigrant	(1) All	(2) US Born	(3) Immigrant
VARIABLES									
Panel A: All refugee-serving schools									
Share of grade-level refugees	-0.773	-1.069*	0.638	-0.113	-0.175*	0.185	0.010	-0.021	0.173*
	(0.548)	(0.640)	(0.535)	(0.095)	(0.105)	(0.124)	(0.081)	(0.089)	(0.098)
Obs.	372,593	344,803	27,790	372,593	344,803	27,790	372,593	344,803	27,790
Panel B: Refugee-serving schools above 1 % concentration									
Share of grade-level refugees	-0.744	-1.032	0.547	-0.087	-0.134	0.122	0.025	-0.003	0.144
	(0.569)	(0.670)	(0.516)	(0.100)	(0.110)	(0.128)	(0.089)	(0.099)	(0.100)
Obs.	71,336	63,333	8,003	71,336	63,333	8,003	71,336	63,333	8,003
School-Year FE	X	X	X	X	X	X	X	X	X
Grade FE	X	X	X	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X	X	X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. Serious disciplinary incidents are defined as incidents that lead to in-school or out-of-school suspensions. Sample includes all students, even those with zero disciplinary incidents. Column (1) reports estimates using the full sample of all non-refugee students. Columns (2) and (3) report estimates using the US Born and Immigrant sub-samples, respectively. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

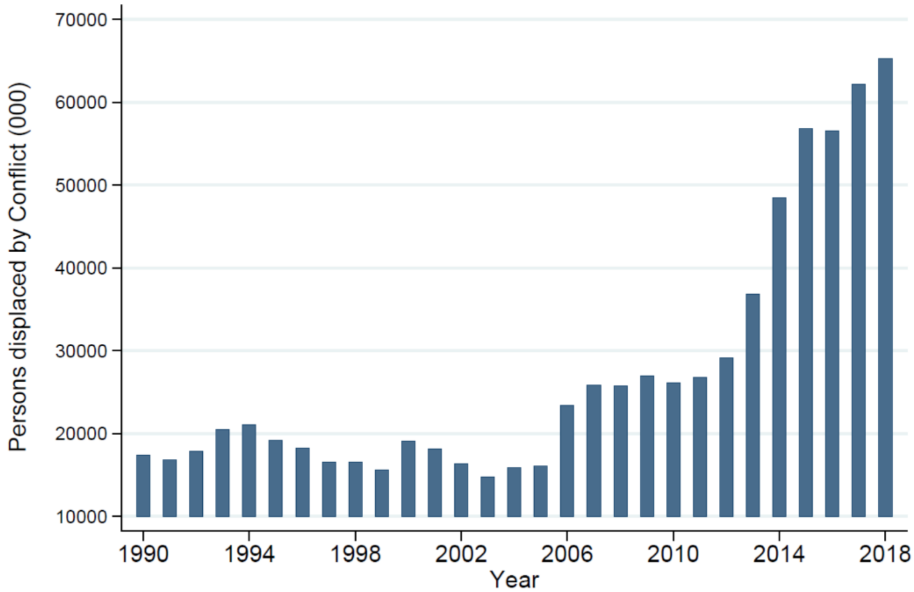
Table 1.11: ELA and Math Achievement of Non-Refugee Students (Grades 4 – 8): Robustness Checks

	Cluster Errors at School-Level		Control for ELA and Math Lagged Scores		Exclude SY 2015		Redefine School Type over Grades 3-8		Include "Refugees" Born in US		Use Change in Norm. Test Score as Outcome Variable		Drop School with Highest Variation	
VARIABLES	(1) ELA	(2) Math	(1) ELA	(2) Math	(1) ELA	(2) Math	(1) ELA	(2) Math	(1) ELA	(2) Math	(1) ELA	(2) Math	(1) ELA	(2) Math
Panel A: All refugee-serving schools														
Share of grade-level refugees	0.102	0.603	0.055	0.542	0.098	0.257	0.109	0.615	0.092	0.641	-0.122	0.427	0.013	0.625
	(0.262)	(0.472)	(0.218)	(0.403)	(0.236)	(0.380)	(0.216)	(0.401)	(0.215)	(0.400)	(0.250)	(0.439)	(0.215)	(0.399)
Obs.	239,789	238,920	239,136	237,506	213,883	213,005	232,245	231,392	242,295	241,425	239,789	238,920	239,080	238,195
Panel B: Refugee-serving schools above 1 % concentration														
Share of grade-level refugees	-0.030	1.002	-0.043	0.975**	-0.038	0.724*	-0.002	1.042**	-0.089	1.112***	-0.124	0.910**	-0.176	0.937**
	(0.296)	(0.607)	(0.247)	(0.412)	(0.268)	(0.386)	(0.241)	(0.404)	(0.247)	(0.414)	(0.278)	(0.444)	(0.239)	(0.434)
Obs.	44,772	44,800	44,627	44,348	40,243	40,265	50,530	50,536	44,607	44,627	44,772	44,800	44,063	44,075
School-Year FE	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Grade FE	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Lagged test score	X	X	X	X	X	X	X	X	X	X			X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. Column (1) reports estimates using the full sample of all non-refugee students using ELA test scores as the outcome variable. Column (2) reports estimates using the full sample of all non-refugee students using math test scores as the outcome variable. ELA and math test scores are normalized to zero mean and unit variance with respect to the state-wide test score distribution by year-subject-grade. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses, except in the first two columns when I cluster errors at the school level.
*** p<0.01, ** p<0.05, * p<0.1

Appendix A1: Additional Figures and Tables

Figure A1.1: Global Trend of Populations Displaced by Conflict (1990-2018)

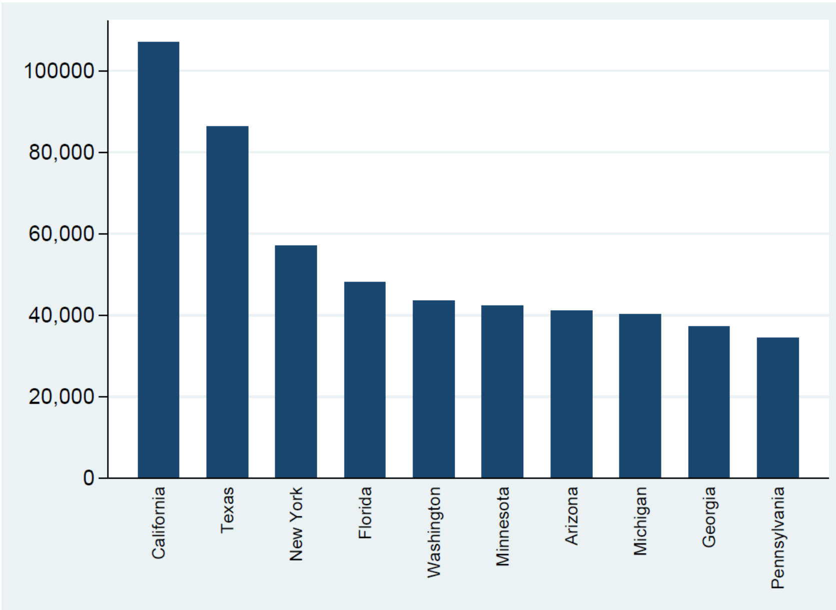


Source: United Nations High Commissioner for Refugees (UNHCR) Population Statistics

Figure A1.2: Refugee Resettlement Flows (1990-2018)

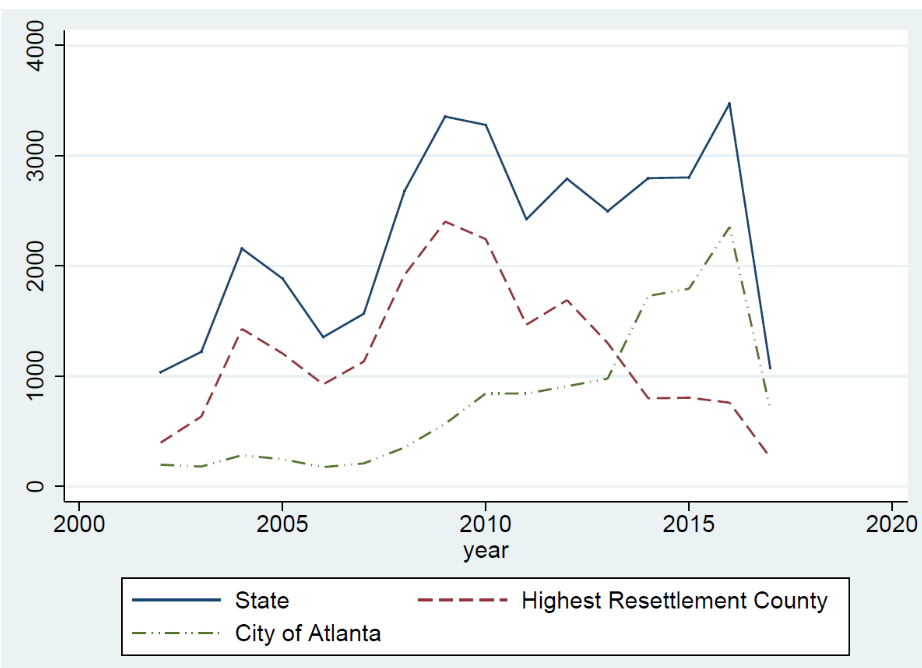
Source: United Nations High Commissioner for Refugees (UNHCR) Population Statistics

Figure A1.3: Cumulative Refugee Arrivals, Top 10 US States (2002-2018)



Source: U.S. Refugee Processing Center

Figure A1.4: Georgia Annual Refugee Arrivals (2002-2017)



Source: U.S. Refugee Processing Center

Table A1.1: ELA and Math Achievement of Non-Refugee Students (Grades 4 – 8): Heterogeneous Effects by Refugees Length of Stay, Recently Arrived

VARIABLES	(1) ELA	(2) Math
Panel A: All refugee-serving schools		
Share of grade-level refugees, recently arrived	-0.139	-0.523
	(0.717)	(1.040)
Share of grade-level refugees, settled	0.120	0.687*
	(0.225)	(0.408)
Obs.	239,789	238,920
Panel B: Refugee-serving schools above 1 % concentration		
Share of grade-level refugees, recently arrived	-0.566	0.508
	(0.717)	(1.014)
Share of grade-level refugees, settled	0.006	1.035**
	(0.253)	(0.425)
Obs.	44,772	44,800
School-Year FE	X	X
Grade FE	X	X
Demographic controls	X	X
Lagged test score	X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. Recently arrived refugees are defined as those whose year of arrival is the same as the school year, i.e. those with less than 1 year since arrival. Settled refugees are defined as those with 1 or more years since arrival. Column (1) reports estimates using the full sample of all non-refugee students using ELA test scores as the outcome variable. Column (2) reports estimates using the full sample of all non-refugee students using math test scores as the outcome variable. ELA and math test scores are normalized to zero mean and unit variance with respect to the state-wide test score distribution by year-subject-grade. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A1.2: ELA and Math Achievement of Non-Refugee Students (Grades 4 – 8): Heterogeneous Effects by Non-refugee EL Status

VARIABLES	(1) ELA	(2) Math
Panel A: All refugee-serving schools		
Share of grade-level refugees	0.109	0.594
	(0.216)	(0.401)
Share of grade-level refugees × ESL	-0.099	0.109
	(0.070)	(0.080)
Obs.	239,789	238,920
Panel B: Refugee-serving schools above 1 % concentration		
Share of grade-level refugees	-0.013	1.008**
	(0.247)	(0.413)
Share of grade-level refugees × ESL	-0.217**	-0.070
	(0.085)	(0.090)
Obs.	44,772	44,800
School-Year FE	X	X
Grade FE	X	X
Demographic controls	X	X
Lagged test score	X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. EL status of non-refugees is determined based on whether the student receives ESL services. Column (1) reports estimates using the full sample of all non-refugee students using ELA test scores as the outcome variable. Column (2) reports estimates using the full sample of all non-refugee students using math test scores as the outcome variable. ELA and math test scores are normalized to zero mean and unit variance with respect to the state-wide test score distribution by year-subject-grade. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A1.3: ELA and Math Achievement of Non-Refugee Students (Grades 4 – 8): Heterogeneous Effects by Grade Levels

VARIABLES	(1) ELA	(2) Math
Panel A: All refugee-serving schools		
Elementary Schools		
Share of Grade-Level Refugees	0.151	0.190
	(0.312)	(0.456)
Obs.	88,714	88,960
Middle Schools		
Share of Grade-Level Refugees	-0.023	0.995
	(0.315)	(0.664)
Obs.	140,642	139,664
Panel B: Refugee-serving schools above 1% concentration		
Elementary Schools		
Share of Grade-Level Refugees	0.143	0.503
	(0.341)	(0.415)
Obs. too	14,897	15,082
Middle Schools		
Share of Grade-Level Refugees	-0.248	1.608**
	(0.353)	(0.730)
Obs.	29,875	29,718
School-Year FE	X	X
Grade FE	X	X
Demographic controls	X	X
Lagged test score	X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. Elementary Schools are classified as schools serving grades K-5 exclusively. Middle Schools are classified as school serving grades 6-8 exclusively. Column (1) reports estimates using the full sample of all non-refugee students using ELA test scores as the outcome variable. Column (2) reports estimates using the full sample of all non-refugee students using math test scores as the outcome variable. ELA and math test scores are normalized to zero mean and unit variance with respect to the state-wide test score distribution by year-subject-grade. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A1.4: ELA and Math Achievement of Non-Refugee Students (Grades 4 – 8): Additional Robustness Checks

	Cluster Errors at School-Year Level		Add Grade-by-Year FEs		Add Grade-by-School FEs		Add Grade-by School and Grade-by-Year FEs		Use Variation in Same-Year Refugee Arrivals	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
VARIABLES	ELA	Math	ELA	Math	ELA	Math	ELA	Math	ELA	Math
Panel A: All refugee-serving schools										
Share of grade-level refugees	0.102	0.603	0.077	0.545*	-0.172	0.737**	-0.207	0.595**	-0.159	-0.647
	(0.274)	(0.517)	(0.200)	(0.314)	(0.209)	(0.348)	(0.201)	(0.274)	(0.711)	(0.989)
Obs.	239,789	238,920	239,789	238,920	239,789	238,920	239,789	238,920	239,789	238,920
Panel B: Refugee-serving schools above 1 % concentration										
Share of grade-level refugees	-0.030	1.002**	-0.118	0.913***	-0.118	0.824**	-0.235	0.708**	-0.568	0.151
	(0.309)	(0.505)	(0.255)	(0.341)	(0.225)	(0.373)	(0.236)	(0.307)	(0.714)	(0.979)
Obs.	44,772	44,800	44,772	44,800	44,772	44,800	44,772	44,800	44,772	44,800
School-Year FE	X	X	X	X	X	X	X	X	X	X
Grade FE	X	X							X	X
Demographic controls	X	X	X	X	X	X	X	X	X	X
Lagged test score	X	X	X	X	X	X	X	X	X	X
Grade-Year FE			X	X			X	X		
Grade-School FE					X	X	X	X		

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. Column (1) reports estimates using the full sample of all non-refugee students using ELA test scores as the outcome variable. Column (2) reports estimates using the full sample of all non-refugee students using math test scores as the outcome variable. ELA and math test scores are normalized to zero mean and unit variance with respect to the state-wide test score distribution by year-subject-grade. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses, except in the first two columns when I cluster errors at the school-year level. The estimates reported in the last two columns, use the share of refugees with same-year arrivals as the variable of interest. That is, it only includes refugee students whose year of arrival is the same as the school year.
*** p<0.01, ** p<0.05, * p<0.1

Table A1.5: ELA Achievement of Non-Refugee Students (Grades 4 – 8): Heterogeneous Effects by Nonlinearities in the Share of Refugees

VARIABLES	(1) ELA	(2) ELA
Panel A: All refugee-serving schools		
Share of grade-level refugees	0.087	-0.000
	(0.354)	(0.314)
Square of Share of grade-level refugees	0.045	
	(0.623)	
Share of grade-level refugees × Total school-level refugees		0.001
		(0.001)
Obs.	239,789	239,789
Panel B: Refugee-serving schools above 1 % concentration		
Share of grade-level refugees	-0.010	-0.066
	(0.419)	(0.370)
Square of Share of grade-level refugees	-0.052	
	(0.672)	
Share of grade-level refugees × Total school-level refugees		0.000
		(0.001)
Obs.	44,772	44,772
School-Year FE	X	X
Grade FE	X	X
Demographic controls	X	X
Lagged test score	X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. All columns report estimates from separate regressions using the full sample of all non-refugee students using ELA test scores as the outcome variable. ELA test scores are normalized to zero mean and unit variance with respect to the state-wide test score distribution by year-subject-grade. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A1.6: Math Achievement of Non-Refugee Students (Grades 4 – 8): Heterogeneous Effects by Nonlinearities in the Share of Refugees

VARIABLES	(1) Math	(2) Math
Panel A: All refugee-serving schools		
Share of grade-level refugees	-0.133	-0.383
	(0.565)	(0.516)
Square of Share of grade-level refugees	2.074*	
	(1.211)	
Share of grade-level refugees × Total school-level refugees		0.006***
		(0.002)
Obs.	238,920	238,920
Panel B: Refugee-serving schools above 1 % concentration		
Share of grade-level refugees	-0.058	-0.281
	(0.619)	(0.547)
Square of Share of grade-level refugees	2.665**	
	(1.311)	
Share of grade-level refugees × Total school-level refugees		0.008***
		(0.002)
Obs.	44,800	44,800
School-Year FE	X	X
Grade FE	X	X
Demographic controls	X	X
Lagged test score	X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. All columns report estimates from separate regressions using the full sample of all non-refugee students using math test scores as the outcome variable. Math test scores are normalized to zero mean and unit variance with respect to the state-wide test score distribution by year-subject-grade. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A1.7: Disciplinary Incidents of Non-Refugee Students (Grades 6 – 8) Middle School Subsample

VARIABLES	(1) Outcome: Number of disciplinary incidents	(2) Outcome: Indicator for serious disciplinary incidents	(3) Outcome: Indicator for fighting incident
Panel A: All refugee-serving schools			
Share of grade-level refugees	-1.557	0.023	0.178
	(1.517)	(0.232)	(0.198)
Obs.	183,776	183,776	183,776
Panel B: Refugee-serving schools above 1 % concentration			
Share of grade-level refugees	-1.532	0.134	0.318
	(1.670)	(0.241)	(0.220)
Obs.	40,440	40,440	40,440
School-Year FE	X	X	X
Grade FE	X	X	X
Demographic controls	X	X	X

Note: Refugee students identified using the self-reported measure, excluding students who self-report as born in the US. Serious disciplinary incidents are defined as incidents that lead to in-school or out-of-school suspensions. Sample includes all students, even those with zero disciplinary incidents. All estimates report results from models using the full sample of all non-refugee students. Demographic controls include indicators for race, gender, free and reduced-price lunch eligibility, and participation in ESL and special education services. Robust standard errors clustered at the school-year-grade level are shown in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix B1: Note on the Identification of Refugee Students

I identify refugee students using data on a self-reported measure of refugee status, allowing me to distinguish between refugees and other foreign-born students directly. While this is the preferred measure, I also explore a second strategy commonly used in the literature (see Cortes 2004, Capps et al. 2015, and Evans and Fitzgerald 2017) to account for possible self-selection arising from voluntary identification as refugees.

This alternative method uses publicly available data on refugee arrivals by year and country of birth to determine whether a foreign-born individual is a refugee. I access city-level data on refugee arrival flows from the Refugee Processing Center (RPC) and construct country-year pairs that indicate the countries from which refugees in Georgia arrived in a particular year. I use this list to identify potential refugee students based on their country of birth. The *potential refugee* student population in my data includes all foreign-born students whose country of origin is on the annual refugee arrivals list within the five years prior to being first observed as a public school student.[53]

The main known source of potential bias in the proxy approach is the likelihood to overcount refugees, especially those who come from countries with a relatively high number of immigrants. There are modified approaches that minimize this issue by restricting the list of likely refugee countries to those with a high refugee-to-foreign-born ratio (see Evans and Fitzgerald 2017 for an example of this approach). To my knowledge, and largely due to data limitations, there has not been a direct assessment of the extent of overcounting that arises from using the proxy approach. Because of the unique data that I use for this paper, I can directly assess the validity of

[53] I allow for a 5-year window to account for the possibility that refugees may arrive as infants and first appear in school records up to 5 years after their arrival. Due to the high annual flow of refugees, most countries do not rely on the 5-year window to be included on the list.

the proxy measure against a self-reported identifier using a longitudinal, administrative data set.

Figure B.1 plots the total refugee counts by the self-reported and proxy measures for a select group of countries to illustrate examples of the differences between the two approaches. A comparison between the two measures reveals that, in addition to overcounting refugees from countries with high immigration, the proxy approach also fails to account for refugees whose country of birth is not on the official list of refugee arrivals to the United States. For example, the proxy measure identifies virtually no refugees from Nepal or Thailand directly from the fact that, in the RPC data, no refugees arrived in the state of Georgia from these countries.[54] However, the self-reported data shows that close to 1,300 refugee students come from these countries. A close investigation into these differences revealed two important facts that helped reconcile these discrepancies. First, none of the self-reported refugees from Thailand or Nepal reported speaking a language commonly spoken in these countries. Instead, they report speaking languages from other refugee-sending countries. Second, both of these countries host large refugee camps, some of which had direct resettlement programs with the United States.[55] Thus, there is evidence to conclude that the self-reported refugees who are not identified using the proxy approach are "true" refugees, likely arriving from camps in transient countries, not students misreporting their refugee status.[56]

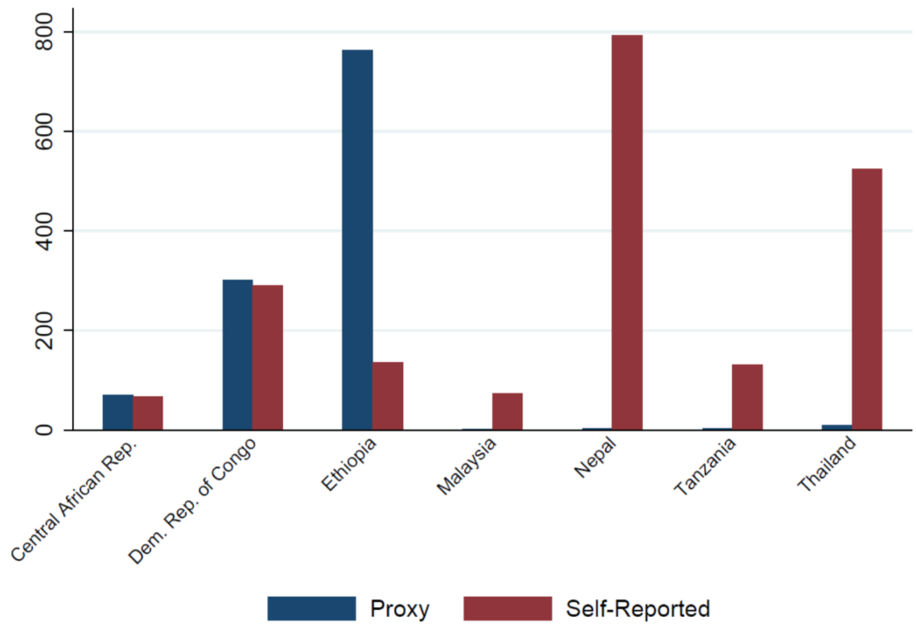
The results presented in the current version of the paper include only those that utilize the self-reported measure to define refugee students, as it provides a more accurate representation of the refugee student population in the sample district. It is important to note, however, that results

[54] I highlight these two countries because they are the ones with the greatest undercounts of refugees. Other countries include Tanzania, Malaysia, and Kenya.
[55] https://www.unhcr.org/en-us/news/latest/2014/1/52e90f8f6/wraps-group-resettlement-myanmar-refugees-thailand.html
[56] I also consulted school administrators and program managers working for refugee-serving organizations in Georgia, who confirmed that many refugees who arrive from camps are likely to mark their "camp country" as their place of birth even if they do not have citizenship from that country.

are sensitive to whether I define refugees using the proxy or self-reported measure. This is, in part, driven by differences in achievement and sociodemographic characteristics. Self-reported refugees have significantly lower test scores and are more likely to be low income and English Learners. Using the proxy method to identify refugees, I find that an increase in the share of refugee students in a grade leads to positive spillovers in ELA and no effect in math. Notably, neither refugee specification shows negative average spillovers in achievement.[57]

[57] Results using the proxy method to identify refugees can be made available upon request.

Figure B1.1: Comparison between Proxy and Self-Reported Counts, Selected Countries



Chapter 2: Effects of an Intensive English Program on the Academic Achievement of English Learners

2.1 Introduction

English Learners (ELs)[58] represent 10 percent of all public school students in the United States and are the fastest-growing student population in the country (Ruiz Soto et al. 2015).[59] ELs are also among the lowest-achieving student subgroups (Snyder et al. 2019).[60] Based on the 2019 NAEP reading scores for 4th graders, the achievement gap between non-ELs and ELs (33 points) is higher than the white-black achievement gap (26 points) and the high-low income achievement gap (28 points).[61] Thus, with their growing prevalence and lag in achievement, understanding the efficacy of policies designed to improve the educational outcomes of ELs is a priority for public schools around the country.

By law, ELs are entitled to adequate grade-appropriate content instruction and English language support; however, there is significant flexibility in policy implementation leading to variation in programs across and within states (Mavrogordato and White 2017). Broadly, the most common policy tools in EL education include initial EL classification, type of English as a Second Language (ESL) instruction,[62] and parameters used to determine EL reclassification.[63] Most causal evidence has focused on the impact of policies that determine EL classification and

---

[58] The National Center for Education Statistics defines English Learners as students who are not proficiency in English due to being born in a non-English-speaking country or living in environments where English is not the dominant language. The literature also refers to ELs as English Language Learners (ELLs), and students with Limited English Proficiency (LEPs) among others.
[59] Figure A2.1 in the appendix shows the percentage of ELs out of total enrollment in US public schools from 2001 to 2017.
[60] Figure A2.2 in the appendix shows the 4th grade reaching achievement gaps based on NAEP test scores from 2000-2017
[61] NAEP, National Assessment of Educational Progress, is a nationally representative assessment administered by the National Center for Education Statistics. See https://www.nationsreportcard.gov/reading/nation/groups/?grade=4 for achievement gaps by groups as of 2019.
[62] It can also be denoted as English for Speakers of Other Languages (ESOL) instruction.
[63] EL reclassification traditionally indicates that a student has achieved adequate English proficiency to stop ESL support services.

reclassification (e.g., Pope 2016; Shin 2018; Johnson 2019), thus leaving a knowledge gap for examining the efficacy of different types of ESL instruction – with the exception of a growing literature on the impact of bilingual education (Steele et al. 2017a; Umansky 2016a; Bibler 2018).[64]

In this paper, I evaluate the impact of an intensive English program on the academic outcomes of ELs with very low English proficiency. Specifically, I study whether program eligibility and enrollment affect test scores in English Language Arts (ELA) and math in the short term. I also explore differential impacts by grade at program eligibility and students' refugee status. To do this, I use individual-level data on the universe of students screened for EL classification in grades 1 through 8 from 2008 to 2018 who attended public schools in one of the largest school districts in Georgia – a state ranking among the top ten states with the fastest-growing EL enrollment as of 2008 and highest total EL enrollment by 2015 (Batalova and McHugh 2010; Ruiz Soto et al. 2015).

As most states in the U.S., students in Georgia are initially classified as ELs based on a test score on an English proficiency exam. However, the studied district also uses the initial proficiency score and grade at EL screening to determine eligibility for an intensive English program aimed at ELs with very low English proficiency. This is a short-term "newcomer" ESL program focused on rapid English-immersion and equipping ELs with basic language skills.[65] Students who participate in the program remain at most one year and later transition to their neighborhood school where they begin traditional ESL instruction - the latter corresponding to a

[64] Most ELs around the country are served by traditional ESL instruction, not bilingual models (U.S. Department of Education 2019).
[65] Newcomer ESL programs can be defined as "specialized academic environments that serve newly arrived, immigrant English language learners for a limited period of time" (Short and Boyson 2012).

combination of push-in or pull-out models combined with regularly scheduled English-only content classes.[66]

I leverage variation in the eligibility criteria to estimate the program effect using intent-to-treat Difference-in-Differences and treatment-on-the-treated Regression Discontinuity designs. In short, in the first empirical approach, I compare the outcomes of students who qualify for the program based on their initial English proficiency score across eligible and ineligible grades.[67] In the second approach, I limit the sample to students who are screened for EL classification in a participating grade and compare short-term outcomes across students whose initial English proficiency score falls within a small window of the cutoff for program eligibility. In both specifications, the control group corresponds to ELs who receive traditional ESL instruction. Therefore, results yield evidence on the impact of an intensive English program relative to typical EL services in neighborhood schools.

I contribute to the literature studying the impact of EL classification and types of ESL instruction in the following three ways.[68] First, I provide novel evidence on the impact of ESL instruction focusing on the subset of ELs with very low English proficiency. In contrast, most of the causal literature focuses on students who score around the cutoff for EL classification. In other words, ELs with relatively high English proficiency.[69] While there is a growing literature in

[66] "Push-in" refers to a model where ELs remain in their core academic class (e.g. reading) where they receive instruction from their content area teacher and a co-teacher who specializes in ESL instruction. "Pull-out" refers to a model where ELs are taken out of the core academic class and receive ESL instruction for a portion of the school day.
[67] I explain the program details in Section 2. Also, see Figure A2.6 for a summary of the program eligibility criteria.
[68] A closely related line of research studies the impact of EL reclassification on academic outcomes. Studies find either positive of null effects. Some find that reclassification leads to gains in academic outcomes whether ELs are reclassified in elementary grades (Pope 2016; Johnson 2019) or high school (Robinson - Cimpian and Thompson 2016; Carlson and Knowles 2016). Others find no impact of EL reclassification (Robinson 2011; Reyes and Hwang 2019).
[69] As an exception, Setren (2019) studies the impact of attending a high-achieving charter on the academic achievement of ELs across the full range of English proficiency. Results show large gains in English exams among ELs with the lowest baseline scores.

this area, prior research on the impact of initial EL classification presents mixed results. Using data from schools in California, Pope (2016) and Shin (2018) show that students who are marginally classified as ELs in Kindergarten have higher test scores in ELA and math, with gains persisting up to middle school (Pope 2016). Studies on longer-term effects, however, find no relationship between initial EL classification and high school graduation or college attendance (Johnson 2019).

In contrast, Umansky (2016a) finds large and sustained negative effects of EL classification in Kindergarten on math and ELA achievement in grades 2 through 10. Explorations into the mechanisms show that the negative effects are concentrated among students who receive English immersion ESL instruction relative to bilingual programs (Umansky 2016a),[70] thus pointing to the importance of differentiating between different types of ESL instructional models.[71] Research also documents unintended consequences of EL classification such as limited access to core content ELA classes and a full academic course load in middle school (Umansky 2016b; Umansky 2018), underrepresentation in honors classes (Umansky 2016b), and lower teacher expectations (Umansky and Dumont 2019).

Second, I present the first rigorous evaluation of a newcomer intensive English program relative to traditional ESL instruction. Extant research on the differential impact of ESL instruction tends to focus on bilingual education relative to English immersion programs (Jepsen 2010; Valentino and Reardon 2015; Umansky 2016a; Steele et al. 2017), with a dearth of evidence examining the effect of different types of English immersion ESL models. Moreover, I study the academic outcomes of students who face different initial types of ESL instruction,

[70] Umansky (2016a) finds no impact of EL classification among students who receive bilingual instruction.
[71] English Immersion corresponds to self-contained grade-level classrooms where ELs receive both ESL and core content instruction from a certified EL teacher, with the goal of fast English acquisition (Wright 2015).

conditional on EL classification. Therefore, I do not confound the EL labeling effect (Umansky and Dumont 2019) with participation in ESL instruction.

Third, I use data from a school district that has a growing and diverse EL population where roughly 40 percent of students screened for EL classification in grades 1 through 12 are refugees. This feature stands in contrast to the vast majority of studies that focus on Hispanic EL students in traditional immigrant-destination states such as California (Umansky and Reardon 2014; Johnson 2019; Umansky 2016; Reyes and Hwang 2019). Thus, my paper provides a test of the external validity of earlier studies while presenting new results that may apply to school districts with a diverse group of ELs and an emerging population of refugees.

From the DID specification, I find that students who are eligible for an intensive English program have significantly lower ELA achievement relative to ELs who receive traditional ESL instruction. Specifically, program eligibility results in a reduction in ELA test scores of 0.17 standard deviations. Additional analyses by grade of eligibility show the negative effects in ELA are large and concentrated among students who are first screened for EL classification in grades 5 and above.

Results from the RD specification also suggest a negative impact of program enrollment on ELA achievement, relative to students whose initial English proficiency is just below the maximum threshold for program participation – these effects are larger than the DID results. In addition, I find large negative effects of program enrollment on math achievement. Together, these results support the hypothesis that part of the negative program effects may be driven by a delay in core-content instruction, which would have greater impacts on students at the threshold of basic English literacy.

Further analyses that allow for heterogenous effects reveals that, unlike the average results, refugee students who are eligible for the intensive English program have higher ELA and math test scores one year after program eligibility. These effects are large and significant, resulting in higher ELA achievement by 0.41 standard deviations and higher math achievement by 0.43 standard deviations. Refugees benefit from program eligibility across all grades of initial EL screening.

The remainder of the paper is organized as follows. Section 2.2 presents a description of the institutional background that determines program eligibility. Section 2.3 describes the data used in the analysis. Section 2.4 outlines the Difference-in-Differences and Regression Discontinuity approaches. Section 2.5 includes a discussion of the results. Section 2.6 concludes.

2.2 Institutional Background

Georgia students in grades 1 through 12 who report speaking a language other than English at home are screened for EL classification using the WIDA Screener, an English proficiency assessment that measures four language domains (listening, speaking, reading, and writing), reporting individual domain scores as well as an overall composite scale score.[72] The latter measure is mapped to proficiency levels ranging from 1.0 to 6.0, which are then used for decision-making: students who score from 1.0 to 4.9 are eligible to be classified as ELs and receive ESL instruction, those with higher scores are not classified as ELs and are scheduled for regular English-only classes.

In addition to initial EL classification, the district uses the WIDA compositive proficiency score in conjunction with the grade at EL screening to determine initial type of ESL services. Students who are first screened in grades 1 or 2 and score 4.9 or below are initially

[72] Students who enroll in Kindergarten are screened for EL classification and ESL services using a different test instrument.

classified as ELs and receive traditional ESL instruction in their neighborhood school, usually a combination of pull-out or push-in services along with a schedule of regular English-only classes.[73] Initial ESL service type is differentiated for students who are first screened in grades 3 and above. Those who score between 2.0 and 4.9 are eligible for the same ESL service type as students screened in grades 1 and 2. However, students who score between 1.0 and 1.9 are eligible for an intensive English program where they receive targeted services for up to one year.[74]

As it is common of most newcomer ESL programs (Short and Boyson 2012), this intensive English program is a short-term intervention designed as a specialized environment for ELs with very low initial English proficiency. The primary goal is to introduce students to the basic English skills – academic and social – that they need prior to enrolling in their neighborhood school. The program is hosted in the district's center dedicated for the education of international students and across eight satellite locations, and participants are tested every nine weeks to determine whether they remain or exit. Thereafter, students move to their assigned neighborhood schools and transition to traditional ESL services.

2.3 Data and Descriptive Statistics

I utilize individual-level administrative data on students in grades 1 through 8 who attended public school in one of the largest and most diverse school districts in metro Atlanta from 2008 to 2018.[75] Over the sample period, the racial/ethnic composition of the district was 70 percent Black, 14 percent White, 13 percent Hispanic, and 6 percent Asian.[76] Approximately 10

---

[73] Most students in Georgia who are initially classified as ELs are provided services in these traditional formats.
[74] Figure A2.6 in the appendix presents a summary of the program eligibility criteria by grade at EL screening.
[75] I denote years by the end of the Spring semester, such that 2008 refers to the 2007-2008 school year. I access data through the Metro Atlanta Policy Lab for Education (MAPLE).
[76] Averages computed by author. Racial categories (white, black, Asian) are mutually exclusive. Hispanic denotes ethnicity, not race.

percent of students were served by ESL instruction, nearly twice the state average of 6.5 percent. A unique feature of the district is that it serves a county where close to 80 percent of refugees who are resettled in the state reside. As of 2017, roughly 4 percent of the students in the district were self-identified as refugees.[77]

The analysis sample consists of students in grades 1 through 8 who at the time of first enrollment in the district report speaking a language other than English at home and are screened for EL classification. Over the sample period, there has been an increase in the number of students who are screened and nearly all have WIDA scores that are low enough to be eligible for EL classification.[78,79] Figure 2.1 shows the distribution of WIDA scores, where it is evident that most students score at the lowest level of proficiency.[80] In particular, nearly 30 percent of the sample have the lowest level of English proficiency and 73 percent fall below the maximum score for program eligibility.

As seen in Figure 2.2, the number of students screened for EL classification peaks in grade 2 and it declines as grade levels increase. In principle, the grade at which students are screened for EL classification depends on the age at which they first enroll in the district. For the vast majority of these students, it is likely that this also coincides with the first time they enter school in the U.S.; 88 percent of students in the sample are foreign-born and 89 percent enter the district from another country or state.[81] Thus, variation in the grade of EL screening is partly driven by differences in age of arrival to the U.S.

[77] The school district does not track students' immigration status. It simply allows for voluntary self-reporting of refugee status in order to better design and target programs that can help refugees integrate into schools.
[78] Figure A2.5 in the appendix shows the number of WIDA takers by school year from 2008-2018.
[79] Exactly 98.46 percent of WIDA takers are eligible for EL classification.
[80] Figure A2.3 in the appendix shows the distribution of WIDA scores for the subset of students who are screened in eligible grades, i.e. the sample used for the RD analysis.
[81] Data limitations do not allow me to distinguish between arrivals from other states and from abroad.

The data I access contain information on students' EL classification, English proficiency screening scores (i.e. WIDA composite scores), grade at EL screening, WIDA test date, and intensive English enrollment status. These data allow me to construct the main variables of interest determining program eligibility (based on initial WIDA score and grade at EL screening) and enrollment. Over the sample period, 2,545 students (57 percent of the sample) were eligible to participate in the program, with an average WIDA score of 1.18.[82] These students make up a very diverse group and are likely to live in low-income households. In particular, the most common racial/ethnic groups are Asian (40 percent), Hispanic (30 percent), and Black (23 percent); 52 percent are self-reported refugees; and 86 percent are eligible for Free or Reduced-Price Lunch (FRL).

As described in Section 2.2, eligibility for program participation is determined by two objective criteria: initial WIDA score and grade of EL screening. However, ultimate enrollment also depends on the consent of parents or legal guardians, as they are allowed to opt-out of the program. As seen in Figure 2.3, program participation varies over time with higher compliance in later years. Specifically, on average 66 percent of eligible students enrolled in the program during the school years 2008-2014. In contrast, 81 percent of eligible students enrolled starting in 2015. I further limit the sample used in the fuzzy RD specification to include only high-compliance years.[83]

As the main outcome variables, I observe individual End-Of-Grade test scores in English Language Arts (ELA) and math. For ease of interpretation, I normalized these variables to mean zero and standard deviation of one with respect to the statewide grade-subject-year test score

[82] WIDA proficiency scores range from 1.0-6.0.
[83] Figure A2.7 in the appendix shows the fuzzy RD first-stage scatter plot for low-compliance years. Notably, there is no change in the likelihood of program enrollment at the cutoff.

distribution. Ideally, I would use concurrent test scores as the preferred outcome; however, this would be most appropriate only if all WIDA takers enroll in the district at the beginning of the school year, such that program exposure at the time of testing would be uniform. Using data on test date, Figure 2.4 shows that this is not the case. In fact, there is significant variation in the month in which students take the WIDA screener, with transitional months (August and January) having the highest counts representing 31 percent of screenings.[84] With this variation in month of WIDA screening, and the fact that I do not observe the length of program participation, I assume that all students participate for one year – the maximum time that students can attend the program. Therefore, I use test scores from the first grade post treatment eligibility to measure the short-term impact of the program.[85]

In addition to variables on EL screening and test scores, I also access information on student attendance, disciplinary records, demographic information (e.g. race and gender), and eligibility to programs such as free and reduced-price lunch (FRL), which I use as control variables in the fully specified models. Table 2.1 reports variable means for the sample used in the DID specification, stratified by program eligibility criteria. Table 2.2 reports variable means for the sample used in the fuzzy RD specification.

2.4 Empirical Strategy

I leverage the eligibility criteria for an intensive English program to employ Difference-in-Differences and fuzzy Regression Discontinuity strategies. These approaches allow me to

[84] There is a more uniform distribution of testing across months within the refugee subsample as seen in Figure A2.8 in the appendix. Given that 90 percent of these students are coming from abroad, this is confirmation of the fact that refugees do not have a choice in the timing of migration to the US.
[85] For example, I use 4th grade test scores to measure of short-term effect of program eligibility for students screened in grade 3.

estimate and compare the Intent-to-Treat (ITT) and Local Average Treatment Effect (LATE) of the program.

2.4.1 Intent-to-Treat Difference-in-Differences

First, I leverage variation in program eligibility based on the maximum WIDA score and grade at the time of EL screening to estimate the following equation:

$$A_{i(t_0+1)} = \alpha + \beta_1 LowWIDA_{it_0} + \beta_2 EligibleGrade_{it_0} + \beta_3 LowWIDA_{it_0} \times EligibleGrade_{it_0}$$
$$+ \gamma WIDA_{it_0} + \boldsymbol{X}'_{i(t_0+1)}\delta + \eta_s + \varepsilon_{i(t_0+1)} \quad (1)$$

where $A_{i(t_0+1)}$ is a normalized achievement measure in ELA or math for student i in the school year post program participation, $t_0 + 1$; $LowWIDA_{it_0}$ is an indicator for the student's WIDA score falling below the maximum threshold for program participation; $EligibleGrade_{it_0}$ is an indicator for the student being first screened for EL services in grade 3 or above; $WIDA_{it_0}$ is a continuous measure of the student's WIDA score; $\boldsymbol{X}_{i(t_0+1)}$ is a vector of student characteristics one year after eligibility for the treatment; η_s is a school fixed effect;[86] and ε is the error term.[87] The parameter of interest, β_3, measures the ITT effect of the program by capturing the achievement difference between students with low initial WIDA scores who are screened for EL classification in an eligible grade, relative to students who do not meet the program eligibility criteria.[88]

There are three potential threats to identifying the causal effect of treatment eligibility. First, by design, this specification compares test scores between older ELs (screened in grades 3

[86] The school FE corresponds to the school with the longest enrollment spell in the year post treatment eligibility.
[87] In my main specification, I cluster the standard errors at the school level. Results are robust to clustering by other variables such as grade, grade-by-school, WIDA score-by-WIDA grade, and WIDA grade.
[88] Note the DID specification does not rely on temporal variation in the existence of the program.

and above) and younger ELs, which may cause the estimates to be biased downward. As documented in prior literature, there is a strong negative correlation between age of arrival to the US and language acquisition (Bleakley and Chin 2004). Thus, it is likely that the estimated coefficient simply reflects the difference in achievement among early and late EL arrivals, irrespective of exposure to the intensive English program. I address this concern by estimating a Regression Discontinuity approach where I compare the outcomes of students who are screened for EL classification in the same grade but face different treatment eligibility based on their initial WIDA score.

The second and third concerns arise from potential endogenous sorting across program eligibility criteria. First, there may be endogenous sorting across grades. For example, it may be the case that parents choose to enroll their child in the district when they are old enough to take advantage of the program. Given that nearly 90 percent of the sample is foreign-born or transferred from abroad, sorting across grades is related to the family's choice of migration. I investigate whether endogenous migration drives differences in results by estimating the program eligibility effects among the subset of refugee students. In principle, these students and their families do not have a choice over the timing of migration, and I observe compelling evidence that this is true. I note differential WIDA screening across grades by EL refugee status, in addition to pronounced differences in the month of WIDA screening by these groups.[89]

Moreover, a concern arises from students being able to manipulate their WIDA score in such a way that it increases their likelihood of program participation. I argue that this is unlikely given that the score used for EL classification and ESL service provision is a composite measure

[89] See Figure A2.4 for the count of WIDA takers by grade and refugee status. See Figure A2.8 for the count of WIDA takers by month and refugee status.

that aggregates four proficiency domains using a function that is unknown to the students.[90] To account for the level of initial English proficiency, however, my preferred specification includes a measure of students' initial WIDA score as an achievement control. Lastly, I also include school fixed effects to account for time invariant characteristics that may drive differences in EL achievement across eligibility groups. For instance, it may be the case that there are differential resources in EL education across schools which may drive student sorting by grade of screening for EL classification. As a robustness check, I also control for student-varying characteristics such as gender, race, eligibility for Free or Reduced-Price Lunch (FRL), and special education status.

2.4.2 Fuzzy Regression Discontinuity

I also estimate the LATE impact of the program by employing a fuzzy Regression Discontinuity at the maximum cutoff for eligibility. I estimate the following two-stage least squares (2SLS) specification:

$$IE_{it} = \alpha_0^{fs} + \beta_1^{fs} LowWIDA_{it_0} + f\left(WIDA_{it_0}\right) + \boldsymbol{X'}_{it}^{fs} \gamma^{fs} + \varepsilon_{it}^{fs} \tag{2}$$

$$A_{it+1} = \beta_0 + \beta_1 \widehat{IE_{it}} + \boldsymbol{X}_{it+1} \gamma + \varepsilon_{it+1} \tag{3}$$

where, in the first-stage equation given by (2), IE_{it} indicates whether student i participates in the temporary intensive English program in year t; $LowWIDA_{it_0}$ indicates whether student i's WIDA score is below the maximum for program eligibility; $f\left(WIDA_{it_0}\right)$ is a continuous function of student i's WIDA score;[91] $\boldsymbol{X}_{it}^{fs}$ is a vector of student time-varying characteristics; and

[90] The WIDA Screener can be taken in a computer or in paper. Irrespective of the format of the test, the composite score is determined by an algorithm unknown to both students and test administrators. https://wida.wisc.edu/sites/default/files/resource/WIDA-Screener-Interpretive-Guide.pdf
[91] I use a linear function in the preferred specification. I also test the sensitivity of the estimator to higher order polynomials.

ε_{it}^{fs} is the error term. The second-stage equation is given by (3) where A_{it+1} is a normalized achievement measure in ELA or math for student i in the school year post program participation, $t+1$; $\widehat{IE_{it}}$ is the predicted program participation as explained by the initial WIDA score; $\boldsymbol{X}_{it+1}$ is comprised of $\boldsymbol{X}_{it+1}^{fs}$ and $f\left(WIDA_{it_0}\right)$; and ε_{it+1} is the error term.

In my preferred specification, I limit the sample to students with initial WIDA scores that fall within one unit of the cutoff for program eligibility. I choose this bandwidth to maximize the number of observations below the cutoff – recall that nearly one-third of the sample obtains the minimum WIDA score. I also run robustness checks using the optimal bandwidth determined by the procedure described in Calonico et al. (2017). I estimate the main model using a linear polynomial specification, a triangular kernel function, and clustering the standard errors at the level of the running variable, i.e. the WIDA scores (Lee and Card 2008).

The main parameter of interest, β_1, measures the effect of participating in the intensive English program as induced by whether the student's WIDA score is below the maximum cutoff for program eligibility. That is, this parameter measures the treatment effect for students whose score is close to the cutoff. As in equation (1), I measure the outcome variables one-year post program participation, thus estimating the short-term effects of the program.

2.5 Results and Discussion

2.5.1 Intent-to-Treat Difference-in-Differences Results

I begin by examining the impact of program eligibility using the pooled sample of students who are screened for EL classification in grades 1 through 7. Table 2.3 shows the results from estimating variants of equation (1) using normalized ELA and math test scores as the outcomes of interest. Each column shows the point estimates from a different specification where control variables are staggered. The last column shows results from fully specified models that include achievement controls, school and grade fixed effects, and demographic controls.[92] In all specifications, standard errors are clustered at the school level.

I find that ELs who are eligible to participate in an intensive English program have lower ELA achievement relative to ELs who receive traditional ESL instruction. I estimate no impact of program eligibility on math test scores. The point estimates for ELA and math are consistent across all model specifications. Specifically, program eligibility leads to lower ELA scores by up to 0.17 standard deviations. This is a large and significant effect, higher than most negative estimates on the impact of EL classification in Kindergarten on ELA scores (Umansky 2016).[93]

I further investigate whether the impact of program eligibility differs by grade of EL screening. Figure 2.5 shows these results using ELA scores as the outcome. I find that the negative impacts on ELA achievement are driven by students who are eligible for the program in grades 5 and above; I find small and statistically insignificant program eligibility estimates for students screened for EL classification in grades 3 and 4. These results align with previous research documenting that EL classification is negatively associated with EL academic outcomes in middle school (Umansky 2018). I find no impact of math test scores irrespective of grade at EL screening.[94]

2.5.2 Heterogeneous Effects by Refugee Status

The EL student population in the district is remarkably diverse in race/ethnicity and immigration status. For example, refugees account for 52 percent of students eligible for the Intensive English program. Previous research points to differences between immigrants and

[92] The full list of demographic controls includes indicators for female, black, Hispanic, Asian, refugee, FRL eligibility, and special education status.
[93] Umansky (2016) estimates that being marginally classified as an EL in Kindergarten leads to lower ELA achievement in grade 2 by 0.061 standard deviations.
[94] See Figure A2.9 in the appendix. While I find the same pattern across grades, zero effects in early grades and negative point estimates in later grades, most estimates are imprecisely estimated.

refugees that may lead to variation in program eligibility effects (Cortes 2004). Thus, I proceed to investigate whether program eligibility impacts students differently based on their refugee status – to my knowledge, this is the first study to isolate the effects of ESL instruction among refugee ELs. Table 2.4 reports the point estimates from fully specified models limiting the sample to students who are self-identified as refugees. Each column shows the impact on ELA and math scores, respectively.

Unlike the average effects, I find large positive impacts of program eligibility on ELA scores and weak evidence of gains in math achievement among the refugee subsample. Specifically, refugee students who are eligible for an intensive English program have higher ELA test scores by 0.41 standard deviations and higher math test scores by 0.43 standard deviations. Further, as seen in Figure 2.6, I find large and positive effects on ELA achievement across all grades of EL screening, with smaller effects as grade level increases.[95] Overall, the impacts of program eligibility among the refugee subsample are exceptionally large, corresponding to roughly four times the impact of a highly effective teacher (Rivkin et al. 2005; Rockoff 2004). One possible explanation is that refugees who enter the district may arrive with very low levels of literacy, even in their own language, possibly as a result of years without access to formal schooling. Therefore, it is possible that participating in a program aimed at increasing their linguistic skills can have very large effects. An important limitation of using the WIDA score to measure initial English proficiency is that scores at the low-end of the distribution mask important skill variation. A student who scores at the lowest level may either have low English proficiency or low literacy even in their native language. Therefore, while I do

[95] I also find positive point estimates for math effects across all grades. See Figure A2.10 in the appendix.

have measures of initial language skill, I am unable to test whether the positive effects among the refugee subsample stem from low literacy, in addition to low English proficiency.

2.5.3 Robustness Checks and Falsification Tests for the DID Results

It is possible that the pooled DID estimates fail to capture the program eligibility effect and rather estimate a negative correlation between having low English proficiency and being a late arrival (Bleakley and Chin 2004), irrespective of program exposure. I test for this spurious correlation using three approaches. First, I narrow the sample used in the estimation to ELs screened in grades 2 and 3 only, so as to compare the outcomes of students across similar in ages. Second, I limit the sample to students with initial WIDA scores within one unit of the cutoff, thus comparing ELs with similar English proficiency. Third, I run a falsification test where I assume that program eligibility is defined for grades 6 and 7 only, using grades 3 through 5 as the "control" group. Simply put, I estimate the model using a placebo cutoff at grade 6.

I report results from the specifications outlined above in Table 2.5. As shown in columns 1 and 2, limiting the sample to more comparable groups reduces the coefficients substantially, leading to statistically insignificant estimates. Moreover, I estimate a that students in grades 6 and 7 who have low initial English proficiency have lower ELA scores by 0.2 standard deviations relative to younger ELs. Because there is no change in program eligibility across these grades, this coefficient simply estimates the correlation between ELA scores and being an older EL. Note the remarkable similarity between these estimates and those from the baseline DID specification. In fact, I cannot reject the null hypothesis of equality between these two effects. Altogether, these tests suggest that the main DID results are strongly driven by differences in language acquisition across ages, irrespective of program exposure.

Notably, results for the refugee subsample are robust to these tests. As seen in Table 2.6, I estimate large gains in ELA achievement even when I limit the sample to students across similar ages or comparable initial English proficiency.[96] As with the pooled sample, results from the falsification test suggest a negative correlation between achievement and age of EL screening. Thus, having a positive finding in the baseline model suggest that the program eligibility effect for this subgroup is large enough to overcome the underlying negative relationship between achievement and grade of EL screening.

2.5.4 Fuzzy Regression Discontinuity Results

Next, I estimate the effect of program enrollment using a fuzzy Regression Discontinuity specification. It improves on the previous approach by comparing the outcomes of students who are screened in the same grade, but face different probabilities of program enrollment as induced by small differences in their initial WIDA score. For this part of the analysis, I narrow the sample to the school years 2015-2018, as these are the years with high program compliance.

I estimate a large discontinuous jump in the probability of program enrollment for students who score within one unit above the cutoff. In particular, having a score below the maximum for program participation increases the likelihood of program enrollment by 94 percentage points. Figure 2.7 shows the first stage plot associated with the fuzzy RD specification. It is worth noting that, although there is a large jump at the cutoff, there is no perfect compliance as students are allowed to opt-out with the approval of parents or legal guardians.

Tables 2.7 and 2.8 provide estimates from the 2SLS specification defined by equations (2) and (3) using ELA and math test scores one year after program eligibility as the main

[96] I find similar effects for math test scores. See Table A2.1 in the appendix.

outcomes, respectively. I limit the sample to students with WIDA scores within one unit of the cutoff and cluster the standard errors at the level of the running variable. Overall, I estimate negative local average treatment effects across both subjects. Results from the 2SLS specification without controls indicate that ELs who enroll in the Intensive English program, as induced by their initial WIDA score, have lower ELA achievement by 0.51 standard deviations, relative to ELs who score just above the cutoff and receive traditional ESL support. These results are qualitatively robust to using the optimal bandwidth derived from the procedure outlined in Calonico et al. (2017).[97] It is worth noting, however, that with the inclusion of grade fixed effects and demographic controls, the point estimates from the pooled sample become smaller and statistically insignificant.

Additional specifications by grade of EL screening indicate that the negative effects in ELA scores are driven by older students, specifically ELs screened in grades 4 through 6.[98] While the overall RD estimates are much larger, and sensitive to controlling for grade level and demographic characteristics, they align with the pooled DID effects qualitatively – neither indicate gains in ELA achievement as a result of eligibility for or enrollment in the Intensive English program.

I find different results for math, however. Unlike the baseline DID specification, I estimate large negative effects in math achievement for students who enroll in the program. Specifically, ELs who participate in the Intensive English program have lower math achievement in the short-term by up to 0.48 standard deviations. These results are robust to controlling for grade of EL screening and demographic characteristics. Further analyses do not provide evidence

[97] See Column (4) in Table 2.7.
[98] ELA fuzzy RD results by grade are shown in Table A2.2 in the appendix.

that these effects are driven by particular grades. Rather, I estimate negative, yet statistically insignificant effects when I estimate the model separately by grade of EL screening.[99]

2.5.5 Sensitivity and Validity Checks for the Fuzzy Regression Discontinuity Design

A key model assumption in this empirical approach is that, around a local neighborhood near the cutoff, treatment and control units are similar in their observable and unobservable characteristics, such that the likelihood of treatment varies across groups as if as random (Lee and Lemieux 2010). While it is not possible to fully test this assumption, one important falsification test involves checking whether predetermined covariates vary systematically around the cutoff. To do this, I estimate variants of equation (2) where I replace the outcome with indicators for predetermined demographic characteristics and program participation variables, such as FRL eligibility. As seen in Figure 2.8, I estimate small and statistically insignificant discontinuities in the likelihood of most predetermined characteristics. The only exception is the race indicator for White where I estimate an increase in the probability of belonging to this group for ELs with WIDA scores within one unit below the maximum cutoff.

The second type of falsification test examines whether there is systematic manipulation of the running variable within a neighborhood around the cutoff. I check for the possibility of precise manipulation following the procedure described in McCrary (2008). Based on the estimates from the McCrary test, I fail to reject the null hypothesis of no difference in the density of the WIDA score across treatment and control groups.[100] Figure 2.9 presents graphical evidence of the overlapping densities across the cutoff with a 95 percent confidence interval.

Lastly, I check the sensitivity of the point estimates with respect to the length of the bandwidth. Specifically, I run variants of the baseline 2SLS specification without covariates

[99] Math fuzzy RD results by grade are shown in Table A.3 in the appendix.
[100] Specifically, I estimate a t-statistic of 1.26 with an associated p-value of 0.21.

where I vary the bandwidth in 0.1 increments. Table 2.9 shows the results from this exercise using ELA and math test scores as the outcome variables. Notably, all the point estimates for math are negative and statistically significant, with their magnitude decreasing as the length of the bandwidth increases. On the other hand, the point estimates for ELA are sensitive to the choice of bandwidth.

2.5.6 Discussion of ITT and LATE Treatment Effects

By design, the DID and fuzzy RD analyses use different samples to estimate the effect of program eligibility and enrollment; thus, they estimate different parameters. The former uses observations on all students for whom the program is intended, and it does not account for program take-up. The second uses only observations from students with WIDA scores within one unit of the threshold and estimates the treatment on the treated within this local neighborhood. Therefore, having results from both analysis enables me to compare the effects of the intensive English program across different types of students.

I estimate negative treatment effects on ELA scores across both specifications. However, the results from the DID analyses are smaller in magnitude, indicating that the effects for the overall sample of eligible ELs are smaller, relative to the impact on ELs for whom the WIDA cutoff is binding. One possible explanation for this difference arises from the potential unintended consequences of the Intensive English program.[101] In addition to access to specialized ESL instruction, students who enroll in the program also delay core-content classes for up to one year. Moreover, they have limited access to general education resources that are only available in their neighborhood school.[102] My results are consistent with the hypothesis that

[101] Umansky (2018) finds that being marginally classified as an EL triggers unintended consequences such as lower exposure to English-fluent peers and limited access to academic resources in middle school.
[102] Setren (2019) finds that non-targeted education interventions can have large positive effects among ELs, especially those with low initial English proficiency.

delaying core content instruction has a larger effect on ELs at the threshold of basic English literacy.

If delayed core-content instruction is a mechanism behind the negative program enrollment effects, ELs should also observe a decrease in math achievement – perhaps even a greater negative effect. While I do not estimate any impact of program eligibility on math test scores, I find large and significant negative effects of program enrollment from the fuzzy RD analysis. This result supports the hypothesis that part of the treatment effect operates through unintended consequences, such as delayed instruction.

2.6 Conclusion

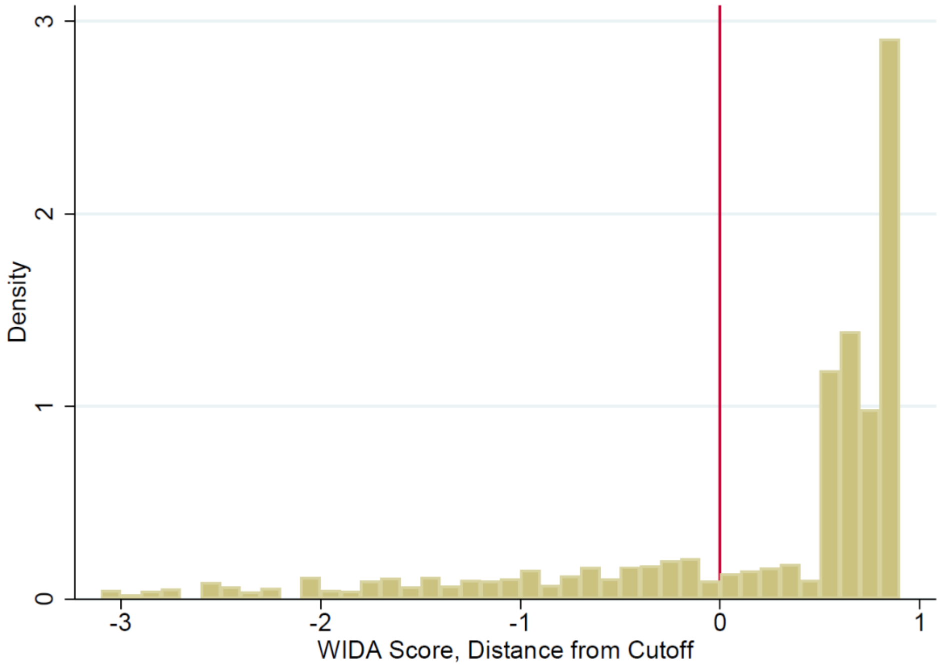
I estimate the impact of an intensive English program aimed at ELs with very low English proficiency. Leveraging program eligibility criteria, I estimate intent-to-treat effects using a Difference-in-Differences approach and local average treatment effects using a fuzzy Regression Discontinuity design. I focus on the short-term impact of the program by estimating the effect on ELA and math test scores. I access individual-level data on the universe of ELs enrolled in public schools in one of the largest schools in Georgia with the second highest enrollment of ELs in the state.

On average, I find that students who are eligible for the Intensive English program have lower ELA test scores one year after program eligibility, relative to ELs who receive traditional ESL support. The negative ELA effects are primarily driven by students who are screened for EL classification in grades 5 and above. However, using the subsample of refugee ELs, I find large and positive program eligibility effects in both ELA and Math.

In line with the intent-to-treat results, estimates from the RD specification indicate that students who enroll in the program, as induced by their initial English proficiency score, have lower ELA achievement. The LATE effects are larger, indicating that the negative treatment effects intensify among the subset of ELs at the threshold of basic English literacy. I also find negative impacts of program enrollment on math achievement.

2.7 Figures and Tables

Figure 2.1: Distribution of Initial WIDA Scores, Grades 1-8, School Years 2008-2018



Note: WIDA scores are rounded to one decimal place. The histogram plots the transformed WIDA scores after centering them at the cutoff and multiplying them by -1, so that all observations above zero are eligible for the program.

Figure 2.2: Number of Students taking the WIDA Screener by Grade Level in School Years 2008-2018

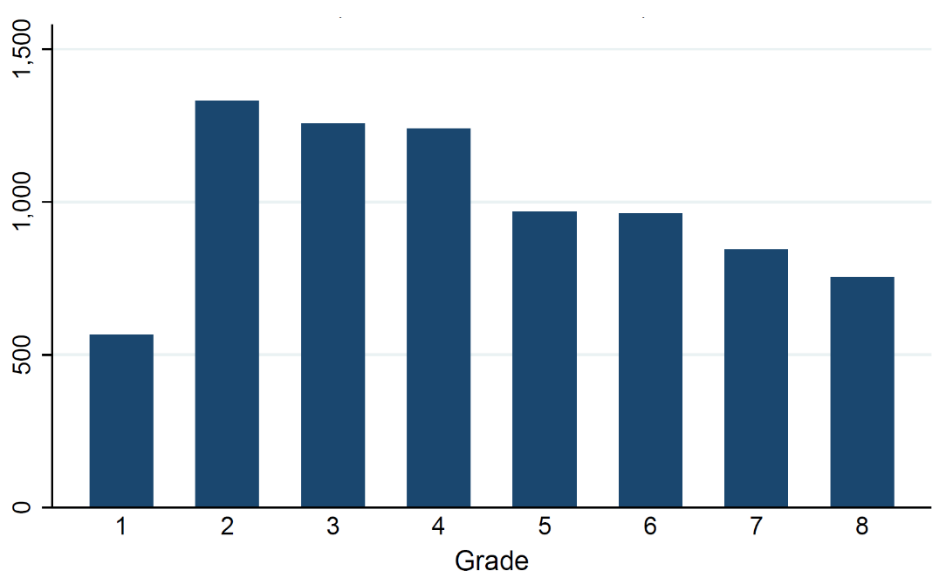


Figure 2.3: Number of Students Eligible and Enrolled in the Intensive English Program by School Year, Grade Levels 3-8

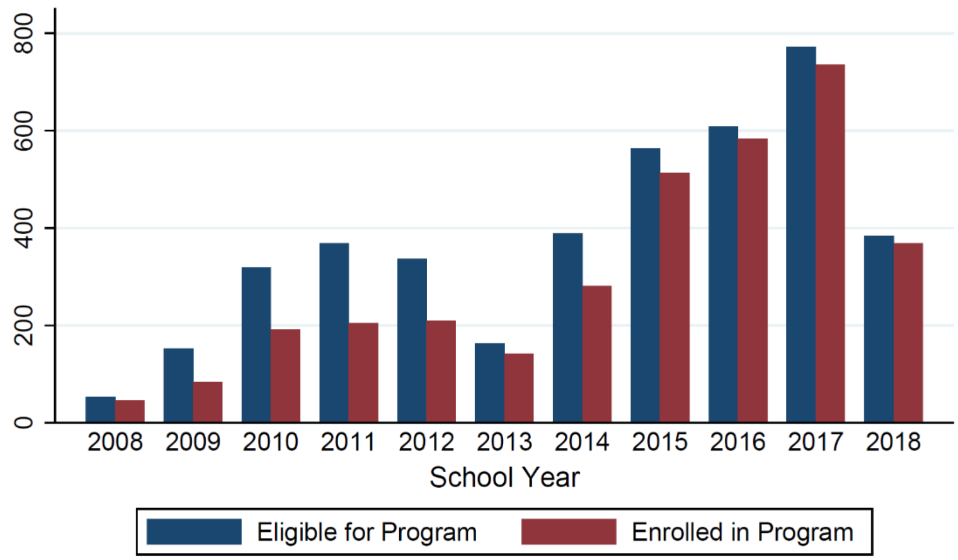


Figure 2.4: Number of Students taking the WIDA Screener by Month, Grades 1-8, and School Years 2008-2018

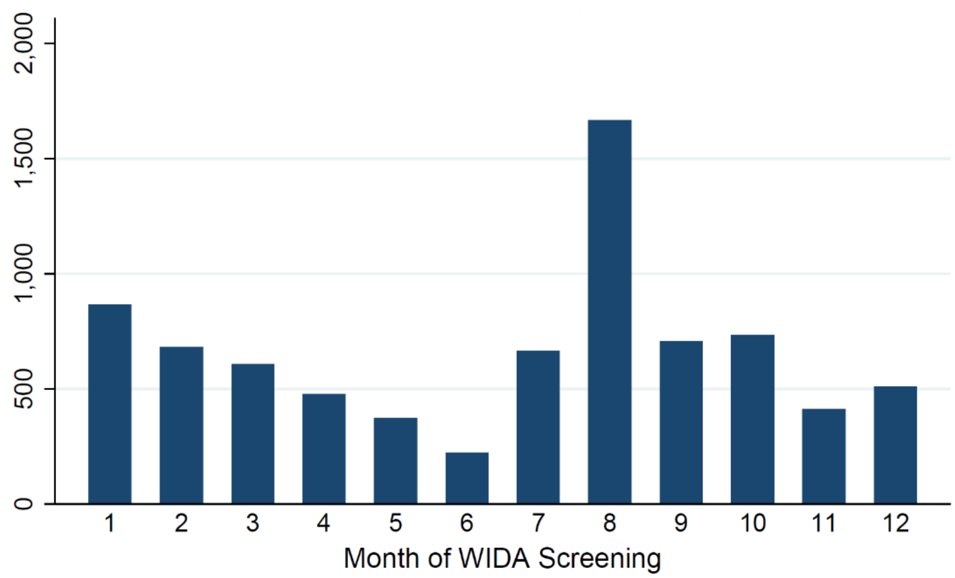
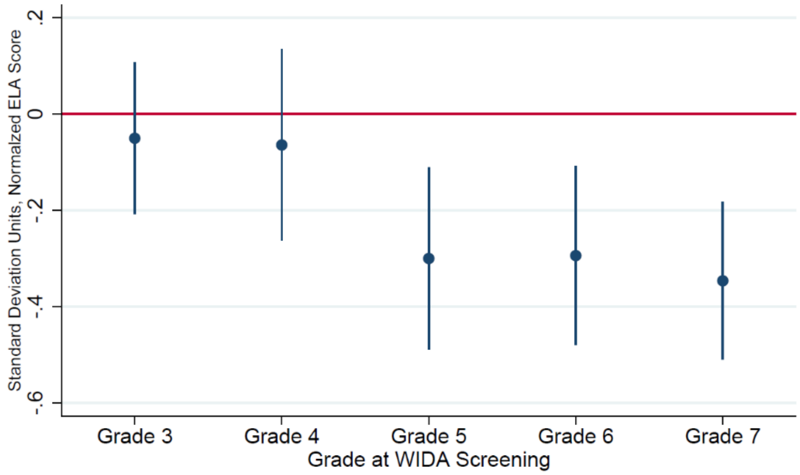
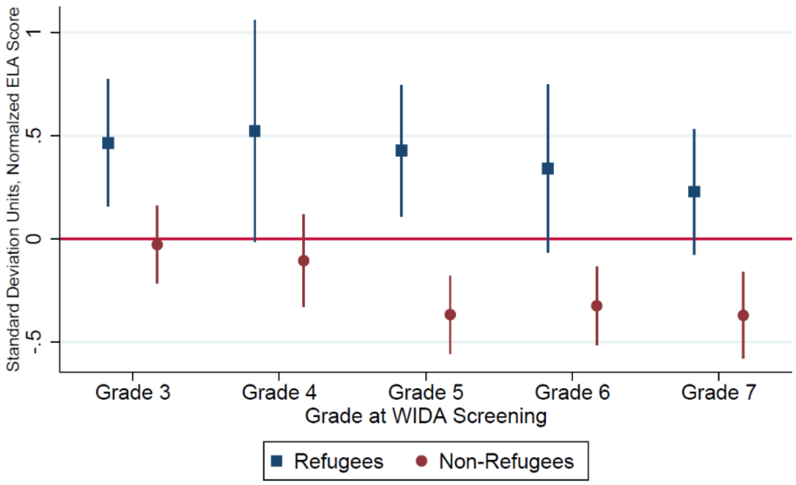


Figure 2.5: Difference-in-Differences Results: Estimated Program Eligibility Effect on ELA Test Scores by Grade of EL Screening



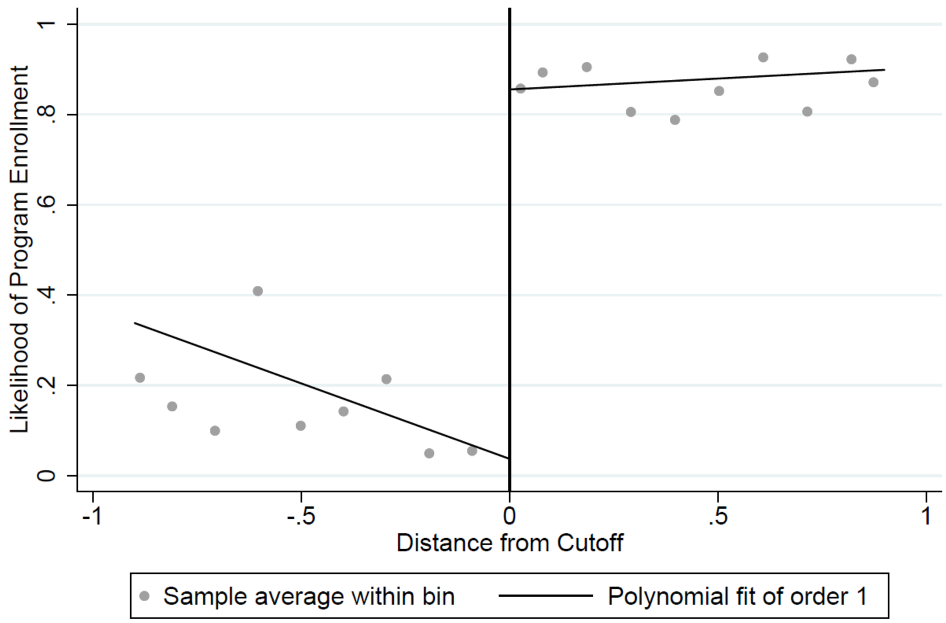
Note: Data comprise students who take the WIDA exam in grades 1-7 and score low enough to be eligible for EL classification. Test scores are normalized to mean zero and standard deviation one with respect to the state-grade-subject distribution. Test scores correspond to the first post-WIDA grade. Estimates are from a fully specified models including achievement controls, school FEs, grade FEs, and demographic controls (indicators for female, black, Hispanic, Asian, refugee, FRL eligibility, and special education status), where the treatment indicator is interacted with indicators by grade of EL screening. The bars represent confidence intervals at the 95% level.

Figure 2.6: Difference-in-Differences Results: Estimated Program Eligibility Effect on ELA Test Scores by Grade of EL Screening and Refugee Status



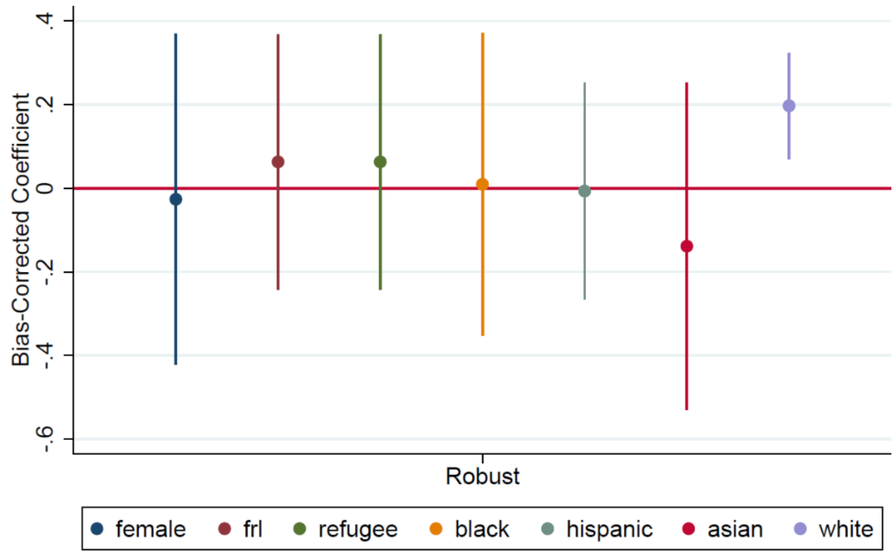
Note: Refugee status is based on a self-reported variable. Data comprise students who take the WIDA exam in grades 1-7 and score low enough to be eligible for EL classification. Test scores are normalized to mean zero and standard deviation one with respect to the state-grade-subject distribution. Test scores correspond to the first post-WIDA grade. Estimates are from fully specified models including achievement controls, school FEs, grade FEs, and demographic controls (indicators for female, black, Hispanic, Asian, FRL eligibility, and special education status). Bar represent confidence intervals at the 95% level.

Figure 2.7: Probability of Program Enrollment Around a 1-Unit Bandwidth from the WIDA Cutoff



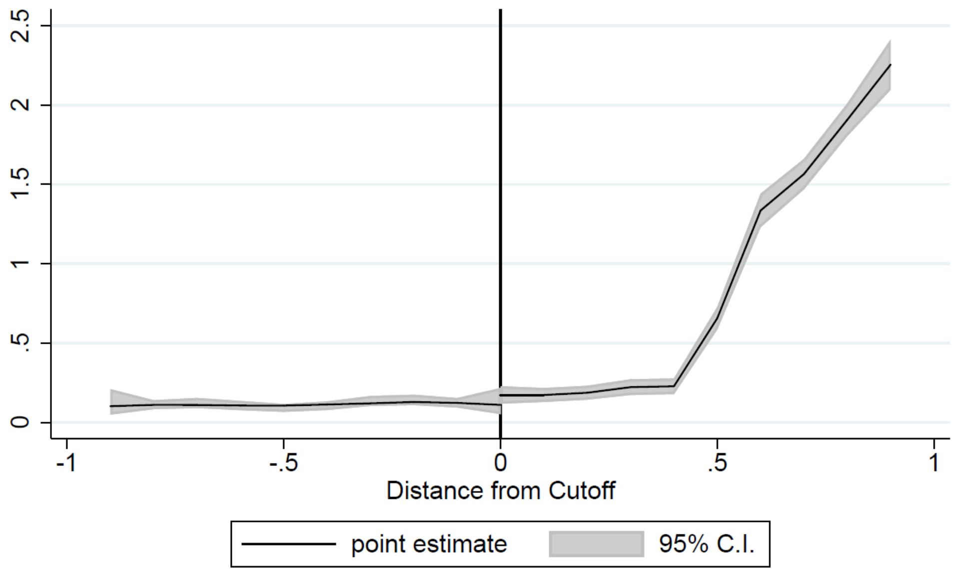
Note: Sample includes students who take the WIDA exam in grades 3-7 and score low enough to be eligible for EL classification. Data is limited to observations from school years 2015-2018. The WIDA score has been transformed such that zero indicates the maximum threshold for program eligibility. Observations above zero are eligible for the program.

Figure 2.8: Falsification Test for Variation in Predetermined Covariates Around the Cutoff



Note: Each point estimate corresponds to a separate regression using each demographic variable as the outcome and the WIDA score as the dependent variable. In each regression, the sample includes students who take the WIDA exam in grades 3-7 and score low enough to be eligible for EL classification. Data is limited to observations from school years 2015-2018. The bars represent confidence intervals at the 95% level

Figure 2.9: McCrary Plot of the Running Variable



Note: Observations from the school years 2015-2018 and limited to those within one unit of the cutoff. The WIDA score has been transformed such that zero indicates the maximum threshold for program eligibility. Observations above zero are eligible for the program.

Table 2.1: Summary Statistics by WIDA Cutoff and Eligible Grades, Grades 1-8, School Years 2008-2018 (DID Sample)

	Eligible Grades (Grades 3-8)				Non-Eligible Grades (Grades 1-2)			
	Low WIDA		High WIDA		Low WIDA		High WIDA	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
WIDA Proficiency Score	1.18	0.21	3.01	0.82	1.15	0.19	3.14	0.83
WIDA Re-taker	0.40	0.49	0.02	0.13	0.04	0.19	0.00	0.07
Aug WIDA	0.19	0.39	0.23	0.42	0.18	0.38	0.33	0.47
Aug/Sep WIDA	0.28	0.45	0.32	0.47	0.26	0.44	0.42	0.49
Fall WIDA	0.64	0.48	0.62	0.49	0.67	0.47	0.69	0.46
English score, normalized	-1.61	0.77	-0.61	0.96	-1.07	0.83	0.00	0.86
Math score, normalized	-1.07	0.78	-0.30	0.97	-0.87	0.88	0.02	1.00
Female	0.47	0.50	0.49	0.50	0.42	0.49	0.46	0.50
Black	0.23	0.42	0.26	0.44	0.27	0.45	0.20	0.40
White	0.14	0.35	0.14	0.34	0.10	0.30	0.21	0.40
Hispanic	0.30	0.46	0.27	0.44	0.20	0.40	0.33	0.47
Asian	0.40	0.49	0.45	0.50	0.46	0.50	0.42	0.49
Refugee	0.52	0.50	0.26	0.44	0.55	0.50	0.08	0.28
FRL Eligible	0.86	0.34	0.77	0.42	0.78	0.41	0.61	0.49
ESOL	0.99	0.12	0.95	0.22	0.96	0.19	0.93	0.25
EL Ever	1.00	0.02	0.98	0.12	1.00	0.07	0.97	0.16
Gifted	0.00	0.02	0.01	0.08	0.00	0.00	0.02	0.13
Special Ed	0.00	0.07	0.01	0.11	0.01	0.09	0.02	0.15
Days Present	113.50	50.98	124.01	50.68	114.87	51.49	127.30	51.95
Days Absent	4.62	5.39	4.35	5.32	4.85	5.13	4.13	4.37
Enrollment Spells	1.14	0.38	1.06	0.25	1.05	0.24	1.02	0.13
Schools Enrolled	1.04	0.20	1.04	0.20	1.04	0.21	1.00	0.07
Observations	2498		981		666		224	

Note: Data comprise students who take the WIDA exam in grades 1-7 and score low enough to be eligible for EL classification. Test score observations correspond to the first post-WIDA grade. All other variables are measured at the time of WIDA screening – i.e. baseline characteristics. DID sample.

Table 2.2: Summary Statistics by WIDA Cutoff, Grades 3-8, School Years 2015-2018 (Fuzzy RD Sample)

	Eligible		Ineligible	
	Mean	SD	Mean	SD
WIDA Proficiency Score	1.20	0.21	2.40	0.27
WIDA Re-taker	0.34	0.47	0.02	0.13
Aug WIDA	0.20	0.40	0.28	0.45
Fall WIDA	0.69	0.46	0.68	0.47
Female	0.48	0.50	0.49	0.50
Black	0.28	0.45	0.31	0.46
White	0.11	0.31	0.07	0.25
Hispanic	0.39	0.49	0.22	0.42
Asian	0.26	0.44	0.46	0.50
Refugee	0.45	0.50	0.29	0.45
FRL Eligible	0.89	0.31	0.78	0.42
ESOL	0.97	0.16	0.90	0.30
Ever EL	1.00	0.00	1.00	0.00
Gifted	0.00	0.00	0.00	0.00
Special Education	0.01	0.10	0.01	0.11
Total Days Present	143.71	42.69	146.50	42.19
Total Days Absent	6.80	7.16	6.84	8.60
Number of Enrollment Spells	1.18	0.41	1.05	0.23
Number of Schools Attended	1.07	0.26	1.04	0.19
Observations	1237		166	

Note: Data comprise students who take the WIDA exam in grades 1-7 and score low enough to be eligible for EL classification. Sample is limited to observations used in the fuzzy RD specification during high-compliance years, 2015-2018. Limited to students with WIDA scores within +-0.9 units from the program eligibility cutoff. Variables are measured at the time of WIDA screening – i.e. baseline characteristics.

Table 2.3: Difference-in-Differences Results: ELA and Math Achievement One Year Post Treatment Eligibility

Variables	Outcome: Normalized ELA Test Scores			Outcome: Normalized Math Test Scores		
Low WIDA	0.10	0.11	0.15**	-0.00	-0.00	0.04
	(0.07)	(0.07)	(0.07)	(0.09)	(0.09)	(0.09)
Eligible Grade	-0.28***	0.11	0.11	-0.10	0.70**	0.79**
	(0.07)	(0.23)	(0.23)	(0.07)	(0.30)	(0.31)
Low WIDA X Eligible Grade	-0.16*	-0.17*	-0.17**	-0.08	-0.07	-0.07
	(0.09)	(0.09)	(0.08)	(0.10)	(0.10)	(0.09)
WIDA Score	0.43***	0.44***	0.42***	0.31***	0.31***	0.28***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)
Achievement controls	X	X	X	X	X	X
School FE	X	X	X	X	X	X
Grade FE		X	X		X	X
Demographic Controls			X			X
Observations	4294	4294	4294	4469	4469	4469

Note: Data comprise students who take the WIDA exam in grades 1-7 and score low enough to be eligible for EL classification. Test scores are normalized to mean zero and standard deviation one with respect to the state-grade-subject distribution. Test scores correspond to the first post-WIDA grade. Demographic controls include indicators for female, black, Hispanic, Asian, refugee, FRL eligibility, and special education status. Grade FEs correspond to the grade level one-year post program eligibility. Clustered standard errors at the school level are shown in parenthesis.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.4: Difference-in-Differences Results: ELA and Math Achievement One Year Post Treatment Eligibility, Refugee Subsample

Variables	Outcome: Normalized ELA Test Scores	Outcome: Normalized Math Test Scores
Low WIDA	-0.20	-0.39
	(0.16)	(0.26)
Eligible Grade	-0.31	0.29
	(0.32)	(0.59)
Low WIDA X Eligible Grade	0.41^{**}	0.43^{*}
	(0.16)	(0.25)
WIDA Score	0.64^{***}	0.44^{***}
	(0.06)	(0.08)
Achievement controls	X	X
School FE	X	X
Grade FE	X	X
Demographic Controls	X	X
Observations	1915	2015

Note: Data comprise refugee students who take the WIDA exam in grades 1-7 and score low enough to be eligible for EL classification. Test scores are normalized to mean zero and standard deviation one with respect to the state-grade-subject distribution. Test scores correspond to the first post-WIDA grade. Demographic controls include indicators for female, black, Hispanic, Asian, FRL eligibility, and special education status. Grade FEs correspond to the grade level one-year post program eligibility. Clustered standard errors at the school level are shown in parenthesis.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.5: Difference-in-Differences Results – Robustness Checks and Falsification Test: ELA Achievement One Year Post Treatment Eligibility

	Robustness Checks		Falsification Test
	(1)	(2)	(3)
Low WIDA Score	0.10	0.28^{***}	0.07
	(0.10)	(0.09)	(0.08)
Eligible Grade	-0.32^{***}	0.23	0.09
	(0.08)	(0.34)	(0.26)
Low WIDA X Eligible Grade	-0.04	-0.05	-0.20^{***}
	(0.08)	(0.10)	(0.06)
WIDA Scale Score	0.39^{***}	0.64^{***}	0.43^{***}
	(0.06)	(0.06)	(0.04)
Achievement controls	X	X	X
School FE	X	X	X
Grade FE	X	X	X
Demographic Controls	X	X	X
Observations	1646	3776	3402

Note: Column (1) limits the sample to students who take the WIDA in grades 2 and 3 only and score low enough to be eligible for EL classification. Column (2) limits the sample to students who take the WIDA in grades 1-7 with WIDA scores within 1 unit from the cutoff. Column (3) limits the sample to students who take the WIDA in grades 3-7 and score low enough to be eligible for EL classification. Treatment is falsely assumed to apply only to grades 6 and 7. In all specifications, test scores are normalized to mean zero and standard deviation one with respect to the state-grade-subject distribution. Test scores correspond to the first post-WIDA grade. Demographic controls include indicators for female, black, Hispanic, Asian, refugee, FRL eligibility, and special education status. Grade FEs correspond to the grade level one-year post program eligibility. Clustered standard errors at the school level are shown in parenthesis.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.6: Difference-in-Differences Results – Robustness Checks and Falsification Test: ELA Achievement One Year Post Treatment Eligibility, Refugee Subsample

	Robustness Checks		Falsification Test
	(1)	(2)	(3)
Low WIDA Score	0.00	-0.03	0.23*
	(0.22)	(0.19)	(0.12)
Eligible Grade	-0.82***	-0.31	0.42
	(0.13)	(0.33)	(0.27)
Low WIDA X Eligible Grade	0.44***	0.38*	-0.20**
	(0.15)	(0.21)	(0.09)
WIDA Scale Score	0.77***	0.78***	0.61***
	(0.08)	(0.07)	(0.06)
Achievement controls	X	X	X
School FE	X	X	X
Grade FE	X	X	X
Demographic Controls	X	X	X
Observations	738	1857	1523

Note: Column (1) limits the sample to refugee students who take the WIDA in grades 2 and 3 only and score low enough to be eligible for EL classification. Column (2) limits the sample to refugee students who take the WIDA in grades 1-7 with WIDA scores within 1 unit from the cutoff. Column (3) limits the sample to refugee students who take the WIDA in grades 3-7 and score low enough to be eligible for EL classification. Treatment is falsely assumed to apply only to grades 6 and 7. In all specifications, test scores are normalized to mean zero and standard deviation one with respect to the state-grade-subject distribution. Test scores correspond to the first post-WIDA grade. Demographic controls include indicators for female, black, Hispanic, Asian, FRL eligibility, and special education status. Grade FEs correspond to the grade level one-year post program eligibility. Clustered standard errors at the school level are shown in parenthesis.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.7: Fuzzy Regression Discontinuity Results: ELA Achievement One Year Post Treatment Eligibility

	Outcome: Normalized ELA Test Scores			
	(1)	(2)	(3)	(4)
First Stage				
WIDA Scale Score	0.94***	0.95***	0.95***	0.86***
	(0.07)	(0.08)	(0.08)	(0.11)
Second Stage				
Enrollment	-0.51**	-0.25	-0.39	-0.63**
	(0.24)	(0.28)	(0.28)	(0.30)
Grade of WIDA FE		X	X	
Demographic Controls			X	
Data-driven optimal BW				X
Eff. N	1352	1352	1352	212

Note: Sample includes students screened who take the WIDA exams in grades 3-7, score low enough to eligible for EL classification. Columns (1)-(3) report estimates using the sample of students with an initial WIDA score within 1 unit of the cutoff for program participation. Column (4) uses the optimal, data-driven bandwidth to determine the number of observations in the sample. Test scores are normalized to mean zero and standard deviation one with respect to the state-grade-subject distribution, and they correspond to the first post-WIDA grade. Robust bias-corrected standard errors clustered at the WIDA score level are shown in parenthesis. All specifications use a linear polynomial function and triangular kernel. Point estimates correspond to the bias-corrected coefficients.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.8: Fuzzy Regression Discontinuity Results: Math Achievement One Year Post Treatment Eligibility

	Outcome: Normalized Math Test Scores		
	(1)	(2)	(3)
First Stage			
WIDA Scale Score	0.94***	0.96***	0.96***
	(0.07)	(0.07)	(0.07)
Second Stage			
Enrollment	-0.37***	-0.30***	-0.48***
	(0.12)	(0.11)	(0.12)
Grade of WIDA FE		X	X
Demographic Controls			X
Eff. N	1413	1413	1413

Note: Sample includes students screened who take the WIDA exams in grades 3-7, score low enough to eligible for EL classification, and have an initial WIDA score within 1 unit of the cutoff for program participation. Test scores are normalized to mean zero and standard deviation one with respect to the state-grade-subject distribution, and they correspond to the first post-WIDA grade. Robust bias-corrected standard errors clustered at the WIDA score level are shown in parenthesis. All specifications use a linear polynomial function and triangular kernel. Point estimates correspond to the bias-corrected coefficients.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

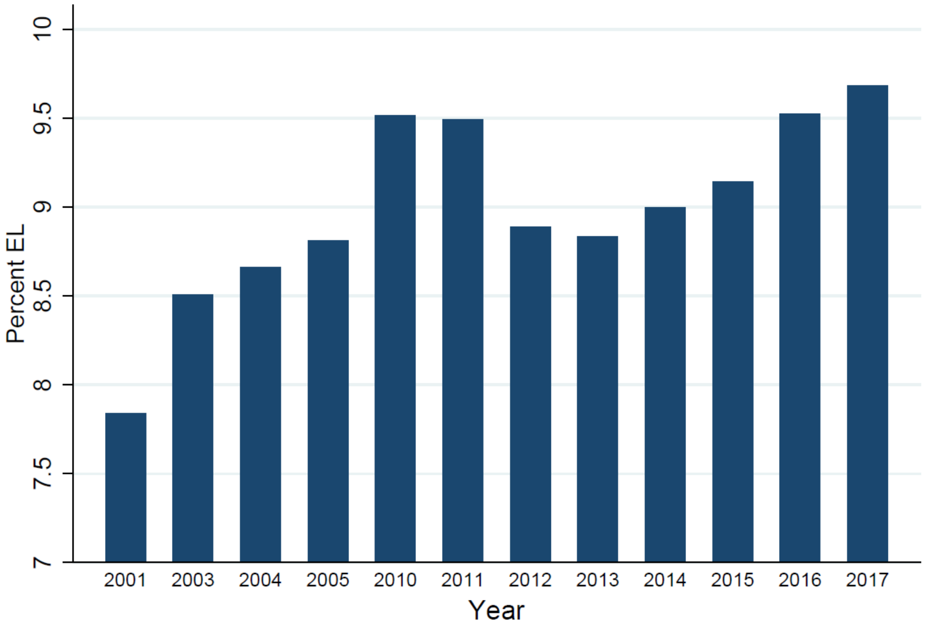
Table 2.9: Fuzzy Regression Discontinuity Results: Sensitivity Analysis Varying the Choice of Bandwidth

	Outcome: Normalized ELA Test Scores					Outcome: Normalized Math Test Scores				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Bandwidth length	0.4	0.5	0.6	0.7	0.8	0.4	0.5	0.6	0.7	0.8
First Stage										
WIDA Scale Score	0.60***	0.86***	0.93***	0.87***	0.91***	0.62***	0.88***	0.95***	0.89***	0.92***
	(0.00)	(0.11)	(0.09)	(0.07)	(0.07)	(0.00)	(0.11)	(0.08)	(0.07)	(0.07)
Second Stage										
Enrollment	0.26***	-0.47	-0.26	-0.35*	-0.52**	-0.61***	-0.62***	-0.37***	-0.48***	-0.40***
	(0.06)	(0.34)	(0.23)	(0.20)	(0.23)	(0.02)	(0.06)	(0.12)	(0.11)	(0.13)
Eff. N	167	214	248	712	712	172	219	255	741	741

Note: Sample includes students screened who take the WIDA exams in grades 3-7, score low enough to eligible for EL classification. Each column shows results from a different specification where the bandwidth increases in 0.1 increments. Test scores are normalized to mean zero and standard deviation one with respect to the state-grade-subject distribution, and they correspond to the first post-WIDA grade. Robust bias-corrected standard errors clustered at the WIDA score level are shown in parenthesis. All specifications use a linear polynomial function and triangular kernel. Point estimates correspond to the bias-corrected coefficients.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

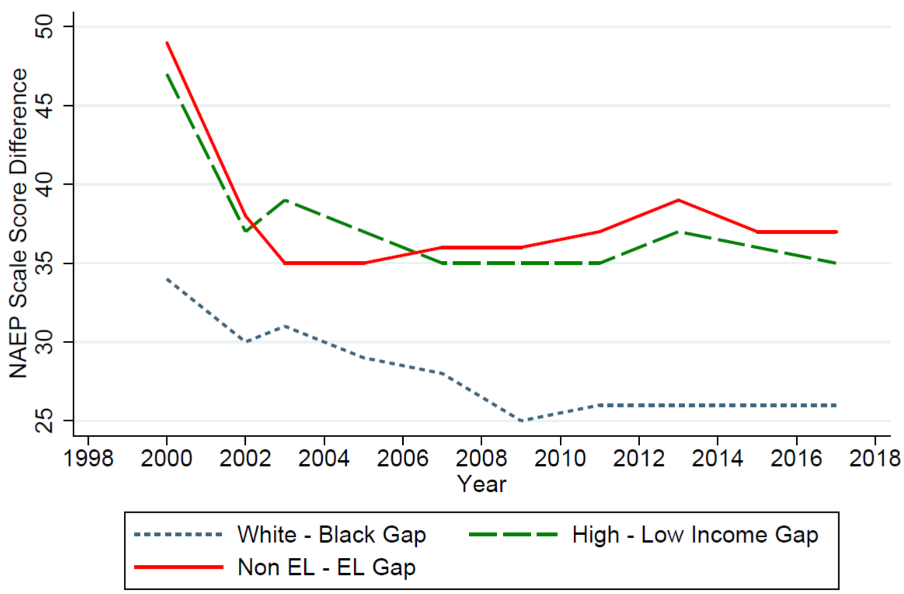
Appendix A2: Additional Figures and Tables

Figure A2.1: Percentage of ELs out of Total Enrollment in U.S. Public Schools, 2001-2017



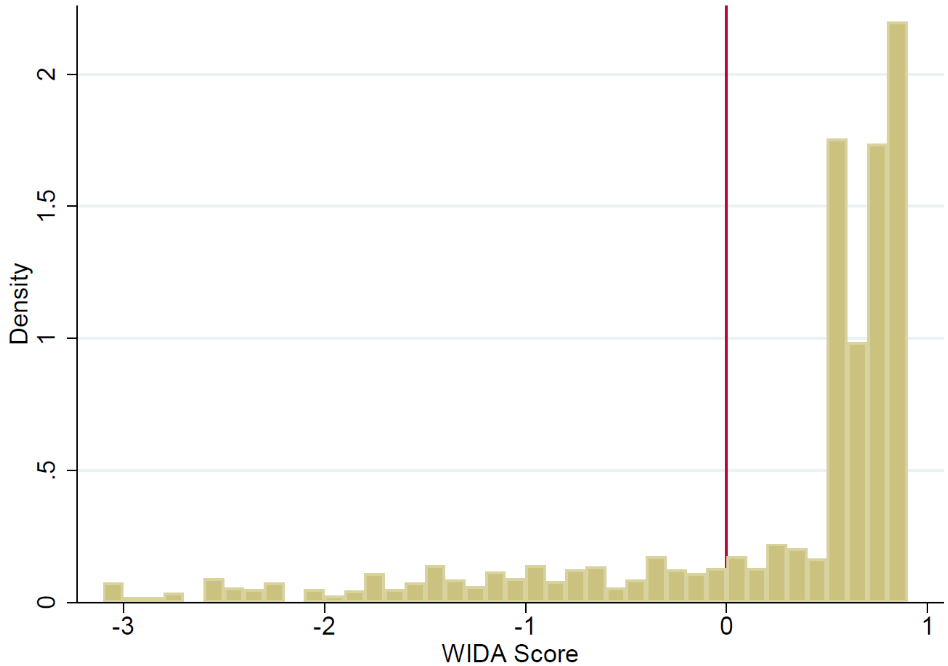
Note: Data obtained from the National Center for Education Statistics

Figure A2.2: Reading Achievement Gaps by Student Subgroups, NEAP Scores Grade 4, 2000-2017



Note: NEAP achievement data by subgroups obtained from the National Center for Education Statistics

Figure A2.3: Distribution of Initial WIDA Scores, Grades 3-8, School Years 2015-2018



Note: WIDA scores are rounded to one decimal place. The histogram plots the transformed WIDA scores after centering them at the cutoff and multiplying them by -1, so that all observations above zero are eligible for the program.

Figure A2.4: Number of Students taking the WIDA Screener by Grade and Refugee Status in School Years 2008-2018

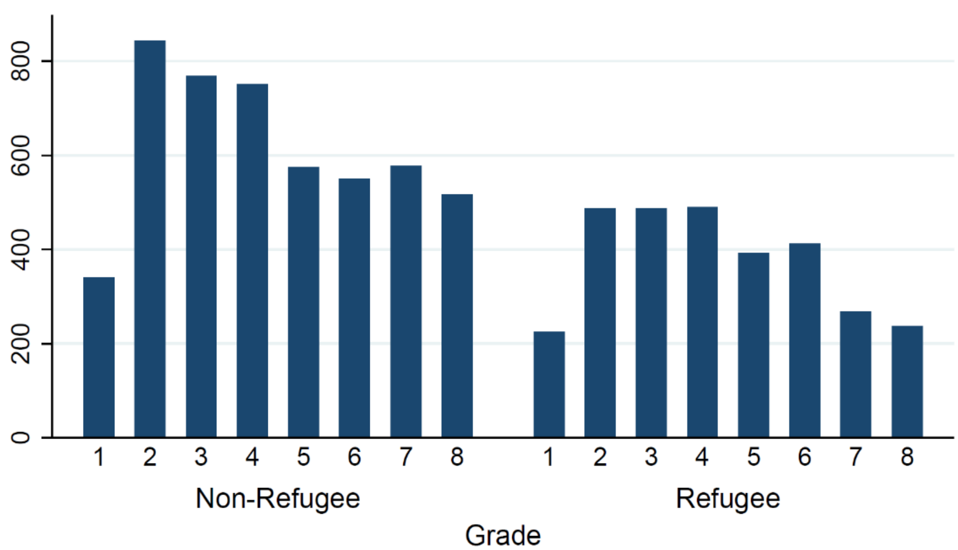


Figure A2.5: Number of Students Taking the WIDA Screener by School Year Across Grades 1-8

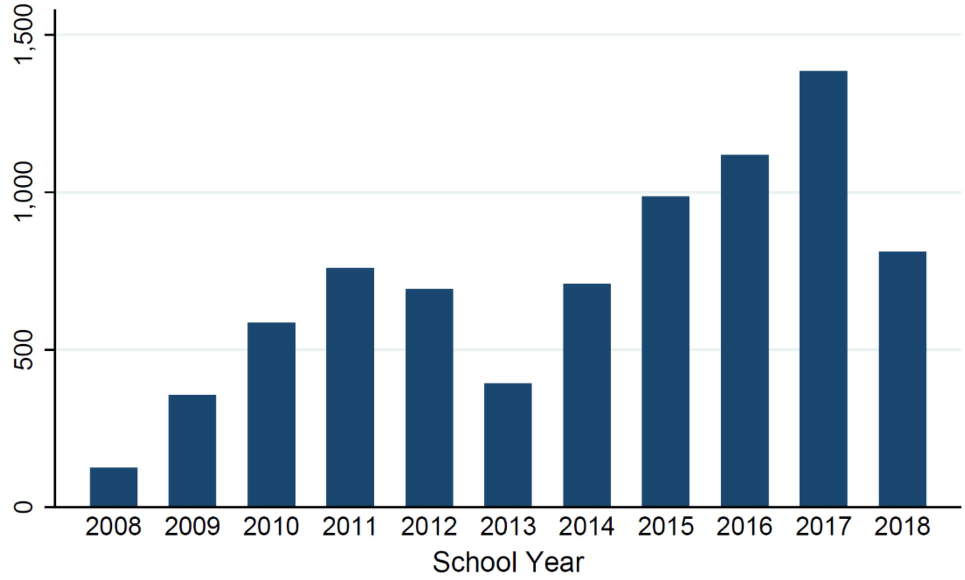
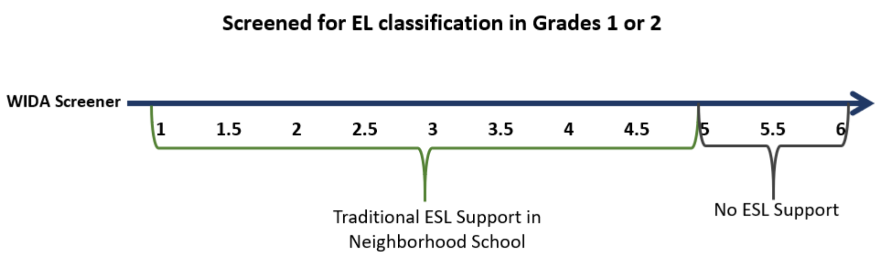


Figure A2.6: Summary of Program Eligibility Criteria by Grade at EL Screening



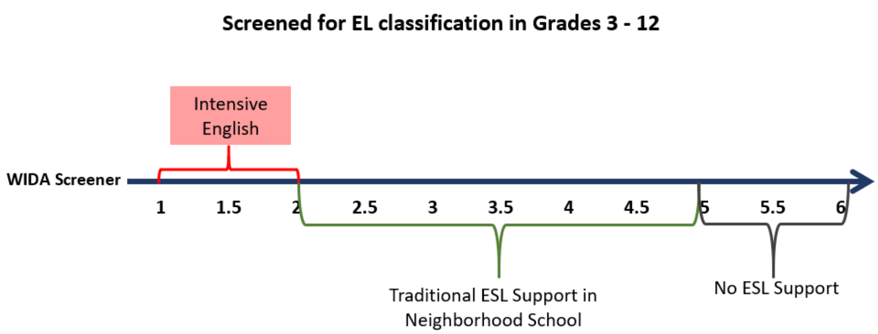
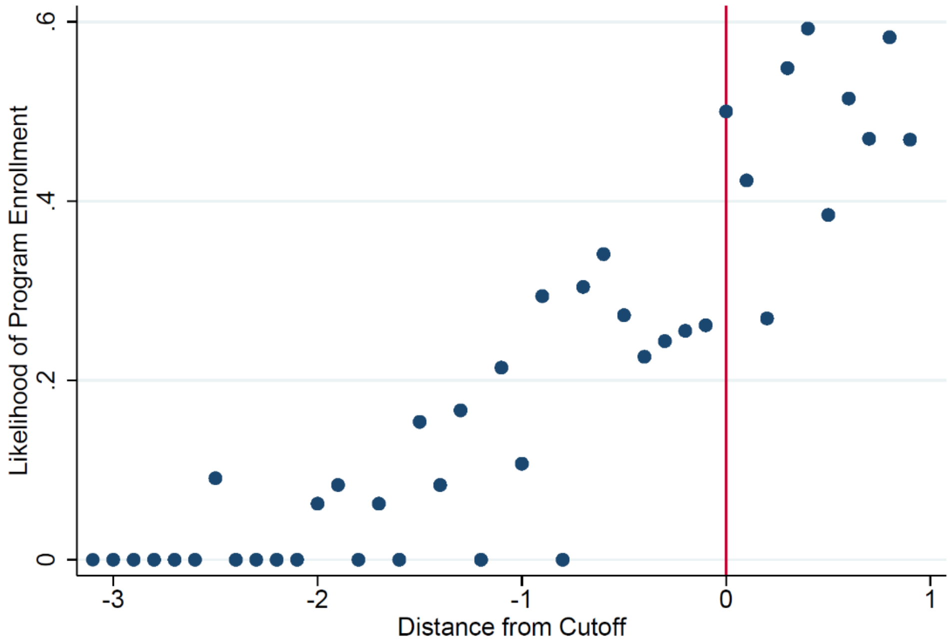


Figure A2.7: Fuzzy RD First Stage Scatter Plot for Low-Compliance Years (2008-2014)



Note: WIDA scores are rounded to one decimal place. The figure plots the transformed WIDA scores after centering them at the cutoff and multiplying them by -1, so that all observations above zero are eligible for the program.

Figure A2.8: Number of Students taking the WIDA Screener by Month and Refugee Status, Grades 1-8, and School Years 2008-2018

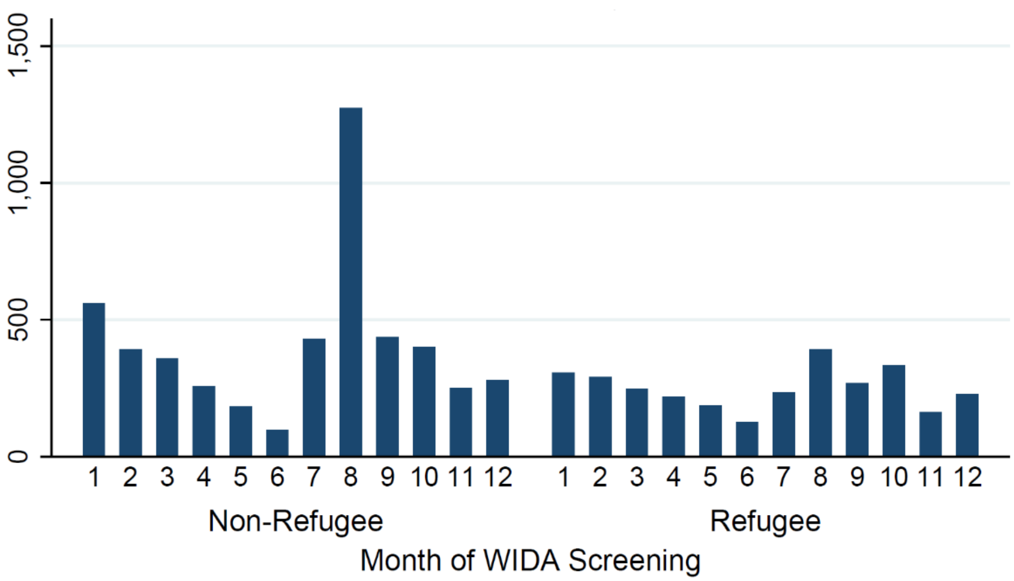
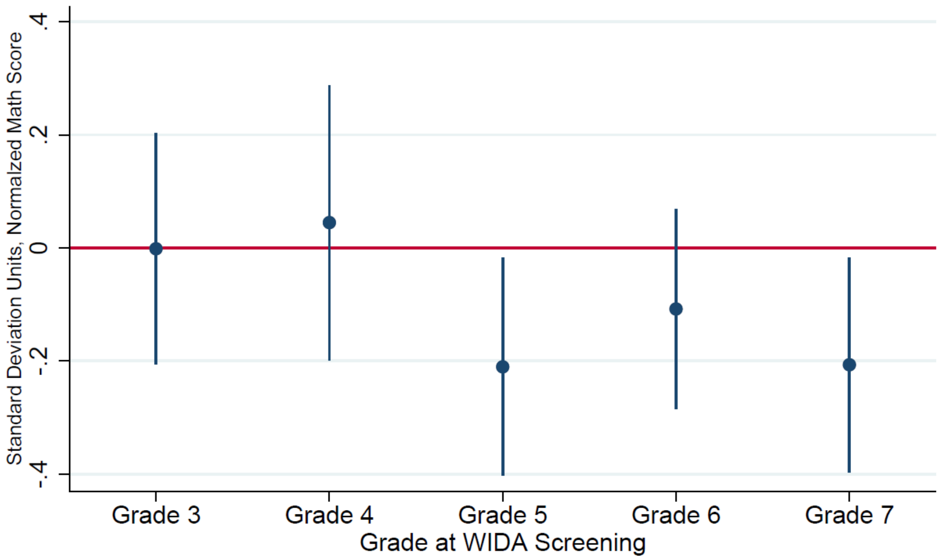
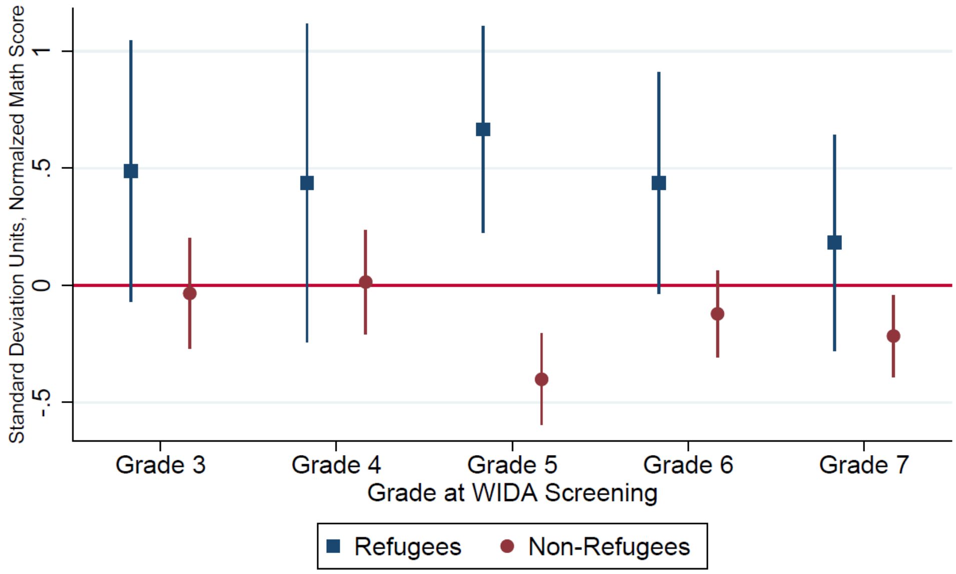


Figure A2.9: Difference-in-Differences Results: Estimated Program Eligibility Effect on Math Test Scores by Grade of EL Screening



Note: Data comprise students who take the WIDA exam in grades 1-7 and score low enough to be eligible for EL classification. Test scores are normalized to mean zero and standard deviation one with respect to the state-grade-subject distribution. Test scores correspond to the first post-WIDA grade. Estimates are from a fully specified models including achievement controls, school FEs, grade FEs, and demographic controls (indicators for female, black, Hispanic, Asian, refugee, FRL eligibility, and special education status), where the treatment indicator is interacted with indicators by grade of EL screening. The bars represent confidence intervals at the 95% level.

Figure A2.10: Difference-in-Differences Results: Estimated Program Eligibility Effect on Math Test Scores by Grade of EL Screening and Refugee Status



Note: Refugee status is based on a self-reported variable. Data comprise students who take the WIDA exam in grades 1-7 and score low enough to be eligible for EL classification. Test scores are normalized to mean zero and standard deviation one with respect to the state-grade-subject distribution. Test scores correspond to the first post-WIDA grade. Estimates are from fully specified models including achievement controls, school FEs, grade FEs, and demographic controls (indicators for female, black, Hispanic, Asian, FRL eligibility, and special education status). Bars represent confidence intervals at the 95% level.

Table A2.1: Difference-in-Differences Results – Robustness Checks and Falsification Test: Math Achievement One Year Post Treatment Eligibility, Refugee Subsample

	Robustness Checks		Falsification Test
	(1)	(2)	(3)
Low WIDA Score	-0.27	-0.19	0.08
	(0.21)	(0.13)	(0.11)
Eligible Grade	-0.72^{***}	0.36	0.95^{**}
	(0.23)	(0.47)	(0.37)
Low WIDA X Eligible Grade	0.48^{*}	0.32^{***}	-0.21^{**}
	(0.26)	(0.12)	(0.11)
WIDA Scale Score	0.53^{***}	0.54^{***}	0.42^{***}
	(0.10)	(0.06)	(0.09)
Achievement controls	X	X	X
School FE	X	X	X
Grade FE	X	X	X
Demographic Controls	X	X	X
Observations	794	1956	1590

Note: Column (1) limits the sample to refugee students who take the WIDA in grades 2 and 3 only and score low enough to be eligible for EL classification. Column (2) limits the sample to refugee students who take the WIDA in grades 1-7 with WIDA scores within 1 unit from the cutoff. Column (3) limits the sample to refugee students who take the WIDA in grades 3-7 and score low enough to be eligible for EL classification. Treatment is falsely assumed to apply only to grades 6 and 7. In all specifications, test scores are normalized to mean zero and standard deviation one with respect to the state-grade-subject distribution. Test scores correspond to the first post-WIDA grade. Demographic controls include indicators for female, black, Hispanic, Asian, FRL eligibility, and special education status. Grade FEs correspond to the grade level one-year post program eligibility. Clustered standard errors at the school level are shown in parenthesis.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.2: Fuzzy Regression Discontinuity Results: ELA Achievement One Year Post Treatment Eligibility by Grade of EL Screening

	(1) Grade 3	(2) Grade 4	(3) Grade 5	(4) Grade 6	(5) Grade 7
First Stage					
WIDA Scale Score	0.98***	1.14***	0.83***	1.10***	0.86***
	(0.18)	(0.08)	(0.16)	(0.27)	(0.13)
Second Stage					
Enrollment	0.83	-0.45***	-1.15**	-1.65**	0.64*
	(0.84)	(0.14)	(0.50)	(0.82)	(0.33)
Eff. N	298	311	271	233	239

Note: Sample includes students screened who take the WIDA exams in grades 3-7, score low enough to eligible for EL classification, and have an initial WIDA score within 1 unit of the cutoff for program participation. Test scores are normalized to mean zero and standard deviation one with respect to the state-grade-subject distribution, and they correspond to the first post-WIDA grade. Robust bias-corrected standard errors clustered at the WIDA score level are shown in parenthesis. All specifications use a linear polynomial function and triangular kernel. Point estimates correspond to the bias-corrected coefficients.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.3: Fuzzy Regression Discontinuity Results: Math Achievement One Year Post Treatment Eligibility by Grade of EL Screening

	(1)	(2)	(3)	(4)	(5)
	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7
First Stage					
WIDA Scale Score	0.99***	1.14***	0.87***	1.02***	0.87***
	(0.17)	(0.07)	(0.14)	(0.21)	(0.13)
Second Stage					
Enrollment	0.42	-0.31	-0.09	-0.76	-0.80*
	(0.69)	(0.21)	(0.48)	(0.97))	(0.43)
Eff. N	319	335	275	239	245

Note: Sample includes students screened who take the WIDA exams in grades 3-7, score low enough to eligible for EL classification, and have an initial WIDA score within 1 unit of the cutoff for program participation. Test scores are normalized to mean zero and standard deviation one with respect to the state-grade-subject distribution, and they correspond to the first post-WIDA grade. Robust bias-corrected standard errors clustered at the WIDA score level are shown in parenthesis. All specifications use a linear polynomial function and triangular kernel. Point estimates correspond to the bias-corrected coefficients.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Chapter 3: The Effect of Dual Language Immersion Programs on Student Achievement: Evidence from Georgia Schools

3.1 Introduction

Dual Language Immersion (DLI) schools offer general education in two languages – English and a target language – with the goal of promoting bilingualism and biliteracy (Boyle et al. 2015).[103] The number of DLIs has increased rapidly in recent years and estimates account for at least 2,000 programs reaching across 35 states (Maxwell 2012; Boyle et al. 2015).[104] This expansion comes as a response to both an increased demand from parents of native English speakers who anticipate positive returns in academics and labor market prospects (Parkes 2008; Ee 2018),[105] and growing evidence on the efficacy of DLI programs among English Learners (ELs) (Umansky and Reardon 2014; Valentino and Reardon 2015; Steele et al. 2017a; Bibler 2018). The dual appeal of these programs also explains the wide variety in DLI designs and implementation across school districts (Boyle et al. 2015).[106,107]

While causal evidence on the impact of bilingual education is rapidly emerging, studies remain relatively sparse and they present mixed evidence. Some studies find that DLI enrollment improves achievement in reading and English Language Arts (ELA), and reduces the time to EL reclassification (Steele et al. 2017a; Bibler 2018), while others find negative test-score effects among ELs, particularly in early grades (Jepsen 2010). Moreover, research remains concentrated

---

[103] I use Dual Language Immersion and "bilingual education" interchangeably.
[104] These estimates are likely to undercount recent expansions in North Carolina, Utah, Delaware, Georgia, and New York City.
[105] In a survey of more than 450 parents with children enrolled in DLI education, "better academic success" was the second most common reason for choosing DLI among parents whose first language did not match the program's target language. The first reason was to develop bilingual skills (Ee 2018).
[106] DLIs can be classified based on the proportion of ELs they enroll. For example, *two-way* DLIs are programs with roughly an equal split between ELs and non-ELs in a classroom. On the other hand, *one-way* DLIs only enroll students who share a first language (e.g. all native English speakers).
[107] As of SY2017-18, there were over 17 target languages offered in DLIs across the U.S. (U.S. Department of Education 2019).

in states with long-established DLI programs. For example, the most recent studies estimate the impact of DLI enrollment in Oregon (Steele et al. 2017) and North Carolina (Bibler 2018) where bilingual education began as early as the late 1980s. Thus, result from current studies may not necessarily apply to ongoing DLI expansions. Lastly, there is a dearth of literature studying the effects of DLI in early grades, specifically among native English speakers.[108]

In this paper, I present new evidence on the causal impact of DLI enrollment on student outcomes. I leverage randomized access to DLIs to estimate intent-to-treat and local-average-treatment effects of program enrollment on test scores in reading, ELA, and math. In contrast with most recent studies, I measure student achievement using early-grade test scores, allowing me to examine effects as early as Grade 2. In addition, I explore whether program effects vary by students' initial English Learner classification. I use individual-level data from Kindergarten cohorts entering DLI from SY2015-SY2017 across five newly established programs in a large school district in Georgia.[109,110]

From the intent-to-treat approach, I find no significant difference in average ELA or reading test scores by Grade 2 between students who "won" an enrollment lottery and gained access to DLI enrollment and those who "lost" a DLI enrollment lottery. However, I find weak evidence of lower math achievement among DLI lottery winners by 0.17 standard deviations. Results from the 2SLS estimation support these findings. Notably, math is one of the subjects that dual language programs teach in the target language. It is possible that these results reflect potential challenges that arise from language discrepancies between instruction and assessment.

[108] Jepsen (2010) and Slavin et al. (2011) study the effects of DLI on early grades, however their focus is on the relative efficacy of bilingual programs among English Learners, not native English speakers.
[109] The oldest program in the study was established in SY2014. I provide additional program details in the Data section.
[110] I denote school years by the calendar year of the Spring semester. That is, SY2015 represents the 2014-2015 school year.

In contrast with the results from the pooled sample, I find achievement gains among native English speakers with access to DLI enrollment, relative to those who do not win a DLI lottery.[111] Specifically, I estimate an increase in reading test scores of 0.31 standard deviations. I also find gains in ELA achievement; however, these results are not robust to all model specifications. Lastly, I do not find significant differences in math achievement between native English speakers with randomized access to dual language education and students who did not win a DLI lottery.

The rest of the paper is organized as follows. Section 2 summarizes relevant previous studies and outlines my contributions to the literature. Section 3 explains the structure of DLI programs in Georgia. Section 4 describes the data used for the study. Section 5 outlines the empirical approaches. Section 6 presents the results and discussion. Section 7 concludes.

3.2 Literature Review and Contribution

Proponents of bilingual education point to robust evidence from cognitive neuroscience that documents a strong relationship between second language acquisition and improvements in executive function (Barac et al. 2014), working memory (Morales et al. 2013; Grundy and Timmer 2017), attention control (Adesope et al. 2010), and task switching (Wiseheart et al. 2016). However, even in controlled laboratory settings, there is no definitive evidence on whether these cognitive gains associated with bilingualism translate to a direct advantage in academic tasks (Barac et al. 2014).[112] Thus, it remains an empirical question whether bilingual education fosters academic gains among students.

[111] Non-ELs make up 81 percent of DLI lottery applicants in the data.
[112] For example, studies have found similar performance between monolinguals and bilinguals in tasks such as letter identification, word reading, and reading comprehension (Lesaux and Siegel 2003; Kang 2012).

Current studies on the impact of bilingual education on student achievement present mixed results, with findings depending on students' EL classification and length of DLI exposure. Among the few studies that leverage randomized access to dual language programs, findings show that students who win a DLI lottery have higher reading and ELA test scores by Grade 3 and above, relative to students without access to DLIs. Steele et al. (2017) and Bibler (2018) each study the impact of oversubscribed DLI programs in Oregon and North Carolina, respectively. Results from both papers point to positive academic gains for both native English speakers and EL students.

On the other hand, studies that examine the impact of DLI access and enrollment in early grades find that ELs in bilingual education programs can experience a delay in English proficiency (Jepsen 2010; Slavin et al. 2011; I. M. Umansky and Reardon 2014).[113] Specifically, there is evidence of lower scores in English proficiency exams in early grades (Jepsen 2010) and slower reclassification patterns relative to ELs in English immersion programs (Umansky and Reardon 2014). However, these studies also find that differences in achievement narrow in later grades. In sum, while ELs in dual language programs may lag in achievement in early grades, research finds that differences diminish as students progress in school.

My contributions to the literature on the impact of dual language education on student achievement are twofold. First, I provide evidence on the effects of DLI enrollment using early-grade test scores. Prior research documents differences in the impact of DLI by grade levels, with evidence of null or negative outcomes in early grades (Jepsen 2010; Slavin et al. 2011). However, these studies focus on the effect of dual language education among ELs, relative to

[113] There is also evidence that ELs enrolled in dual language programs experience no change in test scores, relative to ELs in traditional ESL instruction (Chin et al. 2013)

traditional English as a Second Language instruction.[114] To my knowledge, there are no studies on the academic impact of DLI enrollment in early grades among native English speakers. I provide estimates of the effect of access to dual language education on non-ELs test scores by Grade 2.

Second, I add to a thin body of literature studying the impacts of DLI access and enrollment using data from lotteries for admission to oversubscribed programs. To my knowledge, Steele et al. (2017) and Bibler (2018) are the only papers that provide a rigorous examination of the causal effect of DLI enrollment using randomized access. This study differs from those previously mentioned by estimating the effect of five programs that are relatively new and still in their developing stages. Thus, providing evidence relevant to new expansions of DLI programs.

3.3 Dual Language Immersion Schools in Georgia

Starting with two programs in SY2011, DLI schools in Georgia have grown rapidly to 61 programs reaching across 14 school districts and enrolling 6,713 students as of SY2019.[115] DLIs in the state began as one-way programs, instructing primarily native English speakers in both the foreign target language and English. However, starting in SY2015 DLI programs became a state-approved English as Second Language (ESL) delivery model. As a result, more English Learners are enrolling in DLI thereby turning these programs into two-way models where native speakers of English and the target language are taught in the same classroom.

[114] Traditional ESL instruction refers to a combination of English immersion programs (such as pull-out or push-in methods) which focus on fast English acquisition among ELs, not bilingualism or biliteracy.
[115] https://uat.gadoe.org/Curriculum-Instruction-and-Assessment/Curriculum-and-Instruction/Pages/Dual-Immersion-Language-Programs-in-Georgia.aspx

Non-charter DLI programs use a 50:50 model such that instruction in the target language and English is split evenly during the school day.[116,117] Specifically, the target language is used for instruction in math, science, target language literacy, and sometimes social studies. English is used to teach English Language Arts (ELA) and electives such as music, art, and physical education.[118] Each classroom is supported by two teachers, where one focuses on English instruction and the other instructs exclusively in the target language. The majority of DLI programs in Georgia employ Spanish as the foreign target language (82 percent), with a total enrollment of 5,247 students. The second language with the highest enrollment is French with 587 students, followed by German and Chinese with 462 and 263 students, respectively.

DLI program participation is entirely voluntary and new applications are limited to students entering Kindergarten or Grade 1 to increase maximum target language exposure in early grades.[119] Most programs are housed within traditional public schools where a select number of classrooms are designated as DLI classrooms. Thus, total admission is limited to the number of classrooms available for the program. Students must be registered months in advance to be considered for admission, and in the event of oversubscription, offers are determined via public lotteries where students who live in the DLI school attendance zone are given priority. Separate lotteries for out-of-zone students are also held.

3.4 Data

To estimate the academic impact of access and enrollment to a DLI program, I utilize individual-level data from three cohorts of Kindergarten students who applied for DLI

[116] Other common alternatives include 90:10 target language to English and switching languages across days.
[117] Charter schools have flexibility in choosing the distribution between target language and English instruction across subjects and during the school day. As of SY2019, there are five DLI charter schools in Georgia.
[118] Support for content areas is also provided in English.
[119] DLIs are considered a long-term commitment and it is common for schools to stress that students must remain in the program up to grade 5 to experience the full potential benefits.

enrollment from SY2015-SY2017 in one of the largest school districts in Georgia.[120] The district is among the pioneers in dual language education in the state offering nine DLI programs across three languages, Spanish (7 programs), French (1 program) and Korean (1 program). I limit my sample to schools and years where there was oversubscription, such that access to DLI was determined by lotteries. In total, I include five DLI programs in the analysis sample: four programs in Spanish and one in French.

As seen in Figure 3.1, applications for DLI enrollment have increased rapidly over time – nearly quadrupling from SY2015 to SY2017. In total there were 487 DLI applications subject to a lottery over the sample period, 378 (78 percent) were lotter winners.[121] The number of oversubscribed programs has also risen, as seen in Figure 2. By SY2017, five DLI programs held at least one enrollment lottery. A lottery is held when a DLI program receives more applications than there are available seats for the incoming cohort.[122] Each school can hold two types of lotteries: one for students who live within the DLI school's attendance zone, and another for those who live outside the attendance zone.[123] Thus, randomized access to a dual language program takes place within lottery strata, which is defined by the combination of DLI school, year, and in-zone status. Overall, I identify 27 lottery strata.[124]

The data I access allow me to construct two variables of interest. First, I observe each student's lottery outcomes, which I use to construct an indicator variable denoting randomized access to a DLI program. Second, I define a variable equal to 1 for students who ever enroll in a DLI classroom from Kindergarten up to Grade 2. I require additional information to construct

[120] I access data through the Metro Atlanta Policy Lab for Education (MAPLE),
[121] This count also closely represents the number of individual students interested in DLI. Exactly 92 percent of the sample submitted only one DLI application.
[122] Lotteries are held independently by each school. There is no centralized lottery.
[123] 72 percent of lottery applicants are from within the DLI school's attendance zone.
[124] Siblings of DLI participants and children of DLI teachers are exempt from the lotteries. I omit these students from the analysis sample.

this variable because I do not directly observe DLI classroom enrollment.[125] To circumvent this issue, I identify math classrooms with clusters of DLI lottery winners and denote them as *likely* DLI classrooms.[126,127] All students in a likely DLI classroom are assigned as ever enrolled in DLI.

I merge these records with students' demographic information, such as gender and ethnicity/race, and indicators for eligibility in school programs, such as Free or Reduced-Price lunch (FRL) and English as a Second Language (ESL) instruction. As the main outcome variables, I observe Grade 2 test scores in the Iowa Tests of Basic Skills (ITBS). The ITBS are nationally norm-referenced tests that can be used to measure student growth (Dunbar and Welch 2015). Specifically, I use the standardized Normal Curve Equivalent scores and the National Percentile Rank to measure achievement in reading, English Language Arts, and math.

Applying for a DLI program is entirely voluntary; therefore, students in the analysis sample are a select subset of the overall population in the district. As seen in Table 3.1, students who apply for DLI enrollment are less likely to live in low-income households or ever be classified as English Learners, relative to the average Kindergarten cohort. DLI students are also less likely to be Hispanic or White, and more likely to be Black. However, there is wide variation in the demographic and income composition of DLI students across schools – in part hinting at heterogeneity in the type of DLI instruction. For example, the proportion of ELs across programs varies from 75 to 7 percent, which implies that DLI is primarily used as a type of ESL instruction

[125] I do not observe any direct measure of DLI enrollment. Rather I only observe if students enroll in a DLI-hosting school. Identifying DLI enrollment based on schools would lead to most in-zone applicants identified as "enrolled" simply because their home school is likely the DLI-hosting school.
[126] I use math classrooms because it is a subject all students must take, and it is taught in the target language. Overall, I identify 15 likely DLI classrooms.
[127] Classrooms are defined by the intersection of school year, school code, teacher ID, course number, and course section number.

in the former and an enrichment program in the latter.[128] Among the five DLI programs in the sample, four have fewer than 20 percent EL students.

I proceed to test whether there are significant differences in baseline demographic characteristics across lottery winners and losers, as a check for true random access to DLI enrollment. Comparing unadjusted means shows evidence of nonrandom selection among lottery winners. For example, winners are more likely to be ELs and less likely to receive Special Education services.[129] However, these differences fail to account for the fact that randomization only holds within lottery strata. Thus, I compare regression adjusted means where I control for a set of dummy variables indicating each combination of DLI school, year, and students' attendance zone status. As seen in Table 3.2, I find small and statistically insignificant differences in the regression-adjusted covariate means between lottery winners and losers.

3.5 Empirical Strategy

Leveraging randomized access to oversubscribed programs, I estimate the causal impact of receiving an offer and attending a DLI program on reading, ELA, and math test scores. I begin by estimating the following intent-to-treat specification:

$$A_{it} = \beta_0 + \beta_1^{ITT} Lott_{ij} + \omega_j + X'_{it_0}\gamma + v_{it} \qquad (1)$$

where A_{it} is the outcome of interest for student i at time t; $Lott_{ij}$ is an indicator variable for whether student i was ever offered enrollment in an oversubscribed DLI program resulting from the outcome in lottery j; ω_j indicates lottery strata fixed effects – a set of dummies for all

[128] See Table A3.1 in the appendix for summary statistics by DLI program.
[129] See Table A3.2 in the appendix for unadjusted summary statistics by DLI lottery outcome.

combinations of DLI program, year, and in-zone indicators; X_{it_0} is a vector of demographic characteristics measured at the time of the DLI lottery; and υ_{it} is the error term.

The parameter of interest, β_1^{ITT}, measures the impact of ever getting access to a DLI program, relative to students who were not offered DLI enrollment. Due to the random process by which seats are made available in oversubscribed programs, whether a student has access to a DLI school is uncorrelated with observed and unobserved individual-level characteristics that influence outcomes. Thus, estimates under this specification are unbiased. However, to the extent that lottery compliance is imperfect, results from this regression fail to capture the effect of program participation on test scores.

To examine the causal effect of DLI enrollment, I estimate the following two-stage specification where I leverage random variation in access to the program to instrument for DLI enrollment. The first- and second-stage equations are defined as follows:

$$DLI_i = \alpha_0 + \alpha_1 Lott_{ij} + \omega_j + X'_{it_0}\lambda + \varepsilon_{it} \qquad (2)$$

$$A_{it} = \beta_0 + \beta_1^{LATE}\widehat{DLI_i} + \omega_j + X'_{it_0}\gamma + \upsilon_{it} \qquad (3)$$

where DLI_i is an indicator for whether student i ever attended a DLI program from year t_0 to year t. All other variables are defined as in equation (1). The parameter of interest, β_1^{LATE}, estimates the local average treatment effect on students who are induced to attend a DLI program because of the random lottery outcome.

To better understand the causal effect estimated by β^{LATE}, it is important to clarify the counterfactual to enrolling in a DLI program. The first potential counterfactual scenario is for students to attend the same school where a DLI program is hosted. This is possible given that all DLI programs in the sample are housed within traditional public schools. If this is the case for the control group, then the beta coefficient estimates the effect of the program design and use of the target language to teach core content, while holding constant school-level characteristics. A second plausible counterfactual is for students to enroll in a school different than the DLI school – this may include those outside of the DLI attendance zone.[130] In this case, the treatment involves a change in school-level variables and peers, in addition to receiving instruction in the target language. In practice, the coefficient of interest measures a bundle of these treatments, where the effect encompasses a change in school-wide instructional environment in addition to the use of a foreign language.

A third counterfactual emerges among English Learners (ELs). ELs who do not enroll in a DLI program experience a different type of English as a Second Language (ESL) instruction, namely some type of traditional English immersion program. Therefore, among this subsample, the treatment effect encompasses the impact of DLI programs on English acquisition, relative to traditional ESL instruction.[131] Ideally, I would estimate separate ITT and LATE models by EL status to disentangle these mechanisms. However, there are not enough ELs in the sample to make a credible causal claim. Instead, I estimate variants of the models specified above restricted to the sample of non-ELs, thus excluding potential program effect mechanisms that may operate through English proficiency.

3.6 Results and Discussion

3.6.1 Main Results

I begin by estimating the intent-to-treat effects of access to DLI enrollment on student achievement. Table 3.3 shows the results from estimating variants of equation (1) using normalized test scores in reading, ELA, and math as the outcomes. Each column shows the point

[130] By construction, all the lotteries that admit out-of-zone students compare the outcomes of students who attend different schools.
[131] Traditional ESL instruction includes English immersion models (e.g. push-in and pull-out models) in combination with English-only classes.

estimates from a separate regression; the first column reports results from a model including only lottery strata fixed effects, and the second column shows estimates from fully specified models with controls for gender, race, FRL eligibility, ESL instruction, and an indicator for whether the student applied to multiple DLI lotteries.[132] All standard errors shown in the table are clustered at the school level.

I find no significant difference in reading or ELA performance between students with access to DLI enrollment and those who lost the lottery. Most point estimates are positive, but they are imprecisely estimated. However, I do find weak evidence of lower math performance among DLI lottery winners, relative to those without access to dual language programs, but the effect is imprecisely estimated. Specifically, I estimate a decrease in math achievement by up to 0.17 standard deviations that is statistically significant at a 90 percent confidence level.[133] It is worth noting that math is one of the subjects that are taught in the target language. Thus, it is possible that lower math achievement may be a result of the discrepancy between the language of instruction and the language used in the assessment. These results stand in contrast to the most recent studies that find either null or positive DLI access effects on math test scores (Steele et al. 2017a; Bibler 2018).

Table 3.4 shows the estimates from the 2SLS approach as specified in equations (2) and (3). These estimates represent the treatment on the treated, that is, the impact of program enrollment on those who comply with the outcome from the lottery. In line with the ITT effects, I estimate no impact of DLI enrollment on reading or ELA test scores. I also find negative program enrollment effects on math achievement.[134] Importantly, I find weak lottery compliance

---

[132] I run robustness checks where I limit the sample to students who apply to only one DLI lottery.
[133] Table A3.3 in the appendix shows the ITT results for reading, ELA, and math using the National Percentile Rank measure as an outcome. Results are qualitatively the same as the main specification.
[134] Table A4 in the appendix shows the ITT results for reading, ELA, and math using the National Percentile Rank

among winners. As seen in Panel A, I estimate that winning the lottery increases the likelihood of DLI enrollment by up to 28 percent. This is remarkably low given that enrollment in oversubscribed dual language programs is, in principle, only accessible by lottery (the analysis sample excludes students with guaranteed enrollment, such as siblings of current DLI students and children of DLI teachers). One potential reason for the low estimated lottery compliance is that I do not directly observe DLI enrollment apart from enrollment in a *likely* DLI math classroom. This can lead to mismeasurement in the DLI enrollment variable. For this reason, I prefer the ITT specification because it does not suffer from issues related to the quality of the data. Additionally, this is also the policy-relevant parameter.

3.6.2 Heterogeneity

The main findings were estimated using the pooled sample of students who applied for DLI enrollment and were subject to a lottery. As a result, estimates reflect the impact of a bundle of treatments against varying counterfactuals, making it difficult to isolate potential mechanisms. For example, the impact of DLIs can operate through changes in the pace of English proficiency among ELs (Valentino and Reardon 2015; I. M. Umansky and Reardon 2014). On the other hand, it remains unclear what mechanisms may drive changes in achievement among non ELs (Steele et al. 2017a).[135] While I am not able to disentangle potential mechanisms, I explore whether DLI effects differ by students' EL status. Specifically, I focus on the subset of non-EL students, as they make up over 80 percent of the DLI applicant pool.[136]

measure as an outcome. Results are qualitatively the same as the main specification.

[135] Steele et al. (2017) find modest differences between DLI lottery winners and losers in the characteristics of peers, teachers, and class size. However, none of these differences drive their estimated impacts on reading scores.

[136] Ideally, I would also estimate the effect of DLIs on achievement among ELs. However, I do not have enough observations to make a valid causal claim.

Table 3.5 shows the findings from estimating equation (1) using the subsample of non-EL students. I find positive and significant effects of access to DLI on reading test scores. Specifically, lottery winners have higher reading achievement by 0.31 standard deviations relative to students without access to dual language programs. I also estimate positive impacts on ELA performance, although these estimates are not significant in the fully specified models. Unlike the average results, I do not find evidence of lower math achievement among students with access to DLI programs. These findings are robust to including year fixed effects to account for differences in achievement over time and limiting the sample to students who apply to only one DLI lottery. Results from the robustness checks are reported in Table 3.7. Lastly, I find similar results when I estimate the effect of DLI enrollment among lottery compliers. Results from the 2SLS specification are reported in Table 3.6.[137]

3.7 Conclusion

Dual Language Immersion programs are expanding rapidly around the country, enrolling both native English speakers and English Learners. Despite this growth, research on the impact of DLI access and enrollment is sparse in comparison. I present new evidence on the effect of access and enrollment in dual language education by leveraging data from enrollment lotteries across five DLI programs from a large school district in the Atlanta metropolitan area.

Overall, I find no differences in reading or ELA test scores between students with randomized access to DLIs and those who do not win a DLI lottery. I find weak evidence of lower math achievement among DLI lottery winners, but these negative effects are only significant at a 90 percent confidence level. Given that math is one of the subjects that is taught

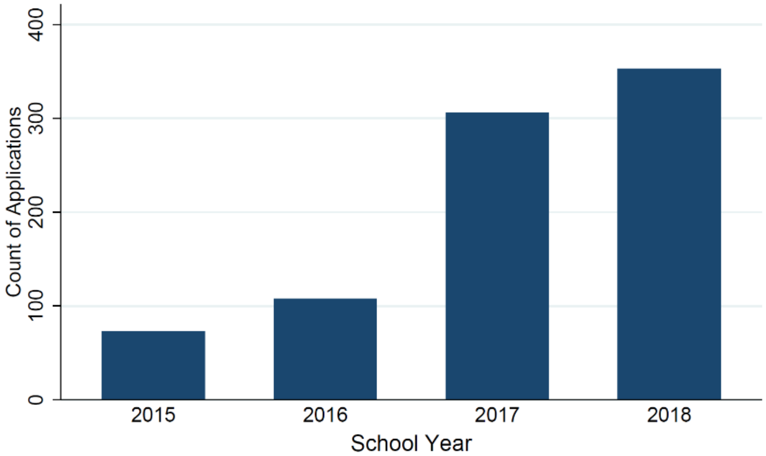
[137] Table A3.5 in the appendix shows the results from the robustness checks using the LATE specification.

in the target language, it is possible these results reflect discrepancies in language of instruction and language of assessment.

Additional analyses limiting the sample by non-EL status show that native English speakers with randomized access to DLI programs have higher reading test scores, relative to students without access. I also find no evidence of lower math achievement among this subset of students. In principle, students in dual language programs are learning bilingual and biliteracy skills. Thus, these effects among native English speakers can be interpreted as an additional benefit to enrolling in DLIs.

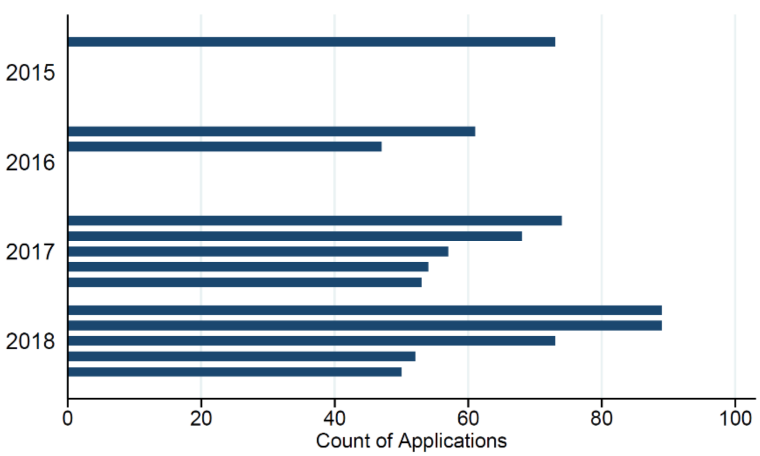
3.8 Figures and Tables

Figure 3.1: Count of DLI Lottery Applications by School Year, All DLI Programs (2015-2018)



Note: Only schools and years when lotteries were held are included in the sample. Observations are measured at the application, not applicant level. Excludes students from preference groups (siblings and children of DLI teachers) and those who apply in Grade 2 and above.

Figure 3.2: Count of DLI Lottery Applications by School Year and DLI Program (2015-2017)



Note: Each bar represents a school. Only schools and years when lotteries were held are included in the sample. Observations are measured at the application, not applicant level. Excludes students from preference groups (siblings and children of DLI teachers) and those who apply in Grade 2 and above.

Table 3.1: Unadjusted Covariate Means: Lottery Applicants and All Entering Kindergarten Cohorts (2015-2017)

	All DLI Lottery Applicants		Entering Kindergarten Cohorts	
	Mean	SD	Mean	SD
Female	0.50	0.50	0.49	0.50
Hispanic	0.24	0.43	0.32	0.47
Multiracial	0.06	0.23	0.05	0.21
White	0.40	0.49	0.50	0.50
Black	0.48	0.50	0.33	0.47
Asian	0.03	0.17	0.09	0.29
Native American	0.03	0.17	0.03	0.17
FRL Eligible	0.44	0.50	0.61	0.61
Gifted	0.00	0.05	0.00	0.00
Ever EL	0.19	0.39	0.31	0.31
ESOL	0.19	0.39	0.31	0.46
Special Ed	0.06	0.24	0.09	0.29
Observations	487		38,826	

Note: DLI lottery applicants include all students who registered for a DLI program and were subject to a lottery. Only schools and years where lottery were held are included. Demographic characteristics and program participation variables correspond to the baseline year, i.e. the year of DLI lottery registration. Observations are measured at the application, not applicant level. Excludes students from preference groups (siblings and children of DLI teachers) and those who apply in Grade 2 and above. The entering Kindergarten cohorts correspond to students who entered the district in this grade from 2015-2017.

Table 3.2: Regression Adjusted Means at Baseline Year by DLI Lottery Outcome (2015-2017)

	Lottery Winners		Lottery Losers		Difference
	Mean	SE	Mean	SE	
Female	0.50	0.06	0.49	0.03	0.01
Hispanic	0.23	0.02	0.26	0.05	-0.03
Multiracial	0.06	0.01	0.05	0.03	0.01
White	0.40	0.03	0.40	0.06	0.00
Black	0.48	0.02	0.49	0.06	-0.01
Asian	0.03	0.01	0.04	0.02	-0.01
Native American	0.03	0.01	0.02	0.02	0.01
FRL Eligible	0.43	0.03	0.48	0.06	-0.05
Ever EL	0.20	0.02	0.15	0.04	0.05
ESOL	0.20	0.02	0.15	0.04	0.05
Special Ed	0.06	0.01	0.09	0.03	-0.03
Observations	378		109		

Note: Sample of all students who registered for a DLI program and were subject to a lottery. Only schools and years where lottery were held are included. Demographic characteristics and program participation variables correspond to the baseline year, i.e. the year of DLI lottery registration. Observations are measured at the application, not applicant level. Excludes students from preference groups (siblings and children of DLI teachers) and those who apply in Grade 2 and above. Means are adjusted by lottery strata.

Table 3.3: Intent-to-Treat Results: Reading, ELA, and Math Test Scores

	Outcome: Normalized Reading Score		Outcome: Normalized ELA Score		Outcome: Normalized Math Score	
Won DLI Lottery	0.14	0.10	0.02	-0.02	-0.11	-0.17*
	(0.17)	(0.12)	(0.16)	(0.11)	(0.10)	(0.09)
Lottery FEs	X	X	X	X	X	X
Demographic Controls		X		X		X
Obs.	423	423	423	423	424	424

Note: Sample includes only the years and schools where DLI lotteries were held. All specifications control for lottery strata. Test scores in the ITBS exam in Grade 2 are used as the outcome variables. Test scores are normalized from the Normal Curve Equivalent (NCE) scores. Clustered standard errors at the school level are shown in parenthesis. This specification controls for baseline covariates. Sample excludes students from preference groups (siblings and children of DLI teachers) and those who apply in Grade 2 and above.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.4: Local Average Treatment Effects: Reading, ELA, and Math Test Scores

	Outcome: Normalized Reading Score		Outcome: Normalized ELA Score		Outcome: Normalized Math Score	
Panel A: First stage						
Won DLI Lottery	0.28***	0.25***	0.28***	0.25***	0.28***	0.25***
	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)	(0.05)
Panel B: Second stage						
Ever Enrolled in DLI	0.50	0.40	0.08	-0.08	-0.39	-0.66*
	(0.95)	(0.45)	(0.49)	(0.36)	(0.35)	(0.35)
Lottery FEs	X	X	X	X	X	X
Demographic Controls		X		X		X
Obs.	423	423	423	423	424	424

Note: DLI enrollment is measured as an indicator variable equal to 1 if the student was ever enrolled in a likely DLI math classroom. Sample includes only the years and schools where DLI lotteries were held. All specifications control for lottery strata. Test scores in the ITBS exam in Grade 2 are used as the outcome variables. Test scores are normalized from the Normal Curve Equivalent (NCE) scores. Clustered standard errors at the school level are shown in parenthesis. This specification controls for baseline covariates. Sample excludes students from preference groups (siblings and children of DLI teachers) and those who apply in Grade 2 and above.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: Intent-to-Treat Results: Reading, ELA, and Math Test Scores – Non ELs Subsample

	Outcome: Normalized Reading Score		Outcome: Normalized ELA Score		Outcome: Normalized Math Score	
Won DLI Lottery	0.37^{***}	0.31^{**}	0.23^{*}	0.16	0.08	0.01
	(0.13)	(0.14)	(0.12)	(0.12)	(0.11)	(0.10)
Lottery FEs	X	X	X	X	X	X
Demographic Controls		X		X		X
Obs.	335	335	335	335	336	336

Note: Sample includes only the years and schools where DLI lotteries were held. All specifications control for lottery strata. Test scores in the ITBS exam in Grade 2 are used as the outcome variables. Test scores are normalized from the Normal Curve Equivalent (NCE) scores. Clustered standard errors at the school level are shown in parenthesis. This specification controls for baseline covariates. Sample excludes students from preference groups (siblings and children of DLI teachers) and those who apply in Grade 2 and above. Subsample of ever non-EL students.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.6: Local Average Treatment Effects: Reading, ELA, and Math Test Scores – Non ELs Subsample

	Outcome: Normalized Reading Score		Outcome: Normalized ELA Score		Outcome: Normalized Math Score	
Panel A: First stage						
Won DLI Lottery	0.35*** (0.05)	0.32*** (0.07)	0.35*** (0.05)	0.32*** (0.07)	0.35*** (0.05)	0.32*** (0.07)
Panel B: Second stage						
Ever Enrolled in DLI	1.06** (0.43)	0.98 (0.62)	0.65* (0.38)	0.51 (0.47)	0.22 (0.32)	0.05 (0.33)
Lottery FEs	X	X	X	X	X	X
Demographic Controls		X		X		X
Obs.	335	335	335	335	336	336

Note: DLI enrollment is measured as an indicator variable equal to 1 if the student was ever enrolled in a likely DLI math classroom. Sample includes only the years and schools where DLI lotteries were held. All specifications control for lottery strata. Test scores in the ITBS exam in Grade 2 are used as the outcome variables. Test scores are normalized from the Normal Curve Equivalent (NCE) scores. Clustered standard errors at the school level are shown in parenthesis. This specification controls for baseline covariates. Sample excludes students from preference groups (siblings and children of DLI teachers) and those who apply in Grade 2 and above. Subsample of ever non-EL students.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7: Intent-to-Treat Results: Reading, ELA, and Math Test Scores – Robustness Checks

	Outcome: Normalized Math Score		Outcome: Normalized Reading Score		Outcome: Normalized ELA Score		Outcome: Normalized Math Score	
Won DLI Lottery	-0.17*	-0.11	0.31**	0.36***	0.16	0.17*	0.02	0.05
	(0.09)	(0.10)	(0.14)	(0.09)	(0.12)	(0.10)	(0.10)	(0.12)
Sample	*Pooled*	*Pooled, One Lottery App*	*Non ELs*	*Non ELs, One Lottery App*	*Non ELs*	*Non ELs, One Lottery App*	*Non ELs*	*Non ELs, One Lottery App*
Lottery Strata FEs	X	X	X	X	X	X	X	X
Demographic Controls	X	X	X	X	X	X	X	X
Year FEs	X		X		X		X	
Obs.	424	405	335	317	335	317	336	318

Note: Sample includes only the years and schools where DLI lotteries were held. All specifications control for lottery strata. Test scores in the ITBS exam in Grade 2 are used as the outcome variables. Test scores are normalized from the Normal Curve Equivalent (NCE) scores. Clustered standard errors at the school level are shown in parenthesis. This specification controls for baseline covariates. Sample excludes students from preference groups (siblings and children of DLI teachers) and those who apply in Grade 2 and above. The "One Lottery App" sample refers to the sample of students that applied to only one DLI lottery.
* $p < 0.10$, ** $p < 0.05$, *** p

Appendix A3: Additional Tables

Table A3.1: Unadjusted Covariate Means by DLI Schools, All Lottery Applicants (2015-2017)

	School A		School B		School C		School D		School E	
Target Language	*Spanish*		*Spanish*		*Spanish*		*Spanish*		*French*	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	0.49	0.50	0.42	0.50	0.49	0.50	0.46	0.50	0.53	0.50
Hispanic	0.18	0.39	0.86	0.35	0.18	0.38	0.24	0.43	0.11	0.31
Multiracial	0.06	0.24	0.04	0.19	0.08	0.27	0.09	0.29	0.04	0.21
White	0.15	0.36	0.65	0.48	0.46	0.50	0.74	0.44	0.34	0.47
Black	0.78	0.42	0.11	0.31	0.38	0.49	0.13	0.34	0.57	0.50
Asian	0.01	0.10	0.04	0.19	0.04	0.20	0.04	0.19	0.03	0.18
Native American	0.00	0.00	0.18	0.38	0.04	0.20	0.00	0.00	0.01	0.10
FRL Eligible	0.50	0.50	0.95	0.23	0.14	0.34	0.26	0.44	0.43	0.50
Ever EL	0.09	0.29	0.75	0.43	0.07	0.25	0.20	0.41	0.12	0.33
ESOL	0.09	0.29	0.75	0.43	0.07	0.25	0.20	0.41	0.12	0.32
Special Ed	0.03	0.17	0.07	0.26	0.07	0.25	0.04	0.19	0.08	0.28
Observations	100		57		74		54		202	

Note: Sample of all students who registered for a DLI program and were subject to a lottery. Only schools and years where lottery were held are included. Demographic characteristics and program participation variables correspond to the baseline year, i.e. the year of DLI lottery registration. Observations are measured at the application, not applicant level. Excludes students from preference groups (siblings and children of DLI teachers) and those who apply in Grade 2 and above.

Table A3.2: Unadjusted Covariate Means at Baseline Year by DLI Lottery Outcome (2015-2017)

	All Lottery Applicants		Lottery Winners		Lottery Losers		Difference
	Mean	SD	Mean	SD	Mean	SD	
Female	0.50	0.50	0.49	0.50	0.51	0.50	-0.022
Hispanic	0.24	0.43	0.24	0.43	0.22	0.42	0.021
Multiracial	0.06	0.23	0.06	0.23	0.06	0.23	0.003
White	0.40	0.49	0.43	0.50	0.30	0.46	0.123**
Black	0.48	0.50	0.45	0.50	0.60	0.49	-0.147***
Asian	0.03	0.17	0.03	0.18	0.02	0.13	0.016
Native American	0.03	0.17	0.03	0.18	0.03	0.16	0.004
FRL Eligible	0.44	0.50	0.45	0.50	0.40	0.49	0.049
Gifted	0.00	0.05	0.00	0.05	0.00	0.00	0.003
Ever EL	0.19	0.39	0.22	0.41	0.10	0.30	0.116***
ESOL	0.19	0.39	0.21	0.41	0.10	0.30	0.113***
Special Ed	0.06	0.24	0.05	0.22	0.10	0.30	-0.048*
One DLI Lottery	0.92	0.27	0.96	0.20	0.78	0.42	
Number of Lotteries	1.13	0.50	1.06	0.33	1.36	0.82	
Observations	487		378		109		

Note: Sample of all students who registered for a DLI program and were subject to a lottery. Only schools and years where lottery were held are included. Demographic characteristics and program participation variables correspond to the baseline year, i.e. the year of DLI lottery registration. Observations are measured at the application, not applicant level. Excludes students from preference groups (siblings and children of DLI teachers) and those who apply in Grade 2 and above.

Table A3.3: Intent-to-Treat Results: Reading, ELA, and Math Test Scores (National Percentile Rank)

	Outcome: Reading National Percentile Rank		Outcome: ELA National Percentile Rank		Outcome: Math National Percentile Rank	
Won DLI Lottery	4.43	3.19	2.03	0.67	-2.75	-4.55**
	(5.22)	(3.91)	(5.10)	(3.57)	(2.58)	(2.16)
Lottery FEs	X	X	X	X	X	X
Demographic Controls		X		X		X
N	423	423	423	423	424	424

Note: Sample includes only the years and schools where DLI lotteries were held. All specifications control for lottery strata. Test scores in the ITBS exam in Grade 2 are used as the outcome variables. Clustered standard errors at the school level are shown in parenthesis. This specification controls for baseline covariates. Sample excludes students from preference groups (siblings and children of DLI teachers) and those who apply in Grade 2 and above.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.4: Local Average Treatment Effects: Reading, ELA, and Math Test Scores (National Percentile Rank)

	Outcome: Reading National Percentile Rank		Outcome: ELA National Percentile Rank		Outcome: Math National Percentile Rank	
Panel A: First stage						
Won DLI Lottery	0.28***	0.25***	0.28***	0.25***	0.28***	0.25***
	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)	(0.05)
Panel B: Second stage						
Ever Enrolled in DLI	15.69	12.51	7.19	2.64	-9.77	-17.86**
	(15.49)	(14.04)	(15.71)	(12.02)	(8.81)	(8.75)
Lottery FEs	X	X	X	X	X	X
Demographic Controls		X		X		X
N	423	423	423	423	424	424

Note: DLI enrollment is measured as an indicator variable equal to 1 if the student was ever enrolled in a likely DLI math classroom. Sample includes only the years and schools where DLI lotteries were held. All specifications control for lottery strata. Test scores in the ITBS exam in Grade 2 are used as the outcome variables. Clustered standard errors at the school level are shown in parenthesis. This specification controls for baseline covariates. Sample excludes students from preference groups (siblings and children of DLI teachers) and those who apply in Grade 2 and above.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.5: Local Average Treatment Effects: Reading, ELA, and Math Test Scores – Robustness Checks

	Outcome: Normalized Math Score	Outcome: Normalized Reading Score	Outcome: Normalized ELA Score	Outcome: Normalized Math Score
Panel A: First stage				
Won DLI Lottery	0.17***	0.19**	0.19**	0.19**
	(0.06)	(0.10)	(0.10)	(0.10)
Panel B: Second stage				
Ever Enrolled in DLI	-0.69	1.91*	0.88	0.29
	(0.75)	(1.08)	(0.67)	(0.51)
Sample	*Pooled, One Lottery App*	*Non ELs, One Lottery App*	*Non ELs, One Lottery App*	*Non ELs, One Lottery App*
Lottery Strata FEs	X	X	X	X
Demographic Controls	X	X	X	X
Obs.	405	317	317	318

Note: DLI enrollment is measured as an indicator variable equal to 1 if the student was ever enrolled in a likely DLI math classroom.Sample includes only the years and schools where DLI lotteries were held. All specifications control for lottery strata. Test scores in the ITBS exam in Grade 2 are used as the outcome variables. Test scores are normalized from the Normal Curve Equivalent (NCE) scores. Clustered standard errors at the school level are shown in parenthesis. This specification controls for baseline covariates. Sample excludes students from preference groups (siblings and children of DLI teachers) and those who apply in Grade 2 and above.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

References

Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. "Teachers and Student Achievement in the Chicago Public High Schools." *Journal of Labor Economics* 25 (1): 95. https://doi.org/10.1086/508733.

Adesope, Olusola O., Tracy Lavin, Terri Thompson, and Charles Ungerleider. 2010. "A Systematic Review and Meta-Analysis of the Cognitive Correlates of Bilingualism." *Review of Educational Research* 80 (2): 207–45. https://doi.org/10.3102/0034654310368803.

Ahn, Thomas, and Christopher Jepsen. 2015. "The Effect of Sharing a Mother Tongue with Peers: Evidence from North Carolina Middle Schools." *IZA Journal of Migration* 4 (1): 5. https://doi.org/10.1186/s40176-015-0030-2.

Amuedo-Dorantes, Catalina, Cynthia Bansak, and Susan Pozo. 2018. "Refugee Admissions and Public Safety: Are Refugee Settlement Areas More Prone to Crime?" *IZA Discussion Paper No. 11612*. https://www.iza.org/publications/dp/11612/refugee-admissions-and-public-safety-are-refugee-settlement-areas-more-prone-to-crime.

Åslund, Olof, Per-Anders Edin, Peter Fredriksson, and Hans Grönqvist. 2011. "Peers, Neighborhoods, and Immigrant Student Achievement: Evidence from a Placement Policy." *American Economic Journal: Applied Economics* 3 (2): 67–95. https://doi.org/10.1257/app.3.2.67.

Ballatore, Rosario Maria, Margherita Fort, and Andrea Ichino. 2018. "Tower of Babel in the Classroom: Immigrants and Natives in Italian Schools." *Journal of Labor Economics* 36 (4): 885–921. https://doi.org/10.1086/697524.

Barac, Raluca, Ellen Bialystok, Dina C. Castro, and Marta Sanchez. 2014. "The Cognitive Development of Young Dual Language Learners: A Critical Review." *Early Childhood Research Quarterly* 29 (4): 699–714. https://doi.org/10.1016/j.ecresq.2014.02.003.

Batalova, Jeanne, and Margie McHugh. 2010. "Number and Growth of Students in U.S. Schools in Need of English Instruction, 2009." Migrationpolicy.Org. August 1, 2010. https://www.migrationpolicy.org/research/number-and-growth-students-us-schools-need-english-instruction-2009.

Beaman, Lori. 2012. "Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S." *The Review of Economics and Statistics* 79 (1): 128–61.

Betts, Julian. 1998. "Educational Crowding Out: Do Immigrants Affect the Educational Attainment of American Minorities?" In *Help or Hindrance? The Economic Implications of Immigration for African-Americans*, edited by Daniel Hamermesh and Frank Bean. Russell Sage Foundation.

Betts, Julian R, and Robert W Fairlie. 2003. "Does Immigration Induce 'Native Flight' from Public Schools into Private Schools?" *Journal of Public Economics* 87 (5): 987–1012. https://doi.org/10.1016/S0047-2727(01)00164-5.

Bibler, Andrew. 2018b. "Dual Language Education and Student Achievement." *SSRN Working Paper*. https://doi.org/10.2139/ssrn.3192647.

Bleakley, Hoyt, and Aimee Chin. 2004. "Language Skills and Earnings: Evidence from Childhood Immigrants." *The Review of Economics and Statistics* 86 (2): 481–96.

Bollinger, Christopher R., and Paul Hagstrom. 2008. “Food Stamp Program Participation of Refugees and Immigrants.” *Southern Economic Journal* 74 (3): 665–92. https://doi.org/10.2307/20111990.

Bossavie, Laurent. 2018. “The Effect of Immigration on Natives ’ School A Chievement : Does Length of Stay in the Host Country Matter ?” In . World Bank Group, Policy Research Working Paper.

Boyle, Andrea, Diane August, Lisa Tabaku, Susan Cole, and Ashley Simpson-Baird. 2015. “Dual Language Education Programs: Current State Policies and Practices.” American Institutes for Research.

Brunello, Giorgio, and Lorenzo Rocco. 2013. “The Effect of Immigration on the School Performance of Natives: Cross Country Evidence Using PISA Test Scores.” *Economics of Education Review* 32 (February): 234–46. https://doi.org/10.1016/j.econedurev.2012.10.006.

Burke, Mary A., and Tim R. Sass. 2013. “Classroom Peer Effects and Student Achievement.” *Journal of Labor Economics* 31 (1): 51–82. https://doi.org/10.1086/666653.

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocío Titiunik. 2017. “Rdrobust: Software for Regression-Discontinuity Designs.” *The Stata Journal: Promoting Communications on Statistics and Stata* 17 (2): 372–404. https://doi.org/10.1177/1536867X1701700208.

Capps, Randy, and Michael Fix. 2015. “Ten Facts About U.S. Refugee Resettlement.” Migrationpolicy.Org. October 20, 2015. https://www.migrationpolicy.org/research/ten-facts-about-us-refugee-resettlement.

Capps, Randy, Kathleen Newland, Susan Fratzke, Susanna Groves, Michael Fix, Margie McHugh, and Gregory Auclair. 2015. “The Integration Outcomes of U.S. Refugees: Successes and Challenges.” Migration Policy Institute. https://www.migrationpolicy.org/research/integration-outcomes-us-refugees-successes-and-challenges.

Carlson, Deven, and Jared E. Knowles. 2016. “The Effect of English Language Learner Reclassification on Student ACT Scores, High School Graduation, and Postsecondary Enrollment: Regression Discontinuity Evidence from Wisconsin.” *Journal of Policy Analysis and Management* 35 (3): 559–86. https://doi.org/10.1002/pam.21908.

Cascio, Elizabeth U., and Ethan G. Lewis. 2012. “Cracks in the Melting Pot: Immigration, School Choice, and Segregation.” *American Economic Journal: Economic Policy* 4 (3): 91–117. https://doi.org/10.1257/pol.4.3.91.

Chin, Aimee, N. Meltem Daysal, and Scott A. Imberman. 2013. “Impact of Bilingual Education Programs on Limited English Proficient Students and Their Peers: Regression Discontinuity Evidence from Texas.” *Journal of Public Economics* 107 (November): 63–78. https://doi.org/10.1016/j.jpubeco.2013.08.008.

Cho, Rosa Minhyo. 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): 629–43. https://doi.org/10.1016/j.econedurev.2012.04.006.

Conger, Dylan. 2015. “Foreign-Born Peers and Academic Performance.” *Demography* 52 (2): 569–92. https://doi.org/10.1007/s13524-015-0369-2.

Cortes, Kalena E. 2004. “Are Refugees Different from Economic Immigrants? Some Empirical Evidence on the Heterogeneity of Immigrant Groups in the United States.” *The Review of Economics and Statistics* 86 (2): 465–80. https://doi.org/10.1162/003465304323031058.

———. 2006. “The Effects of Age at Arrival and Enclave Schools on the Academic Performance of Immigrant Children.” *Economics of Education Review* 25 (2): 121–32. https://doi.org/10.1016/j.econedurev.2004.12.001.

Dagnelie, Olivier, Anna Maria Mayda, and Jean-François Maystadt. 2019. “The Labor Market Integration of Refugees in the United States: Do Entrepreneurs in the Network Help?” *European Economic Review* 111 (January): 257–72. https://doi.org/10.1016/j.euroecorev.2018.10.001.

Diette, Timothy M., and Ruth Uwaifo Oyelere. 2014. “Gender and Race Heterogeneity: The Impact of Students with Limited English on Native Students’ Performance.” *The American Economic Review* 104 (5): 412–17.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. “Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya.” *American Economic Review* 101 (5): 1739–74. https://doi.org/10.1257/aer.101.5.1739.

Dunbar, Stephen, and Catherine Welch. 2015. “Iowa Assessments Research and Development Guide.” The University of Iowa.

Dustmann, Christian, Francesco Fasani, Tommaso Frattini, Luigi Minale, and Uta Schönberg. 2017. “On the Economics and Politics of Refugee Migration.” *Economic Policy* 32 (91): 497–550. https://doi.org/10.1093/epolic/eix008.

Ee, Jongyeon. 2018. “Exploring Korean Dual Language Immersion Programs in the United States: Parents’ Reasons for Enrolling Their Children.” *International Journal of Bilingual Education and Bilingualism* 21 (6): 690–709. https://doi.org/10.1080/13670050.2016.1208144.

Epple, Dennis, and Richard E. Romano. 2011. “Peer Effects in Education: A Survey of the Theory and Evidence.” In *Handbook of Social Economics*, edited by Jess Benhabib, Alberto Bisin, and Matthew O. Jackson, 1:1053–1163. North-Holland. https://doi.org/10.1016/B978-0-444-53707-2.00003-7.

Evans, William, and Daniel Fitzgerald. 2017. “The Economic and Social Outcomes of Refugees in the United States: Evidence from the ACS.” w23498. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w23498.

Fazel, Mina, Ruth V. Reed, Catherine Panter-Brick, and Alan Stein. 2012. “Mental Health of Displaced and Refugee Children Resettled in High-Income Countries: Risk and Protective Factors.” *Lancet (London, England)* 379 (9812): 266–82. https://doi.org/10.1016/S0140-6736(11)60051-2.

Felfe, Christina, Helmut Rainer, and Judith Saurer. 2019. “Why Birthright Citizenship Matters for Immigrant Children: Short- and Long-Run Impacts on Educational Integration.” *Journal of Labor Economics*, May. https://doi.org/10.1086/704570.

Figlio, David, and Umut Özek. 2019. “Unwelcome Guests? The Effects of Refugees on the Educational Outcomes of Incumbent Students.” *Journal of Labor Economics* 37 (4): 36.

Fix, Michael, Kate Hooper, and Jie Zong. 2017. “How Are Refugees Faring? Integration at U.S. and State Levels.” Migrationpolicy.Org. June 5, 2017. https://www.migrationpolicy.org/research/how-are-refugees-faring-integration-us-and-state-levels.

Frattini, Tommaso, and Elena Meschi. 2017. "The Effect of Immigrant Peers in Vocational Schools." 11027. IZA Discussion Papers. Institute of Labor Economics (IZA). https://ideas.repec.org/p/iza/izadps/dp11027.html.

Geay, Charlotte, Sandra McNally, and Shqiponja Telhaj. 2013. "Non-Native Speakers of English in the Classroom: What Are the Effects on Pupil Performance?*." *The Economic Journal* 123 (570): F281–307. https://doi.org/10.1111/ecoj.12054.

Gould, Eric D., Victor Lavy, and M. Daniele Paserman. 2009. "Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence." *The Economic Journal* 119 (540): 1243–69. https://doi.org/10.1111/j.1468-0297.2009.02271.x.

Grundy, John G, and Kalinka Timmer. 2017. "Bilingualism and Working Memory Capacity: A Comprehensive Meta-Analysis." *Second Language Research* 33 (3): 325–40. https://doi.org/10.1177/0267658316678286.

Hoxby, Caroline. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." Working Paper 7867. National Bureau of Economic Research. https://doi.org/10.3386/w7867.

Hoxby, Caroline M., and Gretchen Weingarth. 2006. "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects."

Hunt, Jennifer. 2017. "The Impact of Immigration on the Educational Attainment of Natives." *Journal of Human Resources* 52 (4): 1060–1118. https://doi.org/10.3368/jhr.52.4.0115-6913R1.

Imberman, Scott A., Adriana D. Kugler, and Bruce I. Sacerdote. 2012. "Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees." *The American Economic Review* 102 (5): 2048–82.

Jackson, Osborne. 2015. "Does Immigration Crowd Natives Into or Out of Higher Education?" *Research Department Working Papers. Federal Reserve Bank of Boston*, November. https://www.bostonfed.org/publications/research-department-working-paper/2015/does-immigration-crowd-natives-into-or-out-of-higher-education.aspx.

Jensen, Peter, and Astrid Würtz Rasmussen. 2011. "The Effect of Immigrant Concentration in Schools on Native and Immigrant Children's Reading and Math Skills." *Economics of Education Review*, Special Issue: Economic Returns to Education, 30 (6): 1503–15. https://doi.org/10.1016/j.econedurev.2011.08.002.

Jepsen, Christopher. 2010. "Bilingual Education and English Proficiency." *Education Finance and Policy* 5 (2): 200–227.

Johnson, Angela. 2019. "The Effects of English Learner Classification on High School Graduation and College Attendance." *AERA Open* 5 (2): 233285841985080. https://doi.org/10.1177/2332858419850801.

Kang, Jennifer Yusun. 2012. "Do Bilingual Children Possess Better Phonological Awareness? Investigation of Korean Monolingual and Korean-English Bilingual Children." *Reading and Writing* 25 (2): 411–31. https://doi.org/10.1007/s11145-010-9277-4.

Koedel, Cory, Kata Mihaly, and Jonah E. Rockoff. 2015. "Value-Added Modeling: A Review." *Economics of Education Review* 47 (August): 180–95. https://doi.org/10.1016/j.econedurev.2015.01.006.

Lavy, Victor, and Analia Schlosser. 2011. "Mechanisms and Impacts of Gender Peer Effects at School." *American Economic Journal: Applied Economics* 3 (2): 1–33. https://doi.org/10.1257/app.3.2.1.

Lee, David S., and David Card. 2008. "Regression Discontinuity Inference with Specification Error." *Journal of Econometrics* 142 (2): 655–74. https://doi.org/10.1016/j.jeconom.2007.05.003.
Lee, David S, and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2): 281–355. https://doi.org/10.1257/jel.48.2.281.
Lesaux, Nonie K., and Linda S. Siegel. 2003. "The Development of Reading in Children Who Speak English as a Second Language." *Developmental Psychology* 39 (6): 1005–19. https://doi.org/10.1037/0012-1649.39.6.1005.
Lockwood, J. R., and Daniel F. McCaffrey. 2014. "Correcting for Test Score Measurement Error in ANCOVA Models for Estimating Treatment Effects." *Journal of Educational and Behavioral Statistics* 39 (1): 22–52. https://doi.org/10.3102/1076998613509405.
LoPalo, Melissa. 2019. "The Effects of Cash Assistance on Refugee Outcomes." *Journal of Public Economics* 170 (February): 27–52. https://doi.org/10.1016/j.jpubeco.2018.11.004.
Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies* 60 (3): 531–42. https://doi.org/10.2307/2298123.
Mask, Joshua. 2018. "Consequences of Immigrating During a Recession: Evidence from the US Refugee Resettlement Program." 88492. MPRA Paper. University Library of Munich, Germany. https://ideas.repec.org/p/pra/mprapa/88492.html.
Masterson, Daniel, and Vasil Yasenov. 2019. "Does Halting Refugee Resettlement Reduce Crime? Evidence from the United States Refugee Ban." *IZA Discussion Paper No. 12551*. https://www.iza.org/publications/dp/12551/does-halting-refugee-resettlement-reduce-crime-evidence-from-the-united-states-refugee-ban.
Mavrogordato, Madeline, and Rachel S. White. 2017. "Reclassification Variation: How Policy Implementation Guides the Process of Exiting Students From English Learner Status." *Educational Evaluation and Policy Analysis* 39 (2): 281–310. https://doi.org/10.3102/0162373716687075.
Maxwell, Lesli A. 2012. "'Dual' Classess See Growth in Popularity." *Education Week*, 2012.
McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2): 698–714. https://doi.org/10.1016/j.jeconom.2007.05.005.
Moffitt, Robert. 2004. "Policy Interventions, Low-Level Equilibria, and Social Interactions." In *Social Dynamics*, edited by Steven Durlauf and Peyton Young. MIT Press.
Morales, Julia, Alejandra Calvo, and Ellen Bialystok. 2013. "Working Memory Development in Monolingual and Bilingual Children." *Journal of Experimental Child Psychology* 114 (2): 187–202. https://doi.org/10.1016/j.jecp.2012.09.002.
Mossaad, Nadwa. 2016. "Annual Flow Report Refugees and Asylees: 2014." Office of Immigration Statistics, Policy Directorate. Homeland Security.
———. 2019. "Annual Flow Report Refugees and Asylees: 2017." Office of Immigration Statistics, Office of Strategy, Policy & Plans. Homeland Security. https://www.dhs.gov/sites/default/files/publications/Refugees_Asylees_2017.pdf.
Murray, Thomas J. 2016. "Public or Private? The Influence of Immigration on Native Schooling Choices in the United States." *Economics of Education Review* 53 (August): 268–83. https://doi.org/10.1016/j.econedurev.2016.04.003.

Office of Refugee Resettlement, (ORR). 2018. “Annual Report to Congress 2016.” Administration for Children and Families. https://www.acf.hhs.gov/orr/resource/office-of-refugee-resettlement-annual-report-to-congress-2016.

Ohinata, Asako, and Jan C. van Ours. 2013. “How Immigrant Children Affect the Academic Achievement of Native Dutch Children.” *The Economic Journal* 123 (570): F308–31. https://doi.org/10.1111/ecoj.12052.

Parkes, Jay. 2008. “Who Chooses Dual Language Education for Their Children and Why.” *International Journal of Bilingual Education and Bilingualism* 11 (6): 635–60. https://doi.org/10.1080/13670050802149267.

Pedraja-Chaparro, Francisco, Daniel Santín, and Rosa Simancas. 2016. “The Impact of Immigrant Concentration in Schools on Grade Retention in Spain: A Difference-in-Differences Approach.” *Applied Economics* 48 (21): 1978–90. https://doi.org/10.1080/00036846.2015.1111989.

Pope, Nolan G. 2016. “The Marginal Effect of K-12 English Language Development Programs: Evidence from Los Angeles Schools.” *Economics of Education Review* 53 (August): 311–28. https://doi.org/10.1016/j.econedurev.2016.04.009.

Reyes, Marcela, and NaYoung Hwang. 2019. “Middle School Language Classification Effects on High School Achievement and Behavioral Outcomes:” *Educational Policy*, January. https://doi.org/10.1177/0895904818823747.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. “Teachers, Schools, and Academic Achievement.” *Econometrica* 73 (2): 417–58. https://doi.org/10.1111/j.1468-0262.2005.00584.x.

Robinson - Cimpian, Joseph P., and Karen D. Thompson. 2016. “The Effects of Changing Test-Based Policies for Reclassifying English Learners.” *Journal of Policy Analysis & Management* 35 (2): 279–305. https://doi.org/10.1002/pam.21882.

Robinson, Joseph P. 2011. “Evaluating Criteria for English Learner Reclassification: A Causal-Effects Approach Using a Binding-Score Regression Discontinuity Design with Instrumental Variables.” *Educational Evaluation and Policy Analysis* 33 (3): 267–92. https://doi.org/10.3102/0162373711407912.

Rockoff, Jonah E. 2004. “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data.” *The American Economic Review* 94 (2): 247–52.

Ruiz Soto, Ariel, Sarah Hooker, and Jeanne Batalova. 2015. “States and Districts with the Highest Number and Share of English Language Learners.” Migrationpolicy.Org. 2015. https://www.migrationpolicy.org/research/states-and-districts-highest-number-and-share-english-language-learners.

Sass, Tim R., Anastasia Semykina, and Douglas N. Harris. 2014. “Value-Added Models and the Measurement of Teacher Productivity.” *Economics of Education Review* 38 (February): 9–23. https://doi.org/10.1016/j.econedurev.2013.10.003.

Schneeweis, Nicole. 2015. “Immigrant Concentration in Schools: Consequences for Native and Migrant Students.” *Labour Economics* 35 (August): 63–76. https://doi.org/10.1016/j.labeco.2015.03.004.

Schulz, Jennifer, and Jeanne Batalova. 2017. “Haitian Immigrants in the United States.” *Migration Policy Institute*. https://www.immigrationresearch.org/report/migration-policy-institute/haitian-immigrants-united-states-0.

Schwartz, Amy E., and Leanna Stiefel. 2011. "Immigrants and Inequality in Public Schools." In *Whither Opportunity?: Rising Inequality, Schools, and Children's Life Chances*, edited by Greg Duncan and Richard Murnane. New York: Russell Sage Foundation.

Setren, Elizabeth. 2019. "Targeted vs. General Education Investments: Evidence from Special Education and English Language Learners in Boston Charter Schools." *Journal of Human Resources*, November, 0219-10040R2. https://doi.org/10.3368/jhr.56.4.0219-10040R2.

Shin, Nami. 2018. "The Effects of the Initial English Language Learner Classification on Students' Later Academic Outcomes:" *Educational Evaluation and Policy Analysis*. https://doi.org/10.3102/0162373717737378.

Short, Deborah, and Beverly Boyson. 2012. "Helping Newcomer Students Succeed in Secondary Schools and Beyond." Washington DC: Center for Applied Linguistics. http://www.cal.org/resource-center/publications-products/helping-newcomer-students.

Slavin, Robert E., Nancy Madden, Margarita Calderón, Anne Chamberlain, and Megan Hennessy. 2011. "Reading and Language Outcomes of a Multiyear Randomized Evaluation of Transitional Bilingual Education." *Educational Evaluation and Policy Analysis* 33 (1): 47–58.

Snyder, Thomas D, Cristobal de Brey, and Sally A. Dillow. 2019. "Digest of Education Statistics, 2017," 905.

Steele, Jennifer L., Robert O. Slater, Gema Zamarro, Trey Miller, Jennifer Li, Susan Burkhauser, and Michael Bacon. 2017a. "Effects of Dual-Language Immersion Programs on Student Achievement: Evidence from Lottery Data." *American Educational Research Journal* 54 (1). https://doi.org/10.3102/0002831216634463.

Steele, Jennifer L., Robert Slater, Gema Zamarro, Trey Miller, Jennifer J. Li, Susan Burkhauser, and Michael Bacon. 2017b. "Dual-Language Immersion Programs Raise Student Achievement in English:" Product Page. 2017. https://www.rand.org/pubs/research_briefs/RB9903.html.

Trump, Donald J. 2019. "Executive Order on Enhancing State and Local Involvement in Refugee Resettlement." https://www.whitehouse.gov/presidential-actions/executive-order-enhancing-state-local-involvement-refugee-resettlement/.

Tumen, Semih. 2019. "Refugees and 'Native Flight' from Public to Private Schools." *IZA Discussion Paper No. 12235*. https://www.iza.org/publications/dp/12235/refugees-and-native-flight-from-public-to-private-schools.

Umansky, Ilana, and Hanna Dumont. 2019. "English Learner Labeling: How English Learner Status Shapes Teacher Perceptions of Student Skills & the Moderating Role of Bilingual Instructional Settings." Annenberg Institute at Brown University. https://www.edworkingpapers.com/ai19-94.

Umansky, Ilana M. 2016a. "To Be or Not to Be EL: An Examination of the Impact of Classifying Students as English Learners." *Educational Evaluation and Policy Analysis*, December. https://doi.org/10.3102/0162373716664802.

———. 2016b. "Leveled and Exclusionary Tracking: English Learners' Access to Academic Content in Middle School." *American Educational Research Journal*, December. https://doi.org/10.3102/0002831216675404.

———. 2018. "According to Plan? Examining the Intended and Unintended Treatment Effects of EL Classification in Early Elementary and the Transition to Middle School." *Journal* *of Research on Educational Effectiveness* 11 (4): 588–621. https://doi.org/10.1080/19345747.2018.1490470.

Umansky, Ilana M., and Sean F. Reardon. 2014. “Reclassification Patterns Among Latino English Learner Students in Bilingual, Dual Immersion, and English Immersion Classrooms.” *American Educational Research Journal* 51 (5): 879–912. https://doi.org/10.3102/0002831214545110.

UNHCR. 2019. “Global Trends: Forced Displacement In.” United National High Comissioner for Reugees. https://www.unhcr.org/en-us/statistics/unhcrstats/5d08d7ee7/unhcr-global-trends-2018.html.

U.S. Department of Education. 2019. “Dual Language Learning Programs and English Learners.” Office of English Language Acquisition. https://ncela.ed.gov/files/fast_facts/19-0389_Del4.4_DualLanguagePrograms_122319_508.pdf.

Valentino, Rachel A., and Sean F. Reardon. 2015. “Effectiveness of Four Instructional Programs Designed to Serve English Learners: Variation by Ethnicity and Initial English Proficiency.” *Educational Evaluation and Policy Analysis* 37 (4): 612–37. https://doi.org/10.3102/0162373715573310.

Wiseheart, Melody, Mythili Viswanathan, and Ellen Bialystok. 2016. “Flexibility in Task Switching by Monolinguals and Bilinguals.” *Bilingualism: Language and Cognition* 19 (1): 141–46. https://doi.org/10.1017/S1366728914000273.

Wright, Wayne E. 2015. *Foundations for Teaching English Language Learners: Research, Theory, Policy, and Practice*. Second edition. Philadelphia: Caslon Publishing.

Vita

Maria Camila Morales Leon was born in Bucaramanga, Colombia. She moved to the United States as a teenager and grew up in Clayton County, GA. In the spring of 2011, she transferred from Clayton State University to Georgia State University where she graduated *summa cum laude* with a Bachelor of Science on Economics and a Minor in Mathematics in the spring of 2013. During her undergraduate studies, Camila received the Excellence in Microeconomics Award and the Economics Student Achievement Award.

She began her PhD in Economics at Georgia State University in the fall of 2014. She worked as a research assistant for Dr. Jorge Martinez, Dr. Alberto Chong, and Dr. Tim Sass. During the last two years of her doctoral studies, Camila worked as a Graduate Research Assistant for the Metro Atlanta Policy Lab for Education where she was the lead student researcher in four projects.

During her doctorate, she taught Principles of Macroeconomics and Economic Development, and was awarded the Excellence in Teaching Award in 2019. Camila also received the Second Century Initiative (2CI) University Doctoral Fellowship, the Jack Blicksilver Award, the American Society of Hispanic Economists Dissertation Fellowship, and the Association for Education Finance and Policy New Scholar Award.

Camila was awarded her Ph.D. in Economics in the summer of 2020. Her research lies at the intersection of immigration and education policy. She has accepted a faculty position as an Assistant Professor of Economics at The University of Texas at Dallas.