EXAMINING EDUCATIONAL OPPORTUNITIES AND OUTCOMES FOR STUDENTS CLASSIFIED AS ENGLISH LEARNERS

by

LORNA MARIE PORTER

A DISSERTATION

Presented to the Department of Educational Methodology, Policy and Leadership
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

SEPTEMBER 2022

DISSERTATION APPROVAL PAGE

Student: Lorna Marie Porter

Title: Examining Educational Opportunities and Outcomes for Students Classified as English Learners

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Doctor of Philosophy degree in the Department of Educational Methodology, Policy and Leadership by:

Dr. Ilana Umansky	Chairperson/Advisor
Dr. David Liebowitz	Core Member
Dr. Karen Thompson	Core Member
Dr. Glen Waddell	Institutional Representative

and

Krista Chronister	Vice Provost for Graduate Studies

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded September 2022

DISSERTATION ABSTRACT

Lorna Porter

Doctor of Philosophy

Department of Educational Methodology, Policy and Leadership

September 2022

Title: Examining Educational Opportunities and Outcomes for Students Classified as English Learners

Students classified as English learners (EL) are a protected class of students with core legal rights. There are clear opportunities to strengthen education policies and practices to improve opportunities and outcomes for EL-classified students. This dissertation is comprised of three studies examining key issues in EL education with the goal of generating evidence to inform education policymaking that contributes to stronger education systems for EL-classified students. In Chapter II, I examine the extent to which immigrant students who arrive in grades 6-12 and are EL-classified are enrolled in core content coursework, contextualized within interviews on the policies and practices that shape course placement decisions for the student population. I also estimate the impact of participating in a newcomer program, a specialized program that districts and schools may offer, on the probability of enrollment across core content classes. I find that immigrant EL-classified students are under-enrolled in core content in comparison with peers. I find that participation in newcomer programs is not associated with fuller course access, with evidence that newcomer program participation may constrict access to certain core subjects in select years. In Chapter III I estimate the impact of a state-level EL accountability and support policy, House Bill 3499, on district and student outcomes in Oregon. I find that identified districts spent more on EL expenditures per EL-classified student, but there were no meaningful changes overall in the other outcomes examined. In Chapter IV I explore the extent to which variation in instructional effectiveness of teachers responsible for English language arts (ELA) instruction and teachers

responsible for English language development (ELD) instruction contributes to differences in students' English language arts and English language proficiency performance. I find that the variation in instructional effectiveness of both ELA and ELD teachers has implications for EL-classified students' English language arts and English language proficiency performance. Together, this set of studies provides novel evidence on areas for intervention, as well as impacts of specific interventions, in EL education. Additionally, the body of work provides implications for how this evidence can inform local and state education policy decision-making.

ACKNOWLEDGMENTS

There are many humans without whom my educational journey and this milestone would not be possible. I'm grateful for the opportunity to acknowledge the support that came along the way.

To Ilana—the decision to come to the U of O was one of the best decisions I've ever made. I couldn't have dreamed up a better advisor! You are an incredible, thoughtful scholar whose work changes systems and lives for the better. It is a privilege to know and learn from you as a scholar, mentor, role model, advocate, and friend. Your insights and encouragement pushed me to ask critical questions, to listen, and to think deeply about why we do the work we do. I'm tearing up at the thought of moving forward, but I know that there are many collaborations ahead. Thank you for everything, including believing in me when I sometimes forgot to. You truly made this experience wonderful.

David, thank you for your mentorship. You made space for me to grow and learn while providing support at every step. Karen, the way you approach your work at the intersection of policy and practice is an inspiration for the type of work I hope to do. Kevin, your mentorship at U.C. Davis is a key reason why I decided to pursue a PhD, I'm so grateful for everything you've done, and how you continue to provide mentorship a few states away. Glen, I'm so glad you let me crash your course a few years ago. You called me an "honorary economist" once and you opened up a new world of thinking for me, and I'm better at what I do for it. And Molly—thank you for taking a professional chance on me and hiring me for a dream job! I'm so grateful for your mentorship and excited for all the work we will overcommit to together…cheers to an exciting career ahead!

To Ryan—thank you for everything. I'm so grateful that the subletting neighbor downstairs turned out to be my forever partner. Thank you for extending your three month stay in Eugene to five years to accommodate my presence in your life—I can't imagine going through this process without you. You bring me so much light, inspiration, and joy. I love you.

To Olive—I went into the shelter to adopt a calm, older dog to help me get through the last years of my doctorate. What I got instead, was sandbagged. You are an absolute roller coaster of a creature, but amidst the chaos of the last years you were a bright spot that made every day a little bit better. Unconditional dog love is truly a wonder.

To Kathryn and Chad, I'm so lucky to have siblings that are also my best friends. Thank you for supporting me and keeping me entertained from both near and afar. To Grandmom and Granddad, you are true role models for living life honestly, truly, and with conviction. You've shown so much love and support for my work, thank you. To Grammie, you inspire me every day to ask critical questions, stir up some trouble, and get things done. Our years in Berkeley together were so special, and I'm so grateful for all the chaos we caused at Heinz.

Mom and Dad, it's hard to put into words all the love and appreciation I have for you both. From tearful phone calls to big hugs, it's been hard to move out of state, but a few extra miles don't change how close I am to you both. From the moment I picked up a book, you were there to foster my love for learning, and its truly that which put me on the path towards a career in education. There's never been a moment when you weren't there for me. It's hard to look back to a decade ago and think about how dark things felt for me, but you both pulled me up and out, and I hit the ground running with your love and support. I wouldn't be here without you, and I'm beyond glad I am. Thank you, and love you.

As John Prine said, "How lucky, can one [hu]man get?"

TABLE OF CONTENTS

Chapter

Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER I. INTRODUCTION

Students classified as English learners (EL[1]) are a protected class of students. EL-classified students have a set of core legal rights intended to ensure that districts and schools provide adequate supports and services such that students' English language proficiency does not foreclose their access to equitable educational opportunities (Hakuta, 2020). EL-classified students bring rich assets to their learning (Valenzuela & Rubio, 2018). However, there are many ways in which policy and practice may contribute to inequitable outcomes or fail to address existing inequities for EL-classified students (Robinson-Cimpian et al., 2017). Many EL-classified students experience constricted access to equitable and rigorous learning opportunities (Callahan, 2005; Callahan & Shifrer, 2016; Umansky, 2016), lowered teacher expectations (Umansky & Dumont, 2021), enrollment in under-resourced schooling environments (Gándara et al., 2003), and teachers who report that they are underprepared to provide appropriately modified instruction (Harper & de Jong, 2009; Santibañez & Gándara, 2018). Thus, structural barriers and challenges often create the conditions within which EL-classified students underperform on an array of outcomes in relation to their peers (National Academies of Sciences, Engineering & Medicine, 2017).

The set of three studies included in this dissertation builds from this understanding. The studies are unified by a focus on examining key issues in EL education policy with the goal of generating evidence to inform education policymaking that contributes to stronger education systems for EL-classified students. Together, this set of studies provides novel evidence on potential areas for intervention, as well as potential impacts of specific interventions, in EL education. Additionally, the body of work provides implications for how this evidence can inform local and state education policy

[1] I use the classification "English learner" and abbreviation "EL" throughout to refer to students who've been classified as English learners under federal and state policy. While this is intended to capture the unique policy context for this student group, I recognize English learner is a deficit-oriented term (Dabach & Fones, 2016) and imperfectly captures the linguistic diversity and assets of students who arrive and are in the process of adding English language proficiency to their linguistic repertoire.

decision-making.

The outline of this dissertation is as follows. I start by providing a brief introduction to the EL policy context, outlining key policy parameters and establishing the argument for why it is critical to conduct research on policy levers in EL education. I then describe the three studies that comprise this dissertation and the threads connecting them. I proceed with Chapters II, III, and IV—three unique studies—and end with a brief conclusion.

English Learner Policy Context

EL classification is a federal designation (Every Student Succeeds Act [ESSA], 2016). EL classification has roots in both judicial (*Castañeda v. Pickard*, 1981; *Lau v. Nichols*, 1974) and legislative (Equal Educational Opportunities Act, 1974) efforts to remedy systems that systematically exclude and underserve students not yet proficient in English in U.S. schools and districts. The classification is temporary, intended to ensure that districts and schools have services in place for students to equitably engage in their educational experience as they are developing English language proficiency (Hakuta, 2020). The broad ecosystem of EL policy encompasses identification of potential EL students, classification of EL students, service provision (including instruction on the English language and supports to access the core curriculum), assessment of English proficiency, and policies around reclassification—the point at which it is determined a student has reached a level of English proficiency where the classification is no longer needed (Umansky & Porter, 2020).

Many elements of EL policy are shaped at the federal level. This includes defining potential EL students, requiring states to have standardized entrance and exit criteria for EL classification, outlining key accountability parameters in EL education (ESSA, 2016), and identifying core rights of EL-classified students (Equal Educational Opportunities Act, 1974). However, a great deal of discretion is left to state and local education agencies in how to develop and implement policies for identifying, classifying, serving, and reclassifying EL-classified students. Within this discretion, a range

of policies and practices have been developed and implemented for serving EL-classified students across state and local education agencies (Callahan et al., 2022; Villegas & Pompa, 2020).

Amidst heterogeneity in the policies and practices that shape EL-classified student experiences, some education policies may contribute to inequitable outcomes observed among EL-classified students (National Academies of Sciences, Engineering & Medicine, 2017; Robinson-Cimpian et al., 2017). For example, some states do not provide additional funding resources for EL-classified students, in contrast with other states that provide additional per pupil funds or categorical grants (Education Commission of the States, 2022). As another example, some states do not require that educators participate in specific coursework or training to work with EL-classified students, in contrast with others that require certifications or basic coursework (Callahan et al., 2022). Historically, some states have put in place language policies that restrict schools' ability to provide bilingual education opportunities, while others have put into place policies that mandate extensive English language development (ELD) instruction, foreclosing opportunities to participate in many of the learning opportunities afforded to non-EL peers (Gándara & Hopkins, 2010; Lillie et al., 2012). At the local level, tracking policies and practices that limit EL-classified students' access to the full educational curriculum have been documented (Estrada, 2014; Kangas & Cook, 2020). Policy decisions tied to instruction and course placement have been linked to EL-classified students being provided with less rigorous instruction and less prepared teachers (Dabach, 2015). Further, many state and local policies fail to explicitly attend to the unique needs and strengths of EL-classified students, overlooking how the language-based classification intersects with other education policies in place (Menken & Solorza, 2010; Sugarman, 2019). Doing so can result in challenges both for students and educators who navigate the complex, constantly shifting education policy landscape.

While there are opportunities for stronger EL education policies in many instances, there are also many ways in which policy levers and interventions have been developed and implemented in

service of strengthening supports for EL-classified students. For example, in some states there have been wide swings over time from restrictive to expansive language policy. Many states now offer students the opportunity to access a Seal of Biliteracy or Multiliteracy in recognition of the assets that come with multilingualism, as well as support more widespread bilingual education opportunities (Heineke & Davin, 2020; Moore, 2021). States have developed more coherent frameworks for EL education with the goal of more aligned EL education policies and systems (Briceño & Bergey, 2022). Both state and local education agencies have invested in strengthening the bilingual educator pipeline (Garcia, 2017; Rutherford-Quach et al., 2021). There are also examples of local education agencies pushing for more expansive grassroots bilingual education (Friere et al., 2022) and supporting the implementation of linguistically responsive teaching (Zhang-Wu, 2017).

There is clear evidence that these policy levers matter. For example, more expansive bilingual education is linked to a host of positive outcomes for students (Porter et al., in press), while restrictive language policy approaches are tied to both lost opportunities to learn and stymied growth (Lillie et al., 2012). Similarly, local policies have important implications for EL-classified students. For example, varying reclassification policy approaches can lead to differences in student outcomes (Cimpian et al., 2017) and offering different language instruction programs can shape students' opportunities and outcomes (Morita-Mullaney et al., 2020). Importantly, findings that EL classification can act as both a support (Shin, 2018) or result in negative outcomes for students (Umansky, 2016) across contexts speaks to how important it is to understand the implications of EL policy decisions.

This set of studies is grounded in the idea that a range of education policy levers have important implications in EL education. However, this set of studies is also grounded in the recognition that there are critical gaps in the understanding of how education policies can create the conditions for EL-classified students to be provided with rigorous, culturally and linguistically sustaining instruction that supports equitable access to learning opportunities and timely development

of English proficiency. These studies, while focused on a range of issues in EL education, are unified by the overarching goal of building knowledge to inform EL policymaking.

Summary of Chapters

The overarching research question that guides this dissertation is: *What policy levers can be used to support more equitable outcomes for EL-classified students?* All three studies draw on Oregon statewide student records that track opportunities and outcomes for EL-classified students longitudinally.

The first study (Chapter II) is focused on a key issue in EL education—students' access to academic learning opportunities—for a unique group of EL-classified students. In the mixed methods study, I documented patterns of core content enrollment, a key measure of access to learning opportunities, for a specific group of EL-classified students: immigrant students who arrived in U.S. schools in secondary grades. While a relatively small subset of EL students, secondary newcomers face unique barriers to full content access (Short & Fitzsimmons, 2007; Umansky et al., 2018). This descriptive work lays the foundation for identifying potential areas for policy intervention, as well as provides an important example for state and local education agencies in how descriptive analyses can highlight potential constriction points in accessing opportunities to learn. This work is further contextualized within interviews with those who work in course placement to understand the intersection of individual agency and overarching course placement policies in shaping course access. Finally, I examined whether one policy lever, specialized newcomer programs, appeared to constrict or support core course access among immigrant EL-classified students arriving in secondary grades.

The results from Chapter II inform EL policymaking in multiple ways. First, observed inequities in course access among immigrant EL-classified students in relation to their peers highlight how districts and schools may struggle to ensure that immigrant students have full access to core content learning opportunities, without clear evidence that gaps in initial access are compensated for in the years following arrival. Evidence that there are disparities in access within the arriving immigrant EL-classified student

group highlights how certain student groups, such as those with interrupted formal education and initial lower English proficiency, as well as from specific racial/ethnic backgrounds, may face acute constriction of core content access. This is a clear entry point for policy intervention, with an opportunity for federal and state education agencies to invest in the provision of guidance and support to ensure that arriving immigrant EL-classified students' opportunities are not constricted in relation to their peers. Part of this guidance may be supporting local education agencies in examining their own data to identify certain timepoints, subjects, grades, or student groups for whom this issue is particularly stark. Additionally, qualitative results suggest that individuals draw on their perceptions of which students are "ready" for core content enrollment, as well as their perceptions of which teachers are trained and prepared to support students, to shape course placement decisions. This highlights an opportunity to provide guidance for those engaged in the course placement process. Such guidance can reiterate the core rights of students and provide strategies to develop schedules that ensure students are not systematically excluded, while also not placing students in classrooms they will not be supported in. To this point, there is also a need to invest in teaching supports to ensure that perceived shortages of trained educators does not keep those engaged in course placement from excluding students from core content. Finally, I find evidence that participation in a newcomer program does not support fuller course access. Such programs should be examined to ensure that the important supports provided through these programs are delivered alongside, not at the expense of, access to core content learning.

Chapter III shifts to the full population of EL-classified students, examining how one state policy lever impacts a set of proximal and distal outcomes for EL-classified students. The state policy lever, House Bill (HB) 3499, is a support and accountability policy. HB 3499 is a state policy in Oregon, with one element of the policy being a process to identify a cohort of struggling districts and provide additional funding and technical support to support investments in EL services and supports, while introducing an element of accountability to encourage meaningful change (Carnock, 2017).

Using event study and difference-in-differences analytic techniques, evidence on the estimated effects of being a district identified for HB 3499 provides important information for state policymakers, both in Oregon and those across the U.S. While I found that expenditures increased per EL-classified student at the district level, the increase varied widely in magnitude across districts. This suggests that the funds may have represented very different increases in expenditures across identified districts in comparison with non-identified districts. Therefore, the policy intervention may have allowed for some districts to invest in more substantial changes, while for others the per EL-classified student amount may have not represented a large enough sum to make widespread shifts. When looking at process-based outcomes over the first three years following identification, there was no evidence that identification led to an increase in the ratio of teachers holding ESOL endorsements to EL-classified students, nor the probability that an EL-classified student received bilingual services overall. There was also no meaningful change in students' academic outcomes. While there may be other important policy impacts uncaptured in the chapter, in relation to non-identified districts and trends prior to identification, districts identified for the intervention did not see significant changes, on average, in the outcomes examined. Policy implications include an argument for evaluating the scope of the intervention in relation to the challenges faced at the local level and the true resource costs of meaningful change, while also considering the short timeline for observing positive effects of school improvement policies.

The fourth chapter is focused on teachers of EL-classified students. A core feature of EL education is the provision of instruction on the English language, often delivered through a separate English language development (ELD) class period (Saunders et al., 2006). Many core ELD concepts align or overlap with concepts taught in English language arts (ELA) classrooms (Callahan, 2006). This overlap, along with theoretical arguments that EL-classified students' ELD and core content learning are interrelated processes (Bunch, 2013), open up questions around the different roles that ELA and ELD teachers may play in shaping EL-classified students' ELA and English language proficiency (ELP)

performance. Drawing on ELA and ELP outcomes for EL-classified students in grades 6-8 and value-added methodologies, I modeled the unique contributions of individual ELA and ELD teachers to EL-classified students' performance on ELA and ELP standardized assessments. The analyses are focused on identifying the extent to which variation in ELA and ELD teacher instructional effectiveness contributes to differences in EL-classified students' ELA and ELP performance. The goal is not to identify "good" or "bad" teachers, nor say that some teachers matter and others do not. Rather, this study is designed to add to the discussion around the interrelated nature of language and content learning for EL-classified students and explore the role that teachers across classrooms play in shaping students' outcomes. While there are many policies that shape how EL education is structured, individual teachers are those in the classroom with EL-classified students. Understanding the implications of variation in teacher effectiveness for student outcomes may be an important way to identify entry points into strengthening services and outcomes for EL-classified students.

Findings from Chapter IV suggest that variation in ELA teachers' instructional effectiveness does contribute to meaningful differences in EL-classified students' ELA performance. Variation in ELD teachers' instructional effectiveness also had implications for EL-classified students' ELA performance, although the magnitude of the variability was smaller and sensitive to model specification. Similarly, variation in both ELD and ELA teachers' instructional effectiveness did contribute to meaningful differences in EL-classified students' ELP performance, both overall and for specific domains. However, the variation was of greater magnitude for ELD teachers, while ELA contributions did not always predict meaningful differences across specifications. These findings suggest that strengthening ELA and ELD teachers' instructional effectiveness may be a viable and important policy lever for supporting stronger EL-classified student outcomes. Additionally, contributions from ELA and ELD teachers across subjects, while tentative, highlight the interrelated nature of core content and ELD learning. The study provides preliminary evidence into the importance of aligning and

strengthening instruction across language and content classrooms to support content and ELD teachers in their ability to support content and English language proficiency development among EL-classified students.

These studies, while spanning a range of issues within EL education, all focus on understanding different policy entry points in education systems for strengthening services and opportunities for EL-classified students. These studies also all focus on how evidence on these different entry points can inform policy decisions across state and local levels to strengthen EL education and supports.

REFERENCES CITED

Briceño, A., & Bergey, R. (2022). Implementing policy: Navigating the English learner roadmap for equity. *Journal for Leadership, Equity, and Research*, *8*(1).

Callahan, R. M. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal*, *42*(2), 305-328.

Callahan, R. M. (2006). The intersection of accountability and language: Can reading intervention replace English language development? *Bilingual Research Journal*, *30*(1), 1–21.

Callahan, R., Gautsch, L., Hopkins, M., & Carmen Unda, M. D. (2022). Equity and state immigrant inclusivity: English learner education in ESSA. *Educational Policy*, *36*(5), 1011-1053.

Callahan, R. M., & Shifrer, D. (2016). Equitable access for secondary English learner students: Course taking as evidence of EL program effectiveness. *Educational Administration Quarterly*, *52*(3), 463-496.

Carnock, J.T. (2017). *Pioneering change: Leveraging data to reform English learner education in Oregon*. New America.

Cimpian, J. R., Thompson, K. D., & Makowski, M. B. (2017). Evaluating English learner reclassification policy effects across districts. *American Educational Research Journal*, *54*(1_suppl), 255S-278S.

Castaneda v. Pickard. United States Court of Appeals, Fifth Circuit. 648 F.2d 989 (5th Cir. *1981*)

Dabach, D. B. (2015). Teacher placement into immigrant English learner classrooms: Limiting access in comprehensive high schools. *American Educational Research Journal*, *52*(2), 243-274.

Dabach, D. B., & Fones, A. (2016). Beyond the "English Learner" frame: Transnational funds of knowledge in social studies. *International Journal of Multicultural Education*, *18*(1), 7-27.

Education Commission of the States (2022). *50 State Comparison: English Learner Funding*. Retrieved from: https://reports.ecs.org/comparisons/k-12-and-special-education-funding-05

Estrada, P. (2014). English learner curricular streams in four middle schools: Triage in the trenches. *The Urban Review*, *46*(4), 535-573.

Every Student Succeeds Act, 20 U.S.C. § 6301 (2015).

Equal Educational Opportunities Act (1974) 20 USC Sec. 1701-1758

Freire, J. A., Delavan, M. G., & Valdez, V. E. (2022). Grassroots resistance and activism to one-size-fits-all and separate-but-equal policies by 90: 10 dual language schools en comunidades Latinas. *International Journal of Bilingual Education and Bilingualism, 25*(6), 2124-2141.

Gándara, P., & Hopkins, M. (2010). English learners and restrictive language policies. *New York, Columbia University, Teachers College*, 102-117.

Gandara, P., Rumberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English learners in California schools: Unequal resources, 'unequal outcomes. *education policy analysis archives*, *11*, 1-36.

Garcia, A. (2017). Building a Bilingual Teacher Pipeline: The Portland Public Schools and Portland State University Dual Language Teacher Partnership. *New America.*

Harper, C. A., & de Jong, E. J. (2009). English language teacher expertise: The elephant in the room. *Language and Education*, *23*(2), 137-151.

Hakuta, K. (2020). A policy history of leadership dilemmas in English learner education. *Leadership and Policy in Schools*, *19*(1), 6-9.

Heineke, A. J., & Davin, K. J. (2020). Prioritizing multilingualism in US schools: States' policy journeys to enact the seal of biliteracy. *Educational Policy*, *34*(4), 619-643.

Lau v. Nichols, 414 U.S. 563 (1974).

Lillie, K. E., Markos, A., Arias, M. B., & Wiley, T. G. (2012). Separate and not equal: The implementation of structured English immersion in Arizona's classrooms. *Teachers College Record*, *114*(9), 1-33.

Kangas, S. E., & Cook, M. (2020). Academic tracking of English learners with disabilities in middle school. *American Educational Research Journal*, *57*(6), 2415-2449.

Menken, K., & Solorza, C. (2014). No child left bilingual: Accountability and the elimination of bilingual education programs in New York City schools. *Educational Policy*, *28*(1), 96-125.

Moore, S. C. (2021). *A history of bilingual education in the US: Examining the politics of language policymaking*. Multilingual Matters.

Morita-Mullaney, T., Renn, J., & Chiu, M. M. (2020). Obscuring equity in dual language bilingual education: A longitudinal study of Emergent Bilingual achievement, course placements, and grades. *TESOL Quarterly, 54*(3), 685-718.

National Academies of Sciences, Engineering, and Medicine. (2017). Promoting the educational success of children and youth learning English: Promising futures. *National Academies Press.*

Porter, L., Vazquez Cano, M., & Umansky, I. (In press) *The Case for Bilingual Education as the Standard Service for Students Classified as English Learners.* Civil Rights Project-Proyecto Derechos Civiles.

Santibañez, L., & Gándara, P. (2018). *Teachers of English Language Learners in Secondary Schools: Gaps in Preparation and Support*. Civil Rights Project-Proyecto Derechos Civiles.

Saunders, W. M., Foorman, B. R., & Carlson, C. D. (2006). Is a separate block of time for oral English language development in programs for English learners needed? *Elementary School Journal*, *107*(2), 181–198.

Shin, N. (2018). The effects of the initial English language learner classification on students' later academic outcomes. *Educational Evaluation and Policy Analysis*, *40*(2), 175-195.

Short, D. J., & Fitzsimmons, S. (2007). *Double Work*: *Challenges and Solutions to Acquiring Language and Academic Literacy for Adolescent English Language Learners. A Report to Carnegie Corporation of New York,* Carnegie Corporation.

Sugarman, J. (2019). The unintended consequences for English learners of using the four-year graduation rate for school accountability. *Migration Policy Institute*.

Robinson-Cimpian, J. P., Thompson, K. D., & Umansky, I. M. (2016). Research and policy considerations for English learner equity. *Policy Insights from the Behavioral and Brain Sciences*, *3*(1), 129-137.

Rutherford-Quach, S., Kelly, H., Torre Gibney, D., Ballen Riccards, J., & Parker, C. (2021). *Bilingual education: Lessons from Washington State*. Comprehensive Center Network.

Umansky, I. M. (2016). Leveled and exclusionary tracking: English learners' access to academic content in middle school. *American Educational Research Journal*, *53*(6), 1792-1833.

Umansky, I. M., & Dumont, H. (2021). English learner labeling: How English learner classification in kindergarten shapes teacher perceptions of student skills and the moderating role of bilingual instructional settings. *American Educational Research Journal*, *58*(5), 993-1031.

Umansky, I., Hopkins, M., Dabach, D. B., Porter, L., Thompson, K., & Pompa, D. (2018). *Understanding and Supporting the Educational Needs of Recently Arrived Immigrant English Learner Students: Lessons for State and Local Education Agencies*. Council of Chief State School Officers.

Umansky, I. M., & Porter, L. (2020). State English Learner Education Policy: A Conceptual Framework to Guide Comprehensive Policy Action. *education policy analysis archives*, *28*(17).

Valenzuela, A., & Rubio, B. (2018). Subtractive schooling. *The TESOL encyclopedia of English language teaching*, 1-7.

Villegas, L., & Pompa, D. (2020). *The Patchy Landscape of State English Learner Policies Under ESSA*. Migration Policy Institute.

Zhang-Wu, Q. (2017). Culturally and linguistically responsive teaching in practice: A case study of a fourth-grade mainstream classroom teacher. *Journal of Education*, *197*(1), 33-40.

CHAPTER II. WHERE TO START? SECONDARY-AGE IMMIGRANT EL-CLASSIFIED STUDENTS' COURSE ACCESS AND THE ROLE OF NEWCOMER PROGRAMS

Immigrant students who arrive in secondary grades (grades 6-12) are diverse, enrolling in U.S. schools from all over the globe with varying educational experiences, assets, and needs (Carhill et al., 2008; Ruiz-de-Velasco & Fix, 2000; Thompson et al., 2020). The majority of these students are classified as English learners (EL) upon enrollment (Umansky et al., 2018). EL classification is intended to identify students who would benefit from additional linguistic supports due to being assessed to have English proficiency levels below what is necessary to fully participate in their educational opportunities without accommodations (Hakuta, 2020). As secondary-age immigrant EL-classified students arrive, districts and schools have the opportunity to adapt their programs in linguistically and culturally responsive ways to support students' academic, social, and linguistic needs.

However, districts and schools often struggle to provide appropriate supports and instructional opportunities for these students (Ruiz-de-Velasco & Fix, 2000; Short & Boyson, 2012; Short & Fitzsimmons, 2007). Within the broad challenges that local education agencies face, policymakers and practitioners have expressed concern about appropriate coursework placement for secondary-age immigrant EL-classified students. Students are on tight timelines to graduation and have fewer years in the school system to acquire English while also accessing complex content (Short & Boyson, 2007). Schools face challenges in providing instruction in response to the diversity of students' prior experiences with varying local resources, also while navigating policy pressures (Hopkins et al., 2022; Sugarman, 2019; Umansky et al., 2018). There is limited evidence on how these challenges may translate into constricted course-taking opportunities for immigrant EL-classified students who arrive in secondary grades and what policy levers may support course access.

Understanding course access predictors, levers, and outcomes for secondary-age immigrant EL-classified students is of critical importance given the challenging academic circumstances faced by

many students who arrive as immigrants in later grades. Immigrant students who arrive in later grades face tighter timelines to graduation than earlier arriving peers, more advanced core content with increased English language demands, and a higher likelihood of entering with limited or interrupted formal education (Hopkins et al., 2013; Menken, 2013; Potochnick, 2018; Sugarman, 2017; Walqui, 2000). Students' rich experiences and linguistic assets may not be acknowledged and leveraged in their schooling experiences, as administrators may "rush" students through high school (Sugarman, 2019, p. 25), potentially overlooking students' needs or postsecondary goals (Ruiz-de-Velasco & Fix, 2000). There are also tensions between providing access to core content and ensuring students have access to sufficient English language development instruction (Hopkins et al., 2022). Immigrant students who arrive in later grades face a higher drop-out rate than U.S.-born students or those who immigrated in earlier grades (Fry, 2005). These challenges point to a critical need for more evidence on how access to core content differs among immigrant EL students who arrive in secondary grades, both by student characteristics and educational interventions.

This study comes amidst a growing focus on highlighting inequities in course access among EL-classified students (Vazquez Cano et al., 2021). Federal guidance states that, regardless of a student's English proficiency, "…from enrollment to graduation, EL students are entitled to instruction in the school district's core curriculum" (Dear Colleague Letter: English Learner Students and Limited English Proficient Parents, 2015, p. 18). However, EL-classified students have been found to experience exclusion from core content areas in comparison with non-EL peers (Callahan et al., 2010; Callahan, 2005; Harklau, 1994; Johnson, 2019; Mosqueda, 2010; Umansky, 2016a). Patterns of exclusion from coursework may be evidence of barriers to equity in education systems encountered by EL-classified students (Callahan & Shifrer, 2016). While course placement policies and practices vary across districts and schools, evidence suggests that English instruction for EL-classified students is prioritized over core content (Estrada, 2014). A large body of work also identifies ways in which

students who are low-income and racial/ethnic minorities are often tracked into less rigorous coursework through course placement policies and practices (Attewell & Domina, 2008; Harklau, 1994; Oakes & Guiton, 1995; Yonezawa & Jones, 2006). Such barriers not only exclude students from opportunities to learn, but also have implications for academic growth, secondary school completion, postsecondary eligibility, and even labor market outcomes (Attewell & Domina, 2008; Gamoran & Hannigan, 2000; Giersch, 2018; Joensen & Nielsen, 2009; Long et al., 2012). While immigrant EL-classified students are diverse, many are students of color and face socioeconomic challenges (Portes & Rumbaut, 2014). An examination of course access among immigrant EL-classified students can bring to light issues of access for the population as a whole as well as identify further inequities experienced by historically marginalized and underserved student groups within the diverse immigrant EL-classified student population.

One intervention that may relate to course access for immigrant EL-classified students is recommending placement in a newcomer program. Newcomer programs, while diverse in structures and implementation, are designed to support immigrant EL-classified students during the transition to the new country and schooling environment. Newcomer programs do so by providing targeted instruction to address gaps in formal education, support English language development (ELD), and provide socioemotional services and cultural orientation (Mid-Atlantic Equity Consortium [MAEC], 2019). Typically, newcomer programs are temporary in nature, intended to support students' acute needs upon arrival until it is decided they would be better served with mainstream EL services (Short & Boyson, 2007). While newcomer programs are advocated for as a support (Faltis & Coulter, 2008; Short, 2002; Short & Boyson, 2012), there is little evidence on how they relate to student opportunities and outcomes. This study provides preliminary evidence on whether placement in a newcomer program leads to fuller or more constricted core content access in secondary school.

In this study, through interviews with Oregon administrators and educators as well as

quantitative analyses using data from a statewide sample of immigrant EL-classified students who arrived in Oregon in grades 6-12, I answer the following four research questions:

1. How does core content access for immigrant EL-classified students who arrive in secondary grades compare to access among EL-classified, non-immigrant students and non-EL classified students?
2. How does core content access differ by student characteristics among immigrant EL-classified students who arrive in secondary grades?
3. According to administrators and educators, what policies and practices shape course placement for immigrant EL-classified students who arrive in secondary grades?
4. Among immigrant EL-classified students who arrive in secondary grades, what is the estimated effect of newcomer program participation on the likelihood of being enrolled in an English language arts, math, science, or social studies course?

In the next section I provide an overview of literature on EL-classified students and course access broadly, then focus on recently arrived immigrant EL-classified students more acutely. I also outline the theoretical framework that guides this study. I then provide an overview of the data and methods before presenting findings and ending with a discussion and conclusion.

Literature Review

EL-Classified Students and Course Access

Opportunity to learn, one element of which is operationalized in this study as enrollment in core content coursework, is a critical equity indicator in EL education (Callahan & Shifrer, 2016). It is a core legal right that EL-classified students have access to equitable grade-level core content (*Castañeda v. Pickard,* 1981*; Lau v. Nichols,* 1974). However, there is variation in how this is defined. As outlined in the Dear Colleague letter (2015), to meet this core right districts can "…provide full access to the grade-appropriate core curriculum from the start of the EL program while using appropriate

language assistance strategies in the core instruction so that EL students can participate meaningfully as they acquire English" (p. 18). Alternatively, "…districts may use a curriculum that temporarily emphasizes English language acquisition over other subjects, provided that any interim academic deficits in other subjects are remedied within a reasonable length of time..." (p. 19). As written, schools and districts must ensure that, from the time students enroll, students have equitable core content access, *or* structure EL instruction and services such that there are compensatory opportunities provided after an initial period where English language services displace core content. It is worth noting that research finds EL-classified students' English language development is supported through core content enrollment, with authentic, embedded language development opportunities (Bunch, 2013), while research on intensive ELD finds that such programs can create segregated environments (Lillie et al., 2010).

Overall, there is evidence that EL-classified students are structurally excluded from core content coursework and rigorous instruction (Callahan, 2005; Callahan & Shifrer, 2016; Dabach, 2014; Gandara et al., 2003; Umansky, 2016a). Several studies use national and state datasets as well as longitudinal and causal methods to document EL-classified students' inequitable access to core academic content and rigorous coursework in comparison with peers who are not EL-classified (Callahan, 2005; Callahan & Shifrer, 2016; Estrada, 2014; Mosqueda, 2010; Johnson, 2019; Thompson, 2022; Umansky, 2016a). For example, Umansky (2016a) documents how EL classification negatively impacted access to ELA and a full courseload among middle school students. Additionally, Callahan and Shifrer (2016) find that EL-classified students complete coursework required for high school graduation at lower rates than non-EL peers. Many of these students were not recent arrivals, suggesting that exclusion happens beyond an EL-classified student's initial years in school. Research has also found that EL-classified students are tracked into lower lever courses, even when accounting for prior measured academic achievement (Callahan, 2005; Callahan & Shifrer, 2016; Mosqueda, 2010).

Disparities in course access are evidence that EL-classified students, despite the design of EL classification as a support for students (Hakuta, 2020), are often not provided with equitable or sufficient opportunities to learn, with a host of potential mechanisms that may explain why. Lower enrollment rates in core content may be evidence of course placement policies and practices that are shaped by bias, as practitioners conflate English proficiency with readiness to learn (Harklau, 1994). Exclusion may also be evidence that EL-classified students are perceived to be less academically prepared for grade-level or advanced coursework (Estrada, 2014; Harklau, 1994; Umansky, 2016b). As other potential mechanisms, EL-classified students' exclusion from core content classes may result from the belief that those working in schools are not able to provide the necessary supports for students to succeed in core content classrooms (Harklau, 1994). Further, course placement decisions may reflect school or district prioritization of English language acquisition over core content access (Estrada, 2014; Harklau, 1994; Lowenhaupt et al., 2020; Umansky, 2016a). These inequities in access to learning opportunities have the potential to impact EL-classified students' sense of academic self-efficacy (Kanno & Kangas, 2014; Lee & Soland, 2022), graduation timelines (Sugarman, 2017) and future postsecondary opportunities (Callahan & Shifrer, 2016).

Immigrant EL-Classified Students and Access to Learning Opportunities

While EL-classified students broadly face constricted access to core content courses, course access may be a pressing issue among one group within EL-classified students—those who are secondary-age arriving immigrant students. As districts and schools enroll arriving immigrant EL-classified students in secondary school, they face the question of how to structure student schedules in line with the legal parameters of EL education, as well as in response to students' diverse linguistic and academic backgrounds and their tighter timelines to graduation and postsecondary opportunities (Johnson, 2019; Sugarman, 2019). Immigrant students arriving in secondary grades face a complex set of challenges, with ELD often a key priority for supporting access to the community and educational

opportunities (Hopkins et al., 2022; Umansky et al., 2018). There are limited curricula available that focus explicitly on serving immigrant students with interrupted or limited formal education, and systems and structures can make it difficult to ensure that students' access to core content is provided in a rigorous, accessible format (Hos, 2016). Teachers may feel that immigrant students are not ready to be integrated into mainstream core content classrooms or that it is not their role to adapt instruction to meet the diversity of immigrant student needs (Penn, 2021).

In the face of these challenges and decisions, there is a limited body of research that suggests recently arrived immigrant students have constricted core content access. In research conducted in California, recently arrived immigrant EL-classified students were found to be enrolled in fewer credits than both non-EL-classified students and earlier-arriving or U.S. born EL-classified students in high school (Callahan, 2005; Johnson, 2019). Immigrant students in secondary school also reported that they experience foreclosure from accessing certain core content coursework in their initial arrival years (Hopkins et al., 2013).

There has been very little work to explore what predicts differences in course access among immigrant EL-classified students, but there is a rich body of work that identifies pre-migration, migration, and post-migration factors that impact immigrant student learning experiences and outcomes overall (McBrian, 2005; Suárez-Orozco et al., 2010). Key among them are premigration educational experiences, English proficiency upon arrival, race/ethnicity, and the district and school contexts they encounter upon arrival.

Premigration Educational Experiences

One factor that impacts learning experiences and outcomes for arriving immigrant students is their educational histories. For example, Johnson (2019) finds that differences in course-taking between recent immigrants and never-EL students are largely predicted by differences in eighth grade math and ELA academic performance, which is indicative of educational preparation prior to

migration. Additionally, Callahan (2005) finds that immigrant EL-classified students with higher prior schooling levels are, on average, placed on a higher track than those with lower prior schooling levels. This aligns with qualitative work that documents educators' concerns that later-arriving immigrant EL-classified students who have experienced academic disruptions do not have the foundation to be successful in core content options offered at the school (Ruiz-de-Velasco & Fix, 2000). Case study work by Suárez-Orozco et al. (2010) highlights how immigrant students' academic trajectories were differentiated by those who had strong prior schooling experiences and those with weaker or more limited schooling opportunities prior to migration.

In understanding the role that prior education plays in shaping arriving immigrant students' academic opportunities, research has focused on challenges experienced by students who are identified as having limited or interrupted formal education. Students with limited or interrupted formal education (SLIFE or SIFE[2]) are those who arrive with significant gaps in their education or having not received an education prior to migration (Potochnick, 2018). Educational gaps, limitations, and disruptions may reflect the conditions in their primary country or countries of residence (Browder, 2014; Custodio & O'Loughlin, 2011) or a protracted migration experience (Dryden-Peterson, 2017). SIFE may face challenges in accessing grade-level content if they do not have the expected foundational skills and knowledge (Browder, 2014). Other challenges may include adapting to a formal education environment and navigating new social structures, often with the added challenge of having experienced traumatic or difficult migration experiences (Custodio & O'Loughlin, 2017; Suárez-Orozco et al., 2010).

Initial English Language Proficiency

Another factor that has been found to impact immigrant students' educational opportunities and outcomes is their initial English language proficiency (ELP). In a latent class analysis of arriving

[2] I use SIFE in alignment with the state data used for the study.

immigrant students' academic trajectories, students with higher initial ELP were consistently on more rigorous educational trajectories than those with lower ELP levels (Suarez-Orozco et al., 2010). Research on curricular tracks in middle school revealed that EL-classified students with lower ELP levels were on tracks that included less core content than students with higher ELP levels (Estrada, 2014). Qualitative work has highlighted immigrant EL-classified students' perceptions that they were excluded from certain courses until they had reached certain ELP levels (Hopkins et al., 2013).

Student Race/Ethnicity

Immigrant students are diverse, arriving with different racial/ethnic identities that may, through racialized policies and structures, impact their educational experiences. Incorporating an examination of race within immigration and language studies is critical for understanding how the racialized policies, pratices, and institutions may result in differential experiences across different racial/ethnic immigrant identities (Sáenz & Douglas, 2015). Schools are "central sites of racialization" (Lee et al., 2017; p. 1), as immigrant students encounter persistent and powerful messaging around their racial/ethnic identity through policies, curricula, perspectives, and attitudes (Lee et al., 2017; Rodriguez, 2020). This racialization may translate into students having differential educational experiences as they navigate the different assumptions, attitudes, and biases that can directly impact students' opportunities in school (Rodriguez, 2020), as well as shape students' sense of well-being and belonging (Brown & Chu, 2012; Rodriguez et al., 2022). Race is a consistent predictor of access to higher-track coursework among all students, as Black and Hispanic/Latinx students are disproportionately enrolled in lower-track courses as compared to students with other racial/ethnic identities (Asim et al., 2019; Kelly, 2009; Tyson, 2011).

District and School Context

Postmigration factors are also linked to immigrant student outcomes. Communities with a long history of serving immigrant students may have infrastructure to provide robust services and

supports for immigrant students, translating into stronger student outcomes (Hopkins et al., 2021; Lowenhaupt & Scanlan, 2020). Students arriving in newer destinations or areas with fewer immigrant-origin residents may encounter districts and schools in the process of building their capacity to serve students, many facing scarcity of resources (Lowenhaupt, 2016; Lowenhaupt & Reeves, 2015). Depending on the community, students may also face discrimination and xenophobic attitudes and policies, which can powerfully impact student sense of well-being, safety, and belonging in school (Brown & Chu, 2012; Ee & Gándara, 2020; Gonzales et al., 2013). Instructional programs, such as those used to provide ELD as well as support core content access, are also district and school-level factors that shape student experiences and outcomes. These instructional programs influence student integration, as well as access to rigorous instruction (Dabach, 2014; Hopkins et al., 2015; Hopkins et al., 2013). One such program that may shape student experiences is newcomer programs.

Newcomer Programs. Newcomer programs are an instructional model that can be used to deliver either or both ELD and core content while also providing larger supports and services for arriving immigrant students (MAEC, 2019; Short, 2002; Short & Boyson, 2012). While programs are diverse in design and implementation, the overarching goal is to support arriving immigrant students as they transition to the new school system by providing socioemotional supports, targeted academic instruction, cultural orientation, and instruction on the English language (MAEC, 2019; Short & Boyson, 2012). The Center for Applied Linguistics conducted a national survey and ten case studies on newcomer programs (Short & Boyson, 2012). The study found that programs focused primarily on basic English and academic literacy skills, cultural orientation, and fundamental core content knowledge. Further, the authors found that strong program models included flexible schedules, specialized staff and professional development, academic instruction addressing gaps in content knowledge and basic literacy interventions, engaging families and community organizations, and strategic transition supports (Short & Boyson, 2012).

While newcomer programs are not prevalent nationally, there are common trends in structures and goals across programs. A national survey focused on programs and services for high school EL-classified students found that 16% of districts that served high-school grades with EL-classified students reported having a newcomer program (Lewis & Gray, 2016). Within these districts that provided newcomer programs, 52% of those that had a newcomer program said that it was specifically designed to support EL-classified students with interrupted formal education. Over half of districts reported that students typically spent one year or less in the program, and over 80% reported less than two years. Twenty-two percent of districts that offered newcomer programs reported that the program was a full-day model, ten percent reported a half-day program, and over sixty percent reported that the program was less than half the day (Lewis & Gray, 2016).

There is limited research on the impacts of newcomer program participation on student opportunities and outcomes. Qualitative work has largely focused on the positive pedagogical elements of newcomer programs, emphasizing the importance of providing trauma-informed, culturally responsive language and academic instruction, as well as cultural orientation (Hersi & Watkinson, 2012; Hos, 2016; Jaffe-Walter & Miranda, 2020). Quantitative research has been hindered by limited data, as EL program models are not always included in administrative datasets. Additionally, newcomer programs typically serve small groups of students, yielding small samples that are not always conducive to quantitative analyses. Houston Independent School District, where several newcomer programs are offered, provides annual, public evaluations of their newcomer programs, comparing average outcomes of immigrant students in grades 6-12 enrolled in newcomer programs as compared to immigrant students in grades 6-12 not in newcomer programs. Consistently, evaluations find that students in newcomer programs had, on average, slightly lower English language proficiency and reading score gains than immigrant students who did not participate, although estimates are not adjusted for student-level covariates (Houston Independent School District, 2017; 2018). In contrast,

there is descriptive evidence that comprehensive newcomer schools can support graduation rates and postsecondary outcomes. New York City's Internationals Network for Public Schools, which focuses on recently arrived immigrant students, was found to have higher graduation rates for immigrant EL-classified students than other New York City public schools, as well as a higher proportion of graduating students who planned to pursue a formal postsecondary education (Fine et al., 2005). This current study, using statewide data over multiple years, adds to this emergent evidence by providing key evidence on how newcomer programs shape course access while attempting to address issues of selection through a matching approach.

Theoretical Framework

This research is situated within multiple intersecting theoretical frameworks. These framings include theories of education stratification and structure-agent theory, as well as research on immigrant education and newcomer programs.

I draw on educational stratification literature to frame my first two research questions, which examine course access outcomes for immigrant EL-classified students. Theories of educational stratification and tracking suggest certain course placement policies and practices can be a mechanism for creating or maintaining differential access to learning opportunities in ways that are inequitable (Oakes, 1986). In this study, I first examine how secondary-age arriving immigrant EL-classified students' core content course-taking compares to their peers. Specifically, I look at patterns of exclusionary tracking, which describes instances where students are excluded from enrollment in core content courses (Umansky, 2016a). Exclusionary tracking, if observed, may represent ways in which policies and those working within schools and districts may be enacting or reifying patterns of educational stratification in response to students' experiences or characteristics (Oakes, 1986). The results will shed light on how districts and schools in Oregon are choosing to address the issue of equitable access to learning opportunities for this group of students. If lower enrollment levels are

observed upon arrival in line with what may be expected if schools and districts are choosing to focus intensively on ELD at the expense of core content, then it would be expected that enrollment levels would rise and perhaps even surpass non-immigrant EL-classified students over time in evidence of compensatory core content learning. Evidence on the contrary, of gaps that do not close, can be interpreted as evidence of both educational stratification and potentially constriction of students' right to equitable core content access. If lower enrollment levels are not observed upon arrival overall, that may be evidence that districts and schools are providing simultaneous language and content access, with limited evidence of overall educational stratification for immigrant EL-classified students.

However, comparisons with other, non-immigrant EL-classified peers may mask important patterns in access within the diverse group of arriving immigrant EL-classified students. In answering research question two, I position differences in access to core content among immigrant EL-classified students as potential evidence of ways in which educational opportunities may be stratified to constrict opportunities among certain student groups. Tracking research has found that students' race/ethnicity, socioeconomic status, educational interests, prior educational achievement, and parental preference all predict or contribute to decisions around which courses a student will be enrolled in (Attewell & Domina, 2008; Conger et al., 2009; Harklau, 1994; Oakes & Guiton, 1995; Umansky, 2016a; Yonezawa & Jones, 2006). My analyses are guided by this work, and my interpretation of the observed patterns is contextualized within rich theory that has highlighted how students from historically marginalized communities, particularly those who are Black or Hispanic/Latinx and those who are low-income, often experience biased education practices (Harklau, 1994; Rosa & Flores, 2017; Valenzuela, 1999).

My third research question is also situated within educational stratification and tracking theory, while integrating theory around the role of policy structures and individual agents within schools (Rigby et al., 2016). As with research questions one and two, my focus on course placement is driven

by educational stratification work, with course placement decisions as potential drivers of educational stratification. Within this, my analysis of interviews with those who are involved in course placement decisions for immigrant EL-classified students is framed through theories on the interplay between individuals who implement policy decisions (agents), and the structures that govern their work. Agents working within public agencies hold a great deal of discretion in their role of interpreting and implementing public policy (Rigby et al., 2016). While some research finds that educators implement policy systematically with little variation across individuals (Bray & Russell, 2016), in other contexts individuals' interpretations and enactment of policies may result in varying student experiences (Mavrogordato & White, 2017; White & Mavrogordato, 2019). Individual agents can be strictly confined within structures, or agents can hold autonomy and create their own practices through policy implementation (Rigby et al., 2016). My interview questions and coding analyses focus on understanding the intersection of policy and individuals' actions within course placement processes. I focus on ways in which policy is described to formally shape course placement decisions, while also identifying ways in which agents within the process are exercising their discretion to implement policies through practices that may vary across students or contexts.

The fourth research question is focused on newcomer programs and their role in shaping course access. Research on the design of newcomer programs can be used to support the theory either that newcomer programs would support *or* constrict course access.

There are reasons to believe that the design of newcomer programs, informed by research on the unique needs of arriving immigrant EL-classified students, would lead to fuller course access for those who are placed in them upon arrival. Both teachers and counselors are driving agents in determining course placement (Harklau, 1994). Participation in newcomer programs, which are typically characterized by strong teacher support and individual attention (Hos, 2016), may allow students to connect more closely with teachers and administrators. Authentically caring teachers are

critical resources for immigrant students and may be more experienced in leveraging students' assets to ensure they have access to core coursework (Hopkins et al., 2013). Having a specialized environment to support immigrant students may foster a stronger sense of shared responsibility among staff to support student needs (Hersi & Watkinson, 2012; Hos, 2016; Jaffe-Walter & Miranda, 2020). This may lead to fuller enrollment if teachers and educators provide individualized oversight of student scheduling and make it a priority to monitor if students are on track for graduation. In other instructional settings, EL-classified students may be viewed by their teachers as the responsibility of others (Lowenhaupt et al., 2020; Walker et al., 2004), and deficit-oriented perceptions may lead to exclusion from core coursework (Harklau, 1994).

There are also potentially supportive mechanisms related to academic preparation. Newcomer programs are designed to provide targeted educational supports to address prior education gaps (MAEC, 2019; Short, 2002). Pedagogically, the focus on individualized instruction, language and content supports, and socioemotional services aligns with the field's understanding of the different challenges immigrant students face (Faltis & Coulter, 2008; Short, 2002; Short & Boyson, 2012). The attention towards culturally responsive instruction and specific strategies to support immigrant or refugee students can create an environment in which students feel more comfortable participating in their coursework and activities and find support from peers in navigating the transition process (Hos, 2016; Jaffe-Walter & Miranda, 2020). Thus, participation may support academic readiness, and in turn, course access. Additionally, students who are provided with additional academic and linguistic resources in newcomer programs may be more likely to be perceived as "ready" for core content coursework, either while in newcomer programs due to the concurrent additional supports, or after transitioning out due to the impact of targeted supports. In contrast, students not in newcomer programs may be excluded from core content enrollment if educators enact restrictive course placement policies based on perceived academic deficits.

Conversely, there are reasons to believe participation in a newcomer program may constrict core content access. Newcomer programs may restrict access to a range of credit-bearing courses because they are not prioritized in the program model. It may be that, given the focus on intensive educational supports to address prior educational gaps, ELD, native language instruction, and transition services (MAEC, 2019), students are not placed in certain core content areas given time constraints. Especially if students are not transitioning out of programs, core content access may be limited in comparison with students not participating in newcomer programs even over time. Newcomer programs may also isolate students. Separate programs can limit access to the broader community of teachers and peers as well as cluster students in an under-resourced schooling environment (Chu, 2009; Feinberg, 2000; Garver & Hopkins, 2020). This may inhibit academic growth or access to opportunities to develop English proficiencies, which may be positioned as a gatekeeper to core content coursework. Additionally, EL classification, which is a deficit-oriented label, has been found to negatively impact teachers' perceptions of student ability (Umansky & Dumont, 2021). While all students in the sample for newcomer program estimates are EL-classified, it may be that an additional designation of being in a newcomer program may further negatively impact educators' perceptions of student ability, leading to constricted opportunity to learn.

Data

Quantitative Data

I leverage information on student characteristics, services, and course-taking outcomes from Oregon's longitudinal, student-level dataset from 2013/14 to 2018/19 to answer the quantitative research questions. The main population of interest is immigrant students who arrived in the U.S. in grades 6-12 and were EL-classified during that time frame, although I also draw on the full statewide population of students in grades 6-12 for comparisons in research question one. The state uses a flag to identify students who arrive as immigrants from outside the country. I used this flag, as well as

students' recorded date of entry, to identify the year in which students arrived in the Oregon data system. I dropped any student who had an earlier arrival year, but no data for that year in the system.

I use multiple student-level variables as covariates in this study. This includes eligibility for free/reduced price lunch and whether a student was ever identified for special education. Additionally, I include whether a student was identified as SIFE, defined as having at least two years less schooling than their same-grade peers, scoring two or more years below expected grade level in reading and in math, and may be pre-literate in their native language. For free/reduced price lunch eligibility and SIFE identification I impute missing values with a students' most frequently reported value. I also use data from students' initial ELP screener results. Across the years of data, districts could use one of five approved initial ELP screeners, the English Language Proficiency Assessment (ELPA) screener, the Woodcock-Muñoz, the Language Assessment Screener, the Individual Proficiency Test (IPT), and the Stanford initial English proficiency screener. The data are not converted to comparable proficiency levels in the raw data. To create a variable that was more comparable, I normalized the data using min-max normalization, which re-scales the values within each screener to be between zero and one (Wickham & Seidel, 2019).

Across analyses, I dropped any student-year record that was missing course enrollment data, as well as data on free/reduced price lunch eligibility, gender, race/ethnicity, and home language. For analyses that relied on SIFE data, I dropped any student missing data on whether they were ever identified as SIFE in the data. For analyses that relied on students' initial ELP, I also dropped any record that was missing initial ELP screener results. As seen in the results section, this means that sample sizes shift across research questions and sub-analyses. For example, analyses that look at raw differences by race/ethnicity have a much larger sample size than the analyses that model course enrollment as a function of the host of covariates of interest. This is because data on SIFE, a key covariate of interest, is unavailable before 2015/16 and initial ELP data, another key covariate of

interest, has a high missingness rate (38.36% of the full sample).

Table 1.1 contains summary statistics for the full sample of immigrant EL-classified students in their year of arrival, as well as the subset of students with non-missing SIFE and initial ELP data.

Table 1.1 Student Descriptive Statistics, Year of Arrival

	All Secondary Age Arriving Immigrant EL-Classified Students (N=4,645)	Secondary Age Arriving Immigrant EL-Classified Students with SIFE and Initial ELP Data (N=2,398)
Student demographics		
Grade of Arrival		
6	0.14	0.12
7	0.14	0.14
8	0.13	0.12
9	0.25	0.24
10	0.16	0.15
11	0.11	0.13
12	0.08	0.10
Female	0.47	0.47
Race/Ethnicity		
Multi-Ethnic & AIAN	0.01	0.01
Asian/Pacific Islander	0.29	0.27
Black	0.09	0.08
Hispanic/Latinx	0.45	0.47
White	0.16	0.16
Language		
Spanish	0.43	0.45
Arabic	0.06	0.06
Chinese	0.07	0.06
Vietnamese	0.06	0.05
Somali	0.04	0.03
Other	0.35	0.36
Identified for special education	0.01	0.01
Eligible for FRPL	0.69	0.72
SIFE	--	0.14
Lowest initial ELP level	--	0.70
School/district chars. (in year of arrival)	***Mean***	***Mean***
School % EL	0.09	0.10
School % Ever-EL	0.15	0.18
School % Recent Immigrant	0.02	0.02
School % FRPL	0.56	0.59
District % EL	0.14	0.15
District % Ever-EL	0.18	0.20
District % Recent Immigrant	0.01	0.01
District % FRPL	0.57	0.58

Note. EL=English learner. FRPL=Free/reduced price lunch. ELP = English language proficiency. Recent immigrants are students who've been in the country three years or fewer.

Across the sample, the largest proportion of students arrived in 9th grade, about 25%.[3] The most prevalent racial/ethnic group was students identified as Hispanic/Latinx, just below half, while the second most prevalent was Asian/Pacific Islander students. Very few students were identified for special education at any point in the data. The majority of students were eligible for free/reduced price lunch. Overall, the districts and schools that students arrived in had, on average, over fifty percent of their student population eligible for free/reduced price lunch. The proportion of students who were recent immigrants was low both at the district and school level, while on average about 15% of students were ever EL-classified in the schools students arrived in, and 18% in the districts.

There is a second column in Table 1.1 which presents the same information, but only for students who have complete SIFE and initial ELP data. The SIFE variable was only collected by the state from 2015/16 onward, so for research question two I am restricted to those four years of data. Among the students in the reduced sample, 14% were SIFE. Thirty-four percent of students' initial ELP levels were above the minimum level on the re-scaled assessment scores, meaning that the majority of incoming students' initial assessed ELP was the minimum screener level.[4] Other data patterns were similar to those of the full analytic sample, with slight differences.

Research question four focuses on newcomer programs. The newcomer program variable provided in the data comes from state-level requirements that, for each EL-classified student, schools report the language instruction program and content instruction program used to support ELD and core content access. Language instruction program options for ELD include ELD push-in, ELD pull-out, ELD class period, and newcomer program-ELD. Content instructional programs include two-way immersion, transitional bilingual, developmental bilingual, other bilingual, sheltered instruction,

[3] This may be indicative of districts and schools sorting students into 9th grade upon arrival or of strategic migration timing on the part of families.

[4] The average, rescaled screener score was 0.13 on a scale of 0 to 1.

and newcomer program-core content. I define treatment as having either or both codes be reported as newcomer program in the year of arrival; in almost all cases if one was coded for newcomer program, so was the other. I drop any student whose code was newcomer program, but the school had fewer than five students in a newcomer program in that year.[5] Because program model data changed prior to 2015/16, I only use data from 2015/16 onward. I code treatment as binary, based on placement in the year of arrival regardless of time in program. I drop any student who was not in a newcomer program in the year of arrival but was in a newcomer program in later years. Thus, the treatment captures the effects of being placed in the program upon arrival, regardless of how long students were in the programs, excluding students who were placed in newcomer programs in years after their arrival.

In Figure A1, I plot the proportion of arriving immigrant EL-classified students by their year in school and language instruction program combination. The language instruction program variable was created by combining both the program used to receive ELD services and support core content learning. There were only six combinations of programs in the analytic sample, although an additional two variable values included non-participation and being in former/monitored status. As seen in Figure A1, the majority of students received ELD through a class period and core content instruction through sheltered instruction. Newcomer programs were the second most prevalent program model upon arrival. In the year of arrival, 14% of students were in newcomer programs. The following year this percentage dropped to 10%, 7% in year three, and 4% in year four. Table 1.2 contains summary statistics for the students in the research question four analytic sample in their first year in U.S. schools (n=2,370[6]), by whether a student was placed in a newcomer program upon arrival.

[5] This likely represents a conceptualization of newcomer programs that differs from that of interest in this study or a data reporting error.

[6] This sample is slightly smaller than the right column of Table 1 due to missing language instruction program data, exclusion of students who were in newcomer programs with fewer than five students in a year, or exclusion of students who were in not in newcomer programs in their first year but were in newcomer programs in later years.

Table 1.2 Research Question Four Descriptive Statistics, Year One

	Not in Newcomer Program (N=2,036)	In a Newcomer Program (N=334)	*p*-value
Student demographics			
Grade of arrival			0.4
6	0.13	0.09	
7	0.15	0.13	
8	0.11	0.12	
9	0.24	0.24	
10	0.16	0.15	
11	0.13	0.14	
12	0.09	0.12	
Female	0.47	0.47	0.8
Race/Ethnicity			<0.001
Multi-Ethnic & AIAN	0.01	0.00	
Asian/Pacific Islander	0.28	0.19	
Black	0.08	0.12	
Hispanic/Latinx	0.45	0.60	
White	0.18	0.09	
Language			<0.001
Spanish	0.43	0.56	
Arabic	0.06	0.02	
Chinese	0.06	0.04	
Vietnamese	0.05	0.04	
Somali	0.02	0.05	
Other	0.37	0.28	
Special education	0.01	0.00	0.5
SIFE	0.14	0.14	>0.9
Initial ELP, lowest level	0.67	0.91	
Eligible for FRPL	0.69	0.92	<0.001
Average # years in Newcomer Program	--	1.65	
School and district characteristics (in year of arrival)	***M***	***M***	
School % EL	0.09	0.15	<0.001
School % Ever-EL	0.16	0.28	<0.001
School % Recent Immigrant	0.02	0.02	<0.001
School % FRPL	0.56	0.77	<0.001
District % EL	0.14	0.16	<0.001
District % Ever-EL	0.20	0.23	<0.001
District % Recent Immigrant	0.01	0.01	<0.001
District % FRPL	0.57	0.66	<0.001

Note. EL=English learner. FRPL=Free/reduced price lunch. SIFE= Student with interrupted formal education. ELP=English language proficiency. Recent immigrants are students who've been in the country three years or fewer.

A minority of arriving students were placed in a newcomer program upon arrival, 14% (*n*=334) of immigrant EL-classified students who arrived in grades 6-12, statewide. There were differences in the student populations placed in newcomer programs and those who were not. A larger proportion of students in newcomer programs were Hispanic/Latinx and Black than students not in newcomer

programs, where a larger proportion of students were Asian/Pacific Islander and White. A larger proportion of students in newcomer programs were eligible for free/reduced price lunch (92% compared to 69%), although the same percentage were SIFE and there were no differences by gender. Students in newcomer programs were in districts and schools that, on average, had higher proportions of EL- and ever-EL-classified students, as well as higher proportions of students eligible for free/reduced price lunch.

Tables A1 and A2 in the Appendix also present descriptive statistics by newcomer program participation, but in Table A1 the sample is reduced to students who arrived in the 15 districts that offered newcomer programs (n=1,394), and in Table A2 to students who were enrolled upon arrival in the 28 schools that offered newcomer programs (n=647). Within districts that offered newcomer programs, 24% of students in the sample were placed in a newcomer program, while 52% of students who arrived in schools that offered newcomer programs were placed in newcomer programs. As with the full sample, in both reduced samples a higher percentage of students in newcomer programs were eligible for free/reduced price lunch and identified as Hispanic/Latinx and Black than students not in newcomer programs. In contrast with the full sample, a lower percentage of students in newcomer programs were identified as SIFE. Students in newcomer programs, on average, attended schools with higher proportions of EL-classified students, ever-EL classified students, and students eligible for free/reduced price lunch.

The key outcomes of interest in this study were course enrollment variables, part of a data file which collects a student level observation for each course enrolled in over the academic year. I merged course-taking data with the main student file, dropping any course enrollment observation where the student missed more than two thirds of the course enrollment period and the course end date was

different from the students' end date in the course.[7] I then created multiple binary indicators capturing if a student was enrolled (1) or not (0) in a given year in each of the following four content areas: ELA, math, science, and social studies. I used National Center for Education Statistics (2021) subject area classifications to determine if an enrollment record counted as an ELA, math, science, social studies, or other course, then created a dichotomous variable for each student-year record to capture if they had an enrollment record within that subject area in a given academic year, taking the maximum value in a given year across multiple enrollment records.[8]

Qualitative Data

I conducted interviews with eleven staff from six districts across the state of Oregon focused on enrollment processes for arriving secondary age immigrant EL-classified students. In total, of fifteen districts selected for geographic variation and variation in the proportion of their school that was EL-classified, six districts were willing and able to participate. District personnel identified key administrators or educators involved in immigrant EL-classified student policy or course placement policy. I contacted those personnel for interviews, also asking them to recommend additional personnel to interview. In three districts I spoke to only one person, in two districts I spoke to two, and in one district I spoke to four.

Interviews, some in-person and some over the phone, ranged in length from twenty minutes to an hour and covered several topics on enrollment policies and practices for registering and creating schedules for secondary age immigrant EL-classified students. Interviewees held a range of positions,

[7] This was meant to exclude observations that captured initial, short-term enrollment periods where students bounced between courses, as is common in secondary school, but to retain enrollment records where students arrived partway through a term or year but were enrolled through the end of the term or year.

[8] This approach is blunt. One analytic risk is that comes with this approach is overestimating course enrollment, given that I may be capturing instances where a student is enrolled for a short amount of time. Additionally, schools and districts used varying reporting methods, with some reporting in year-long course records, while others reported enrollment by semesters, trimesters, or quarters, with inconsistencies in reporting length within schools as well. The binary indicators described then have the potential to overestimate enrollment for schools and districts that report classes on a quarter or trimester schedule, given that there are more opportunities for a student to be enrolled, even if it does not translate into as much time accessing the subject matter.

including overseeing various federal programs, teachers, counselors, and an EL program coordinator, with some holding roles at the district central office and others on school campuses. The six districts, which will remain anonymous, represented varying locales in Oregon, including districts classified as rural, town, suburban, and urban. Additionally, while some served large immigrant populations, others served smaller populations. Interviews were transcribed and transcripts uploaded to Dedoose.

Analytic Approach

I utilized a concurrent, mixed-methods research design to answer the four research questions (Creswell & Clark, 2007). For research questions one and two, I quantitatively examined differences in course enrollment to identify key patterns in differences in course access. For research question three, I analyzed interviews for information about the policies and practices that shaped course access. This analysis provided context for, and information on, the patterns observed in research questions one and two. For research question four, I used quantitative data to examine how newcomer programs impacted course access.

Research Question 1: Descriptive Comparisons of Course Access

For research question one I relied on descriptive statistics to describe course access patterns. I first compared enrollment levels of immigrant EL-classified students who arrived in secondary grades as students progressed through school in comparison with their peers, broken out by grade of arrival. For each grade of arrival (i.e., 9th grade), I calculated the proportion of arriving immigrant EL-classified students enrolled in a given core content area in the year of arrival and subsequent grades, in comparison with students who were EL-classified in that same grade but not recently arrived immigrant students, as well as students in the same grade who were not EL-classified. I then plotted results. As an example, I calculated the ELA enrollment rate for immigrant EL-classified students who arrived in 6th grade in their year of arrival, as well as for each subsequent grade they were observed in over the data. I compared this enrollment rate over time with the ELA enrollment rate for all students

who were, in 6th grade, EL-classified but not recent immigrants (regardless of if they were reclassified to fluent English proficient over the next six years), as well as all students who were not EL-classified in 6th grade (including both students who had never been EL-classified and those who had been EL-classified but who had reclassified out of EL status at some point prior to 6th grade).

Research Question 2: Differences in Course Access Among Secondary-Age Arriving Immigrant EL-Classified Students

It is also valuable to also understand differences in rates of course access within the population of secondary-age arriving immigrant EL-classified students. This can help to illuminate the diversity within the student group and probe for potential inequities. I first ran simple descriptive statistics and created data visualizations to highlight raw differences by student characteristics. While raw comparisons do not account for the intersection of multiple factors that may drive differences in course access, they do provide important insights into inequities as observed, which sometimes are "control[ed] away" (Castillo & Gillborn, 2022, p. 7) through modeling decisions. Decisions to simply control for student characteristics or contextual characteristics can obscure or overlook important sources of inequity that are often tied to racism and racialized policies (Castillo & Gillborn, 2022).

To complement raw descriptive statistics, I ran a set of linear probability models to predict course enrollment by student characteristics (n=2,398 students, 5,156 student-year observations). I fit separate models predicting enrollment in the four primary content areas (ELA, math, science, social studies) from a host of student covariates, including race/ethnicity, SIFE, eligibility for free/reduced price lunch, initial ELP, ever-identification for special education services, and gender. I included fixed effects for grade of arrival, time, and district.[9] This allowed me to focus on the predictive role of student-level characteristics while controlling for variation attributable to what were likely differential

[9] I also run all the above specified models as logit models. Results were substantively similar in almost all instances, although for select outcomes in select years statistical precision varied. I report linear probability results for ease of interpretation, but logit results are available upon request.

course placement practices by grade as well as differences attributable to time-invariant district characteristics and time. I clustered standard errors at the district level.

As a further way to explore heterogeneity in course access among immigrant EL-classified students, I also ran a series of hierarchical linear models predicting the likelihood of being enrolled in each course area for secondary-age arriving immigrant EL-classified students. The first model included student characteristics (the same that were in the fixed effect specification) and fixed effects for time, grade, and cohort. The second added in a random intercept for the school attended, and the third included both a school and district random intercept. These multilevel models are similar to the fixed effect approach described above, but the inclusion of school and district random intercepts allowed me to estimate the variation in course access attributable to the school and district levels, rather than explaining away this variation. This compliments the fixed effects approach which may more conservatively identify the role of student-level characteristics while explaining away the variance attributable to the nested nature of the data by identifying variation attributable to nesting within schools and districts in a meaningful way (Gibbons et al., 2019). Thus, I discuss results from the fixed effect models when focused on variation by student characteristics but discuss the level of variation attributable to the school and district as modeled in the hierarchical linear models, focusing primarily on the intraclass correlation coefficient (ICC) as the parameter of interest.

Research Question 3: Course Placement Policies and Practices

The third research question builds from the first two by moving from descriptive course access analyses to documenting the policies and practices used that might help to explain course access patterns observed. Drawing on the interview transcripts, I conducted two rounds of qualitative coding using Dedoose, Version 8. The first round I applied a single parent code "course placement". This code was applied to all excerpts that mentioned ways in which it was determined or decided what courses an arriving secondary-age immigrant EL-classified student would be placed in.

I then conducted a secondary round of coding that was more inductive in nature, reviewing the coded excerpts while allowing themes to emerge as eventual child codes (Miles et al., 2018). These child codes included: formal policy, practice, factors informing placement decisions, core content, English language development, challenges, and strong practices or successes. Two important code definitions to highlight are *formal policy* and *practice*. While this is not a perfect dichotomy, the code conceptualizations were informed, in part, by the work of scholars who have examined education policy implementation and attempted to distinguish between *formal policies* and *practices that reflect agents' policy interpretation* (Bray & Russell, 2016; Lipsky, 2007; Mavrogordato & White, 2017; Rigby et al., 2016; White & Mavrogordato, 2019). I code excerpts as *formal policy* if the interviewee is describing a standardized step in course placement that has been developed at the district or school level and is generally perceived as being applied across arriving immigrant students. One example, which is described in the findings section, is a policy that determines what proportion of a student's day will be spent in an ELD class period. Typically, the *formal policy* code was applied when an interviewee explicitly used the word "policy" to describe a step, or used terms such as rules or guidelines. In contrast, the *practice* code was applied when interviewees described the course placement process in ways that referenced individual processes, choices, and adaptations of policies. Typically, this code was applied when interviewees described steps they or their colleagues took in ways that conveyed a sense of individual interpretation of what was appropriate.

Research Question 4: Newcomer Programs and Course Access

For the fourth research question, I estimated the impact of participating in a newcomer program on the likelihood that a student was enrolled in an ELA, math, science, or social studies course. I used coarsened exact matching (CEM) as a preprocessing step, matching the two groups (students in newcomer programs in their first year and those not) on observable district- and student-level variables, as measured in students' year of arrival. Matching relies on the researcher to match

treatment and control groups such that they are reduced to more comparable groups, based on observable characteristics (Iacus et al., 2012).

I chose matching variables that I theorize are predictive of being in the treatment group (enrolled in a newcomer program) and correlated with the outcome (course access). For student-level variables, I required exact matches on grade of arrival, race/ethnicity, SIFE, and eligibility for free/reduced price lunch. I also matched on initial ELP, coarsened into bins of the minimum screener value (0) or above (>0).[10] I also matched on district percent of students who were eligible for free/reduced price lunch and district percent of students identified as recent immigrants, the first coarsened into quartiles and the second into below the median (<.01) or above the median (>=0.01). As discussed in the Literature Review section, students' socioeconomic status, race/ethnicity, academic histories and exposure to the English language are linked to differential educational pathways and outcomes, and thus are critical matching variables. Grade of arrival is also important, given that enrollment patterns may vary based on when students arrive in relation to their expected graduation year. District context also likely matters, as districts with larger immigrant populations may have stronger support systems, while districts with larger low-income student populations may face compounding challenges in offering educational opportunities due to constricted resources and competing needs.

I specified matching of many to one. Matching was done using the Matchit package in R (Ho et al., 2011). The sample size was very small in the fourth year observed (with fewer than 20 students who participated in newcomer programs) so I only conducted analyses for students' first three years in U.S. schools. Once the CEM weights were generated, any unmatched student was dropped from the dataset. As seen in Table A5, which presents the weighted average covariate values for the year one, two, and three samples by treatment and control, there is tight alignment across control and

[10] Due to both sample sizes and the range of screener types, I could not match more precisely with confidence.

treatment groups on matched variables, although small, sometimes statistically significant differences remain in district and school-level covariates. This sample includes all districts that enrolled arriving immigrant EL-classified students, regardless of whether they offered a newcomer program.

Using the matched sample, I fit a series of linear probability models regressing course enrollment indicators on newcomer program participation for each of the three years observed in the data. I fit and present three model specifications for each subject area and year. The first specification includes student-level covariates of whether a student was identified for free/reduced price lunch, ever identified for special education, race/ethnicity, SIFE, initial ELP, and grade of arrival, as well as district- and school-level proportions of free/reduced price lunch, immigrant, and ever-EL students. The second includes the same covariates but adds in cohort fixed effects. The third model, which is my preferred model, adds in district fixed effects. The district fixed-effects model specification is:

$$Y_{it} = \beta_0 + \beta_1 (NP_i) + \boldsymbol{X}_i + \boldsymbol{V}_s + \Gamma_c + \Delta_d + \varepsilon_{it}$$

where Y is a binary indicator, for student i in school s and cohort c in district d in year t, for whether a student is enrolled in a given core content area, and β_1 is the estimated relationship between newcomer program participation and the likelihood of core content enrollment. Covariates include the vector of student covariates ($\boldsymbol{X}_i$) outlined above, as well as school covariates ($\boldsymbol{V}_s$) and cohort (Γ_c) and district (Δ_d) fixed effects. In all models I cluster standard errors at the district level. All models were estimated using the fixest packages in R (Bergé, 2018).

I focus my discussion on results from two models—the cohort fixed effects model and the cohort and district fixed effects model. The two approaches answer slightly different questions, given that the inclusion of fixed effects changes the variation of interest to within fixed effect units (Gibbons et al., 2019). The model that includes cohort fixed effects but not district fixed effects allows me to estimate the impact of newcomer program participation on the probability of course enrollment comparing matched students across districts that both offer and do not offer newcomer programs.

The inclusion of district fixed effects changes this comparison to estimating the impact of newcomer program participation only within districts that offer newcomer programs where some arriving students are placed in these programs and others are not. The benefit of including district fixed effects is accounting for district characteristics that may impact students' probability of course enrollment. However, the limitation that comes with including district fixed effects is that I am only identifying the effect of newcomer program participation on course access off of variation within districts that offer newcomer programs, a limited sample from an already relatively small sample of students.

As with the inclusion of different fixed effects, matching also comes with benefits and limitations. CEM is an approach that reduces the treatment and control groups to be more comparable on observables prior to intervention (Iacus et al., 2012). A key limitation, though, is that regardless of the matching specification, there is no way to account for unobserved variables or factors that may be impacting selection into treatment or the outcome of interest. As is discussed further in the Limitations section, newcomer programs are largely designed to support students with gaps in their educational history and with low initial ELP (MAEC, 2019). Thus, it is tenable that a student's educational history (as proxied by the SIFE indicator) and initial ELP would be the most salient predictors of newcomer program participation. I include both as matching variables. Other included matching variables such race/ethnicity and eligibility for free/reduced price lunch attempt to account for other ways in which local actors may be making decisions about newcomer program placement or course placement. The decision to include these variables is grounded in theory and prior research identifying those as important factors in shaping students' educational opportunities and outcomes (Kelly, 2009; Suarez-Orozco et al., 2010; Tyson, 2011). While I am confident that the matching approach is creating more comparable groups and I am including the best set of covariates given data limitations, I cannot fully rule out the threat of omitted variable bias. As such, I am cautious in the interpretation of the estimated effect of newcomer programs. I frame these as tenative estimates—a first approach at answering an

important question while limited in the ability to firmly isolate the causal effect.

I suspect that bias, if present, comes as students who are placed in newcomer programs are those who would have been less likely to be enrolled in the courses examined. Given that newcomer programs are intensive interventions typically framed as only appropriate for students who it is thought are not "ready" for traditional educational services, I'd anticipate these students are likely those who would be less likely to be placed in core content. In this way, selection issues may arise from school administrators' decisions to recommend newcomer program placement to students based on unobservable characteristics that also lead to lower core content enrollment levels. Selection issues may also arise from students who themselves are more interested (or disinterested) in the newcomer program strcuture or family preference about whether their student should particpate. If biased, the estimates in this paper are likely overstating negative effects or masking positive effects through a downward bias.

Newcomer Program Robustness Checks

To strengthen my conclusion about newcomer program effects, I conducted four robustness checks. The first two were more restrictive matching approaches. I first re-ran the models with a sample that required an exact match on district, as well as the other student-level covariates specified above. I then required binned matches on school-level proportions of recent immigrant students and students eligible for free/reduced price lunch, rather than district-level. The third was less restrictive than the main approach, only requiring matches on student-level covariates. The fourth was the least restrictive, as I ran the same models, but with the full, unmatched sample. Less restrictive matching specifications are more generalizable and draw on more variation, while more restrictive approaches potentially account for more bias, but restrict the sample. For each of these robustness checks, I only present the two main specifications (cohort fixed effects and then cohort and district fixed effects).

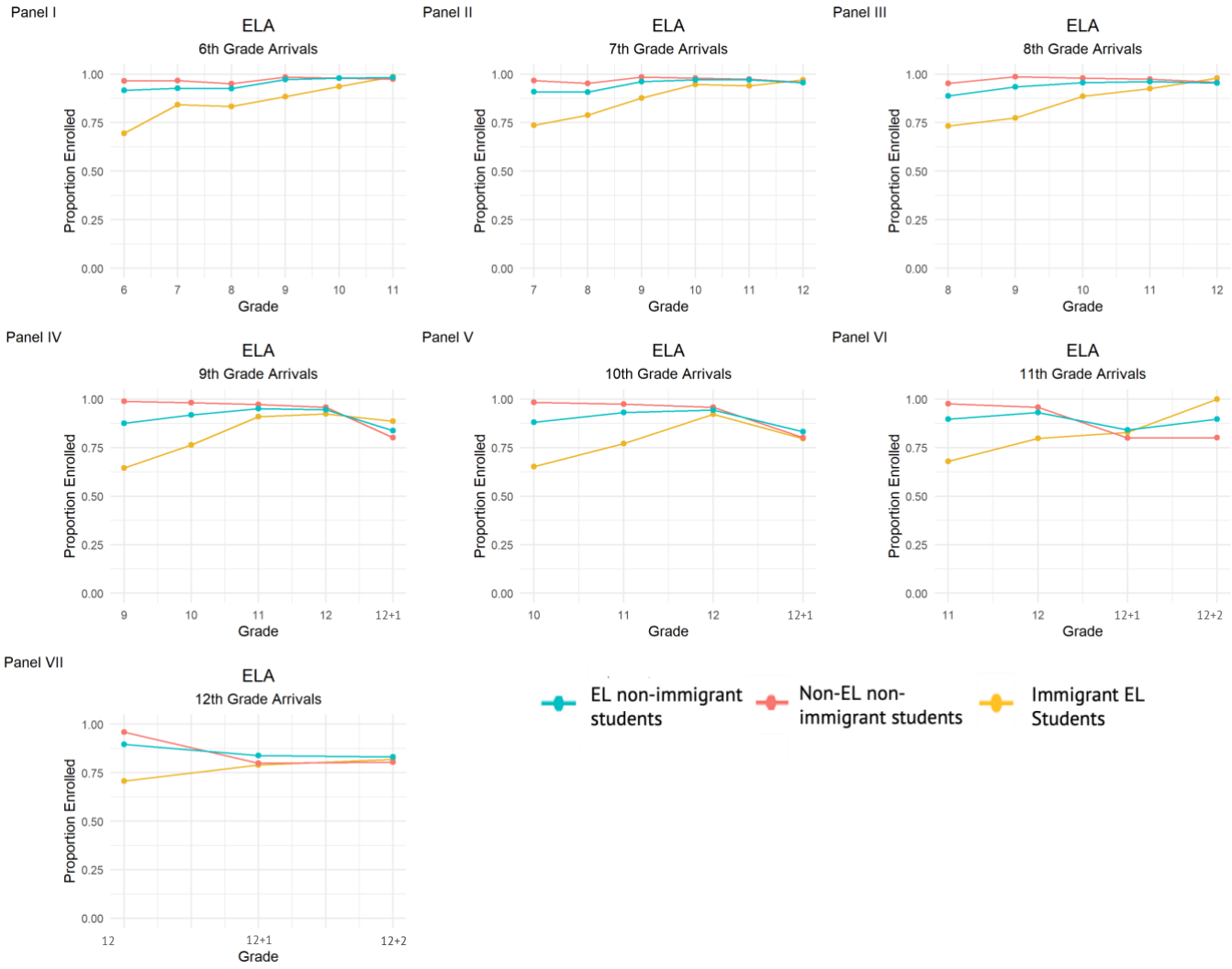
Findings

Research Question 1: Descriptive Comparisons of Course Access

Overall, immigrant EL-classified students who arrived in grades 6-12 had lower enrollment rates in ELA, science, and social studies than their EL and non-EL peers in the year they arrived. Gaps tended to narrow over the following two years, although in many cases enrollment remained lower through the third year after arrival. Arriving immigrant EL-classified students were also enrolled in math at lower rates, but the gaps were smaller and closed in most instances the following year. Across subjects there was heterogeneity in enrollment rates by grade of arrival. In Figures 1.1 and 1.2 I present ELA and math enrollment rates over time for immigrant EL-classified students by their grade of arrival, in comparison with trajectories for students who were EL-classified in that grade but not recent immigrant arrivals and students who were not EL-classified in that grade. Math and science results are presented in Figures A2 and A3.

As seen in Figure 1.1, regardless of grade of arrival, immigrant EL-classified students were enrolled in ELA coursework at substantially lower rates in the year they arrived than both students who were EL-classified in that year and those who were not. With the exception of 12th grade arrivals (Panel VII), there was also a large gap in the following year as students progressed through grades. For students who arrived in 6th-8th grades (Panels I, II and III) the gaps persisted through middle school and into high school, narrowing or closing around 10th grade. For students who arrived in 9th, 10th, and 11th grade (Panels IV, V and VI) the ELA enrollment gap was present for two years before narrowing or closing. It is not uncommon for later arriving immigrant EL-classified students to stay beyond their 12th grade year. For these students, we see that immigrant EL-classified students were enrolled in ELA at the same or higher rates than students who were EL-classified in their same grade of arrival as well as students who were not.

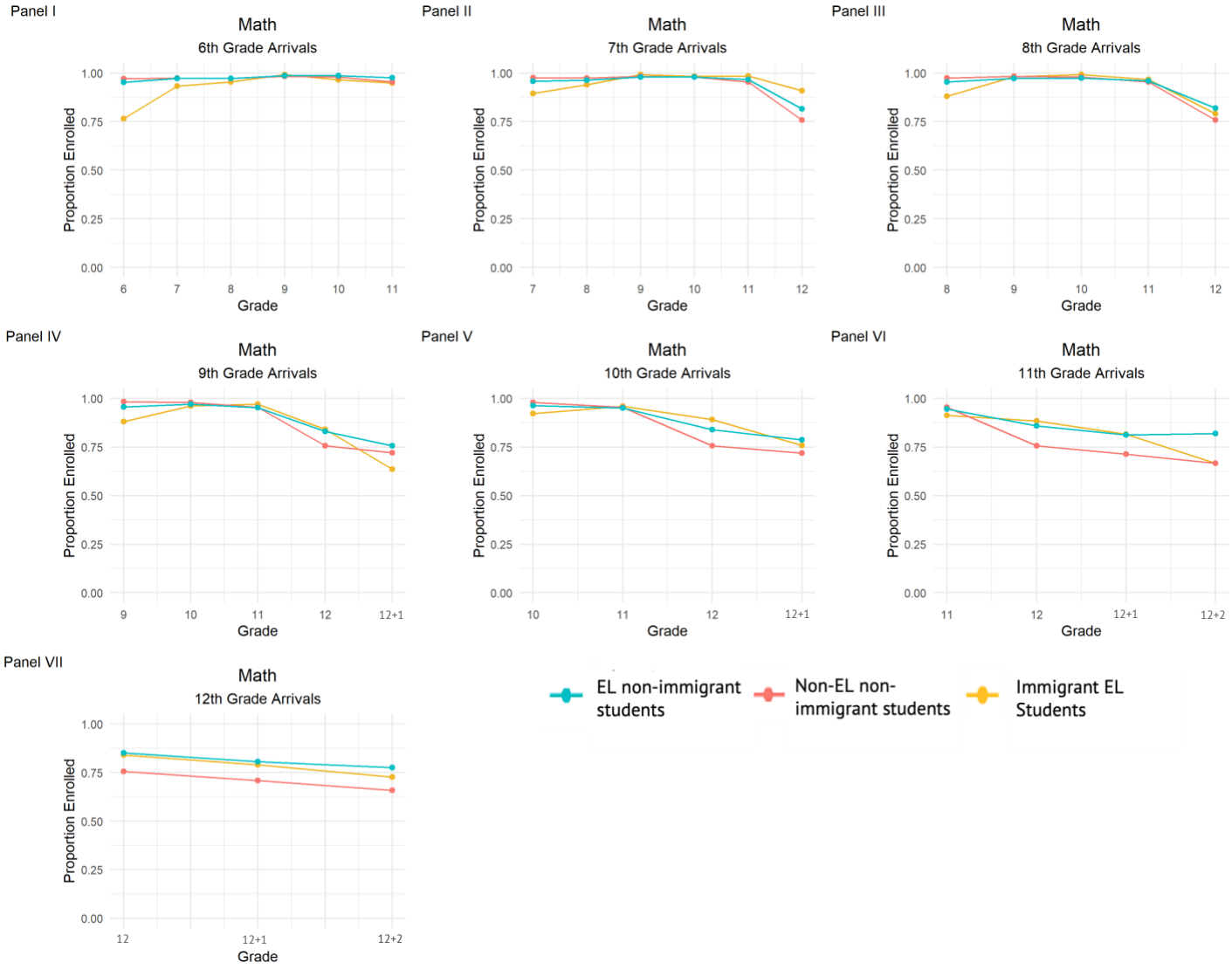
Figure 1.1 ELA Course Access for Arriving Immigrant EL-Classified Students, EL-Classified Students, and Students who are neither EL-Classified, nor Arriving Immigrant Students



Of the four subjects examined, math was the core content area where there were the smallest differences in enrollment rates for immigrant EL-classified students and their peers. As plotted in Figure 1.2, Panel I, 6th grade arrivals had the lowest rate of enrollment in math across grades of arrival, about 75% in comparison with almost 100% of students who were EL-classified in 6th grade and students who were not EL-classified in 6th grade. Students who arrived in 7th-12th grade all had similar enrollment rates upon arrival, around 85%, which was slightly lower than their peers for all arrival grades other than 12th grade arrivals, who actually had higher enrollment rates than students who were not EL-classified in 12th grade and comparable rates to EL-classified students. By the second year after arrival, for almost all immigrant EL-classified students the gaps between their enrollment levels and the two comparison groups had either shrunk to a few percentage points or closed. In later observations, students who had arrived in 10th-12th grade had, in many instances, higher enrollment rates than students who were not EL-classified in the same grade they had arrived in.

Immigrant EL-classified students were uniformly less likely to be enrolled in science upon arrival, across grades of arrival (Figure A2). Gaps were most striking for 6th, 9th, 10th, and 11th grade arrivals. For example, only about half of arriving 9th grade students were enrolled in science upon arrival, in comparison with about 85% of students who were EL-classified in 9th grade and almost 100% of students who were not EL-classified. For 6th and 7th grade arrivals, enrollment gaps narrowed as they progressed through grades, but did not close until 10th grade, while for 8th and 9th grade arrivals enrollment levels converged at 11th grade. Unlike ELA, there is no evidence that immigrant EL-classified students who stayed beyond their 12th grade year were more likely to take a science course than other non-immigrant students who also continued past their 12th grade year.

Figure 1.2 Math Course Access for Arriving Immigrant EL-Classified Students, EL-Classified Students, and Students who are neither EL-Classified, nor Arriving Immigrant Students



As with the other subjects examined, social studies enrollment upon arrival was substantially lower among arriving immigrant EL-classified students than students who were EL-classified in the same grade, with gaps ranging from about 20 percentage points (6th grade arrivals; Figure A3) to about eight percentage points (12th grade arrivals; Figure A3). The gaps were larger in comparison with non-EL students. For all grades other than 12th grade arrivals the gaps in enrollment rates shrunk, but persisted the following year, and into a third year for 6th and 9th grade arrivals. In almost all instances immigrant EL-classified students had closer enrollment levels to students who were EL-classified in the same grade as the immigrant students arrived, while students who were not EL-classified had continuously higher enrollment levels, although rates became more similar over time.

Research Question 2: Differences in Course Access Among Secondary Age Arriving Immigrant EL-Classified Students

While comparisons to students who were not arriving immigrant EL-classified students provide important context for how enrollment patterns for arriving immigrant EL-classified students differ in comparison with their peers, there is incredible diversity among immigrant students. It is critical, in the interest of probing potential inequities, to examine how access differed among immigrant EL-classified students. I examined the relationship between student characteristics, including race/ethnicity, SIFE, initial ELP, and eligibility for free/reduced price lunch, and course enrollment. I first look at raw differences, then use a regression framework to identify the degree to which student characteristics predict differences in access when accounting for differences attributable to grade of arrival, district, school, and time.

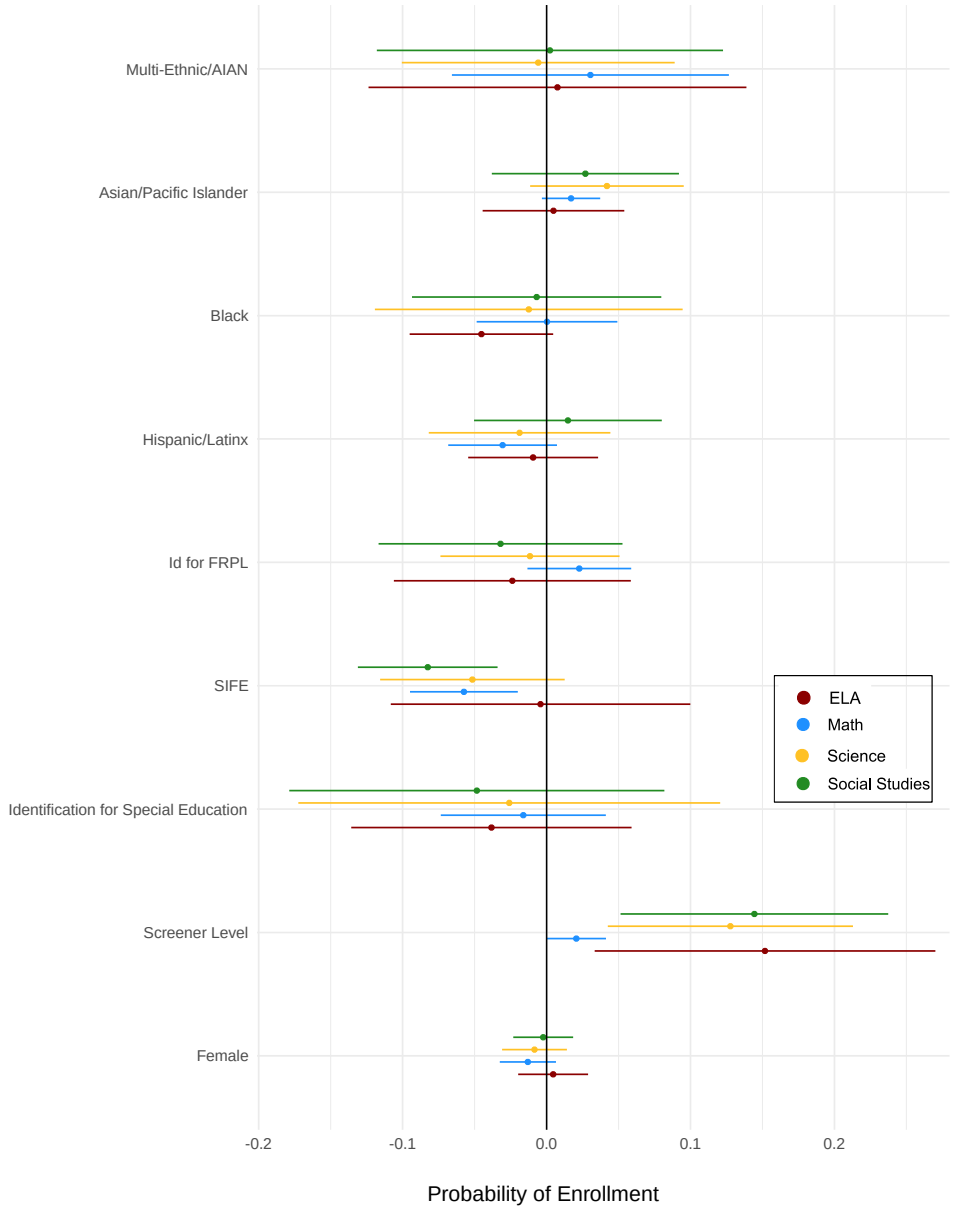
In examining raw differences in the proportion of students enrolled in core content areas, I found gaps in average enrollment rates by student socioeconomic status, race/ethnicity, academic history, and initial ELP. Figures A4 and A5 plot the raw differences in course enrollment between students who were eligible for free/reduced price lunch and those who were not (Figure A4, Panel I),

by race/ethnicity (Figure A4, Panel II), by SIFE identification (Figure A5, Panel I), and whether students' initial ELP was the minimum screener level or above (Figure A5, Panel II).

Overall, students eligible for free/reduced price lunch were enrolled in core subjects at lower rates in their arrival year and the following year than students not eligible for free/reduced price lunch (Figure A4, Panel I). These gaps persisted in ELA enrollment over the time students were enrolled, although shrunk in later years, while gaps in science and social studies persisted in all years except the fourth year after arrival in science and the fifth year after arrival in social studies. Course enrollment gaps were smaller in math. From arrival through the third year after arrival, Black and Hispanic/Latinx students were enrolled in ELA, science, and social studies at lower rates than students from other racial/ethnic groups (Figure A4, Panel II). Black and Hispanic/Latinx students were also less likely to be enrolled in math classes upon arrival and in the following year, although, as with gaps by free/reduced price lunch eligibility, differences were smaller. SIFE were less likely to be enrolled in the subjects examined for the first three years of enrollment, while in the final year observed differences were much smaller and even reversed in ELA (Figure A5, Panel I). Finally, differences were also observed by initial ELP, as students whose ELP was lowest upon arrival were less consistently likely to be enrolled across subjects than students assessed above the lowest level (Figure A5, Panel II).

While raw differences are important on their own, it is important to understand how course access varies by these characteristics once holding other characteristics constant. For students who arrived from 2015/16 onward, I fit linear probability models estimating the relationship of the above characteristics with the probability of enrollment in ELA, science, and social studies classes, including fixed effects for grade of arrival, district, and year in school in the first four years of arrival. Point estimates with corresponding 95% confidence intervals are plotted in Figure 1.3 and presented in Table 1.3. While some raw differences held, others were not significant predictors of course access.

Figure 1.3 Relationship between Secondary-Age Immigrant EL-classified Students' Characteristics and Course Enrollment, 2015/16-2018/19



Note. Figure represents the average relationship between the listed characteristics and the probability of enrollment in a given core content area from linear probability models that predict course access from the listed characteristics as well as fixed effects for grade of arrival, district, and year in school for 2,298 students who arrived in grades 6-12 from 2015/16-2018/19 with complete data (5,156 student-year observations). ID for FRPL=Eligible for free/reduced price lunch, SIFE=Student with interrupted formal education, AIAN=American Indian/Alaska Native. Screener level=Initial ELP level is above the minimum.

Overall, SIFE identification and initial ELP ("Screener Level") were the only significant predictors of the probability of being enrolled in a given core content course, and the significance level was not constant across core content subjects. Holding constant other student characteristics,

SIFE identification was associated with a lower probability of enrollment in a math and social studies course, with an estimated 5.74 percentage point decrease in the probability of being in a math class, and an 8.25 percentage point decrease in the likelihood of being in a social studies class. Students whose ELP scores were above the lowest screener level were significantly more likely to be enrolled in all core content areas other than math, with the largest magnitude reaching a 15.18 percentage point increase in the probability of being in an ELA course, while there was an estimated 14.4 percentage point increase in the probability of being in a social studies course and 12.8 percentage point increase in the probability of being in a science course.

Table 1.3. Relationship between Student Characteristics and the Probability of Enrollment across the First Four Years of Enrollment, 2015/16-2018/19

	ELA	Math	Science	Social Studies
SIFE [95% CI]	-0.004 [-0.108; 0.100]	-0.057** [-0.095; -0.020]	-0.052 [-0.116; 0.013]	-0.083** [-0.131; -0.034]
Initial ELP [95% CI]	0.152* [0.033; 0.270]	0.021. [0.001; 0.041]	0.128** [0.043; 0.213]	0.144** [0.051; 0.237]
Multi-Ethnic/AIAN [95% CI]	0.008 [-0.124; 0.139]	0.0305 [-0.066; 0.127]	-0.006 [-0.101; 0.089]	0.002 [-0.118; 0.123]
Asian/PI [95% CI]	0.005 [-0.044; 0.054]	0.017 [-0.003; 0.037]	0.042 [-0.011; 0.095]	0.027 [-0.038; 0.092]
Black [95% CI]	-0.045. [-0.095; 0.005]	0.001 [-0.049; 0.049]	-0.012 [-0.119; 0.095]	-0.007 [-0.094; 0.080]
Hispanic/Latinx [95% CI]	-0.009 [-0.055; 0.036]	-0.031 [-0.068; 0.007]	-0.019 [-0.082; 0.044]	0.015 [-0.050; 0.080]
FRPL Eligible [95% CI]	-0.024 [-0.106; 0.059]	0.023 [-0.013; 0.059]	-0.012 [-0.074; 0.051]	-0.032 [-0.117; 0.053]
Identified for Special education [95% CI]	-0.038 [-0.136; 0.059]	-0.016 [-0.074; 0.041]	-0.026 [-0.172; 0.121]	-0.048 [-0.179; 0.082]
Non-Male [95% CI]	0.005 [-0.020; 0.029]	-0.013 [-0.033; 0.007]	-0.008 [-0.031; 0.014]	-0.002 [-0.023; 0.018]
Grade FE	X	X	X	X
Yr. in Sch FE	X	X	X	X
District FE	X	X	X	X
N	5,156	5,156	5,156	5,156

Note. *p<.05, **p<.01, ***p<.001. N represents the total student-year observations in the model. Estimates are accompanied by 95% confidence intervals. EL=English learner. FRPL=Free/reduced price lunch. SIFE= Student with interrupted formal education. ELP=English language proficiency, AIAN=American Indian/Alaska Native.

I also regressed the probability of course enrollment on student characteristics while including random effects for both the school and district students attended. Tables A3 and A4 present coefficients from the mixed effects models predicting course access. Both tables present coefficients from a series of models, including a specification that does not account for students' school or district (Model I), a specification with a school-specific intercept (Model II), and a specification that has both a school- and a district-specific intercept (Model III), Table A3 presents coefficients from models predicting the probability of ELA and math enrollment, and Table 4, science and social studies enrollment. As evidenced by the ICCs calculated from models that include both school and district random intercepts (Model III), the school a student attended consistently explained more variance in the probability of being in a given core content course than the district a student was in. For example, for ELA enrollment the school intercept ICC was 0.27. This suggests that, as modeled, the school level accounted for about 27% of the variation in probability of ELA enrollment. This jumps to 30% for math, then declines to 17% for science and 22% for social studies. In contrast, district-level ICCs ranged from 2% to 11%. This suggests that there was a large degree of variability between schools in immigrant EL-classified students' course access, while districts played less of an explanatory role.

In sum, raw differences highlight how students' course access differed by student characteristics such as race/ethnicity, SIFE, socioeconomic status, and initial ELP. Modeled differences in course access further details how access to core content coursework was largely predicted by SIFE and initial ELP once accounting for other factors, while other raw differences by student characteristics were not significant when modeled. The school a student attended plays a large role in explaining variation in the probability of course enrollment, while the district a student attended holds less of an explanatory role. Importantly, these descriptive results are not estimates of the "effect" of student characteristics on enrollment, but rather patterns that likely capture some way in which

districts and schools, and those working within them, are responding to these characteristics that results in constricted course access. The next section explores the role of school and district policies, as well as individuals' practices, in shaping course access for secondary age immigrant EL-classified students.

Research Question 3: Course Placement Policies and Practices

Drawing on interviews with 11 staff from six school districts in Oregon, I analyzed how those working closely with course placement decisions for immigrant EL-classified students in middle and high school settings described the policies and practices that shape course placement for secondary-age immigrant EL-classified students. Overall, responses suggested that there were some formal policies in place that shaped the placement process, but that course placement was predominantly a process governed by individuals' practices. This resulted in an approach that one federal programs coordinator described as "different students having different experiences depending on what school they go to". Even in a district context where responses suggested there was a higher degree of policy formalization, the EL program director described how "We have policy, we have guidelines, we have steps we take…then, in the end, it's really up those of us who work most closely with the students…"

In this section, I synthesize themes that emerged, focusing first on themes that emerged when interviewees were discussing the limited formal policies that shaped course placement for arriving immigrant EL-classified students. I then describe the practices discussed.

Formal Policy: Prioritizing ELD and Math Classes in Schedule-Building

There was a clear policy focus on creating schedules that prioritized ELD and math content access for arriving immigrant EL-classified students. For example, interviewees from five of the six districts described formal policies around how much of an arriving immigrant EL-classified students' day would be comprised of ELD classes. In two districts, interviewees described a half-day of ELD as the standard approach for arriving immigrant EL-classified students. In contrast, interviewees from three

districts described a standard policy of enrolling students in one ELD course, although an interviewee from one of those districts did describe how they would try to "get students extra courses of ELD if possible" as part of their approach. Another interviewee described how students arriving in grades 9-12 were placed in a double block of ELD in their school, while students arriving in grades 6-8 were placed in a mixed-age ELD block from morning till lunch as part of the newcomer program.

When describing policies, interviewees from three districts also discussed how math was the one core class prioritized in the development of schedules. For example, one EL program coordinator described the policy of placing all arriving students in a math class and explained, "They will be in a math class. There's just no way around it, you will be in a math class. Because we would be doing you a disservice if we didn't put you in a math class". An additional two interviewees referred to math as a "universal language" when describing how it was often the core class that students were placed in upon arrival, regardless of English language proficiency.

Formal Policy: Collecting Information on Academic Histories to Inform Course Placement

There were also policies focused on gathering information about students' academic histories to determine the level of coursework a student would be placed in. Interviewees from all districts described policies in place that required students' transcripts be evaluated to determine what credits could be awarded and, in turn, what courses to place students in. However, while this was described as a formal overarching policy, it varied across districts how strictly the approach to transcript translation and evaluation was guided by policy and how much instead was guided by individual decisions and available resources. For example, one interviewee described how the district contracted the work out to an outside organization to translate transcripts and determine credit equivalencies. In contrast, another described how their district relied on other staff who spoke the language to try to translate the transcripts and determine equivalencies. Transcript evaluation policies that didn't include formal policy guidelines and supports were often seen as creating challenging circumstances, as there

was confusion about which classes could be counted towards graduation requirements, what the equivalencies were between classes taken and U.S. classes, and overall translation difficulties. This resulted in variability in course placement, as confusion around what courses could be counted toward graduation informed what courses students were placed in.

Not all policies focused on gathering information on prior academic experiences were focused on transcript evaluation. Outside of transcript translation and evaluation, interviewees from two districts described how their district policy was to use short math assessments to determine which math course students would be placed in, while another described how their intake process included questions about students' academic histories. Together, policies focused on gathering information on students' academic histories guided course placement, as decisions were made in response to prior learning, as well as what courses counted towards those necessary for graduation.

Practices: Course Placement Decisions in Response to Perceptions of Students' Needs and Readiness

One practice that shaped course placement was making decisions in response to perceptions of students' needs and readiness. Many interviewees expressed that they often felt students would be overwhelmed in core content coursework. They cited a lack of academic skills as well as English proficiency as reasons they would not place students in certain core content courses. As an example of this, the Homeless/Migrant Education Coordinator in one district described how their team approached placement:

> We know that their language is low, so we try to put them in classes where they will interact and just have…that experience with the language…And so putting that student directly into, say a 7th grade social studies class when they won't understand a word of what the teacher is saying is not ideal. That's not going to actually help that student.

Similarly, another interviewee described how many staff in their district had "the perspective of, we just get them immersed in language that first year, we don't worry about core classes", while another

described how they "didn't want to lie", but arriving immigrant EL-classified students would not be enrolled in core content classes until it was determined they could be "released out" from a schedule that was comprised of ELD classes, physical education, and electives when they had developed sufficient English proficiency to participate. Another interviewee described how they recognized students needed core content coursework to graduate, yet they didn't want to push core content too hard, stating, "We don't want to lose these students either. We want to be able to support them in a way where they do graduate, but which path makes the most sense?"

Practices: Course Placement Decisions in Response to Student and Family Input

Another practice focused on making course placement decisions in response to student and family input. While no interviewees described a formal way in which student and family input on course placement was integrated into the course placement process, there were examples of practices that were discussed. The most common was informal conversations during intake. For example, one counselor described how their approach to course placement was shaped in response to students' stated postsecondary goals, sharing,

> If they really are pushing for four-year university or even community college, I'm always telling them they're more than eligible and we have all these resources to help them get there. So, then I'm wanting to make sure they're taking the right classes to be on track to be able to do that.

This was the only interviewee who brought up students' postsecondary goals as informing the course placement process, as others shared that they asked students about their interests to determine elective placement. Of note, this quote captures how integrating student goals into course placement relies on students' expressing an interest in this path—with no mention of systematically or structurally working this exchange of information into the course placement process. Similarly, family input was talked about sparingly, and often was more to ascertain students' academic histories, rather than ask for input into how schedules could be developed to meet student or familial goals around educational opportunities. However, in select interviews, it was talked about how families brought up educational

goals during the enrollment process. One interviewee's description highlights how this varied by student and family background, saying "…the education level of the parent and their own exposure to educational opportunities, I think, drives that conversation…it just really does depend on sort of where that family was educationally when they were in their home country".

Practices: Course Placement Decisions in Response to Local Capacity

Finally, course placement decisions were shaped in response to perceived local resources available to support students in core content coursework. It was often described that those working in placement felt many core content teachers did not have the training to support the diversity of linguistic backgrounds or socioemotional needs, therefore students were not placed in those classes. This highlights how limited capacity constricted placement options. As another way in which capacity shaped course placement, in one district that had limited EL support staff one interviewee described how they would "clump" EL-classified students into certain core classes so that they could direct EL support staff to target those classes. This also limited which courses they would put students in.

Overall, course placement was discussed as a challenging process for interviewees and their districts and schools. A complex set of factors informed both policy and practice, but ultimately there was a large degree of individual discretion. The resulting process was described as varying widely across students. In discussing the different factors that went into the individual decisions, one interviewee described it as "a balancing act. And we don't always get it right".

Research Question 4: Newcomer Programs and Course Access

I estimated the impact of participating in a newcomer program on ELA, math, science, and social studies enrollment outcomes for students arriving in 2015/16 or later. I focus on results from the analyses that use a smaller subset of students who make up more comparable treatment and control groups for the first three years of enrollment, as determined through the CEM pre-processing approach. I discuss the cohort-adjusted specification (Figure 1.4, Panel I and Table 1.4, Model II) and

the district- and cohort-adjusted specification (Figure 1.4, Panel II and Table 1.4, Model III). I also highlight results from robustness checks, which include the full, unmatched sample, as well as samples that came from different matching setups.

Table 1.4. Newcomer Program Participation and Course Enrollment, Matched Sample

		Model I: Cov. Adj			**Model II: Cov. Adj. w/Chrt FE**			**Model III: Cov. Adj. w/Chrt & Dist. FE**		
		Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
ELA										
	Newcomer Program [95 % CI]	-0.303** [-0.490; -0.116]	-0.107 [-0.246; 0.031]	-0.060 [-0.207; 0.087]	-0.274** [-0.458; -0.089]	-0.084 [-0.218; 0.051]	-0.053 [-0.199; 0.093]	-0.175* [-0.340; -0.011]	-0.014 [-0.083; 0.054]	-0.045 [-0.154; 0.064]
Math										
	Newcomer Program [95 % CI]	-0.024 [-0.126; 0.078]	0.017 [-0.034; 0.068]	-0.022 [-0.064; 0.020]	-0.033 [-0.140; 0.074]	0.016 [-0.040; 0.071]	-0.020 [-0.061; 0.020]	-0.025 [-0.111; 0.061]	0.009 [-0.024; 0.042]	-0.075** [-0.126; -0.023]
Science										
	Newcomer Program [95 % CI]	-0.149 [-0.385; 0.087]	-0.083 [-0.224; 0.058]	-0.005 [-0.146; 0.137]	-0.148 [-0.389; 0.093]	-0.084 [-0.229; 0.062]	0.009 [-0.121; 0.138]	-0.043 [-0.248; 0.161]	-0.101 [-0.273; 0.071]	-0.013 [-0.179; 0.152]
Social Studies										
	Newcomer Program [95 % CI]	-0.211. [-0.435; 0.013]	-0.097. [-0.198; 0.004]	-0.125* [-0.238; -0.011]	-0.194. [-0.418; 0.031]	-0.090. [-0.180; -0.0005]	-0.121* [-0.232; -0.010]	-0.134 [-0.368; 0.101]	-0.080 [-0.180; 0.020]	-0.218* [-0.380; -0.056]
	Cov.	X	X	X	X	X	X	X	X	X
	Chrt. FE				X	X	X	X	X	X
	Dist. FE							X	X	X
	N	956	527	289	956	527	289	956	527	289

Note. .*p*<.01, **p*<.05, ***p*<.01, ****p*<.001. N represents the total student observations in the model. Estimates are accompanied by 95% confidence intervals. Results are presented for three model specifications, each estimated for each separate subject and year in the data. Cov.= model that includes student, school, and district-level covariates. Chrt FE=Cohort fixed effects, Dist. FE=District fixed effects.

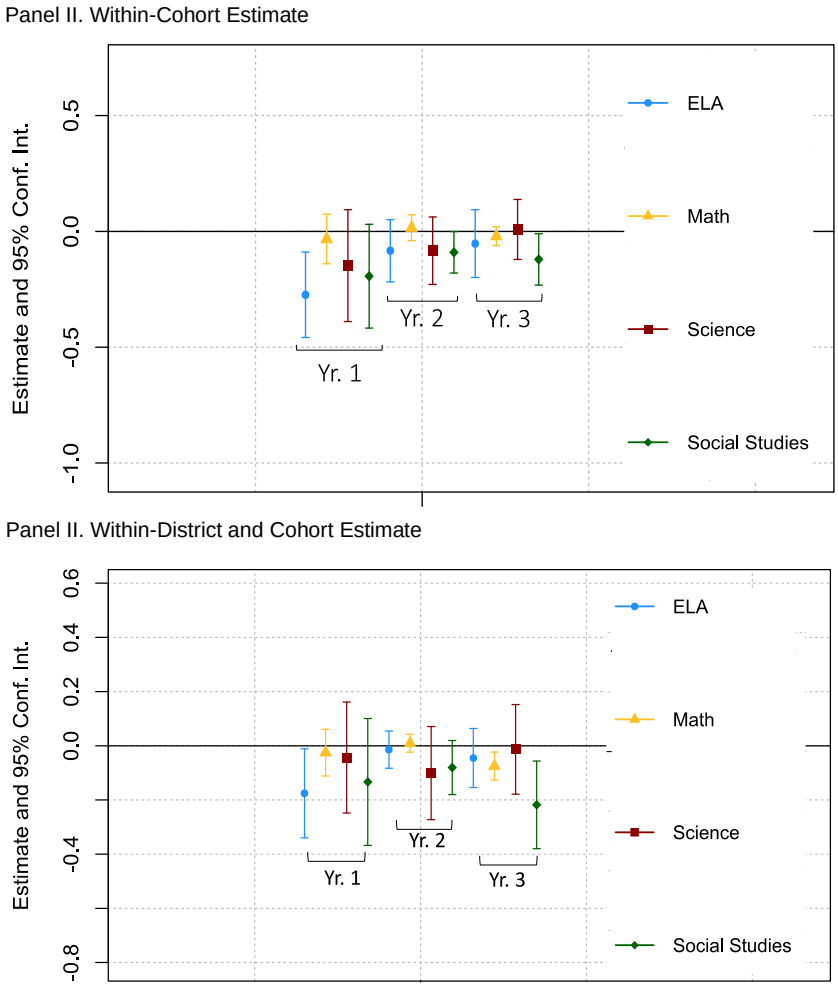
Among students in the matched sample, the estimated impact of newcomer program participation on ELA enrollment was consistently negative in the year of arrival across specifications. In the specification that draws on variation across districts while accounting for cohort-specific effects (Model II), the likelihood of ELA enrollment dropped by 27.4 percentage points in the year of arrival, although confidence intervals span moderate to large effects. In the cohort and district-adjusted specification (Model III), newcomer program participation had an estimated effect of -0.175 [95% CI:

-0.340; -0.011] on the likelihood of ELA enrollment in the year of arrival. This translates into an estimated decline in the probability of enrollment in an ELA course of 17.5 percentage points within districts that offer newcomer programs, although the confidence interval spans small negative estimates to much larger negative estimates. Estimates of the relationship between newcomer program participation and ELA course enrollment among the matched samples were also negative in years two and three across specifications, although coefficients were much smaller and not statistically distinguishable from zero.

There was no consistent, significant effect of newcomer program participation on the probability of enrolling in a math course. Coefficients on newcomer program participation were consistently small in magnitude, with confidence intervals that spanned small to modest negative or positive estimates. The exception, however, was the within-district (Model III) year three estimate, where there was a significant decrease in the probability of math enrollment, -0.075 [-0.126; 0.023]. While the cohort-adjusted specification that drew on cross-district variation was also negative, it was smaller in magnitude and not statistically significant.

There was no significant impact of newcomer program participation on the probability of being enrolled in a science class. Estimates were negative and imprecise in all years and specifications, with confidence intervals that spanned large negative estimated effects to small or moderate positive estimated effects. There was also no significant estimated impact of newcomer program participation on social studies enrollment in the first two years observed. However, there was a significant negative estimated effect in year three across all three model specifications. The estimated effect was largest in the within-district specification (Model III), an estimated decline of 21.8 percentage points in the likelihood of science enrollment in year three, -0.218 [-0.380; -0.056], while the estimated decline in the cohort-adjusted specification (Model II) was about 12 percentage points, -0.121 [-0.232; -0.010].

Figure 1.4 Newcomer Program Participation and Course Access, Matched Sample



Note. Each figure represents the outcomes for students who arrived in grades 6-12 from 2015/16-2018/19 matched on observables. The year one sample includes 956 students, year two 527 students, and year three 289 students. Estimates are from models that include student and school covariates, as well as cohort fixed effects.

Robustness Checks

Estimates largely aligned across the three alternative matching approaches and the estimates generated using the unmatched sample, although the statistical significance was sensitive to the specification in select instances.

Figure A6 presents estimates from the within-district estimates from samples that followed three different matching specifications, one that required an exact match on district (Panel I), one that matched on school-level covariates (Panel II), and one that matched only on the student-level

covariates (Panel III). Estimates from the exact-match sample differed in that both the year one and year two ELA estimate were similar in magnitude to the main estimate, but both were significantly different from zero. Additionally, the third-year estimate on the probability of math enrollment was not statistically distinguishable from zero. Estimates from the school-level covariate matching specification were very similar to the main estimates, but the negative association between newcomer program participation and social studies enrollment in year three was similar in magnitude, but no longer significant. Matching just on student-level covariates results in estimates that were similar in magnitude, but the negative relationship between newcomer program participation and ELA enrollment was not significant in the first year, and neither was the math estimate in the third year. I also ran the same models, but for unmatched estimates (Panel IV), which drew on the full sample. Unmatched estimates largely aligned, with coefficients aligning with the main estimates as having similar direction and magnitude. However, the negative coefficients on ELA enrollment were not statistically significant in this specification.

Figure A7 presents the same set of robustness checks but drawing on variation across districts that both do and do not have newcomer programs. These estimates also largely align with the main estimates in terms of direction and magnitude, although in more instances estimated coefficients are significantly different from zero. Across all alternative specifications the first-year estimate on ELA enrollment was negative and statistically significant, while in the exact district match there was also an estimated negative effect in years two and three, and in the unmatched estimates there was a negative estimated effect in year three. Math estimates were not significant across specifications, while there was a significant negative estimate on the probability of science enrollment in the second year from the exact district match specification. There was also an estimated negative effect on social studies enrollment in the third year using the exact district match specification, as in the main specification, although the estimate was null for all other specifications.

Limitations

There are key limitations of this work. I did not have access to information on all the student characteristics and experiences that may drive course access and placement in newcomer programs. For analyses focused on differences within the group of immigrant EL-classified students there may be unobserved covariates that were correlated with the outcome and other predictors in ways that resulted in biased estimates. For example, the significant differences attributable to initial ELP and SIFE identification may have been driven by unobserved covariates. However, given the important role of these characteristics identified in prior literature, as well as the steps taken to account for other factors, I argue that these estimates are important and yield critical information on how observed differences in access can be explained by student characteristics.

For research question four, data limitations introduce bias concerns. The validity of the matching design rests on the assumption that the matching approach accounts for issues of selection into treatment. Program designs and descriptions of who programs are for support the idea that students' academic history (as proxied through SIFE) and initial ELP are the two key drivers behind recommendations for placement in a newcomer program within districts that have newcomer programs (MAEC, 2019). However, these blunt measures may not capture important nuances that impact selection into treatment and the probability of course enrollment. For example, SIFE has a standard statewide definition, but the identification process may vary across districts and schools, resulting in imperfect quantifications of students' prior exposure to academic instruction and settings. Further, the use of varying screeners and dichotomization of ELP levels into the minimum and above the minimum in the matching approach may mask important differences in initial ELP. As a result, just relying on these two variables may not have fully captured drivers of selection into treatment.

Beyond blunt measures, there were also other factors I did not have data on. I did not have access to course-taking data from students' previous country of residence, which may have impacted course

access outcomes and newcomer program placement. Additionally, students and parents may have had preferences about newcomer program placement that were correlated with the outcomes. If the estimates were biased, it was likely a negative bias, as other factors that predicted selection into newcomer programs were likely negatively associated with course access. While I still believe the approach has reasonably addressed issues of selection and bias, these concerns lead me to frame these estimates as initial causal estimates of the impact of newcomer program enrollment.

As a further limitation, the sample was relatively small, with a small subset of students served by newcomer programs. Program model data only extended back to 2015/16 in the current categories, and samples were too small to use the matching approach to estimate the effect of newcomer program participation on course enrollment in year four. This made it impossible to see whether newcomer program effects differed when looking over a longer time horizon. The shortened panel also limited the number of students included in the analyses. Combined with high levels of data missingness for students' initial ELP, this mean the sample was relatively small, which can impact the ability to detect statistical significance. Quantitative research on small, structurally vulnerable populations can be challenging, given that many statistical properties and assumptions require large samples. However, this type of research is critically important to conduct (Etz & Arroyo, 2015). In the face of these challenges, I drew on multiple strategies suggested for working with small samples, such as including rich descriptive statistics and visualizations to complement predictive and causal models and use of numerous covariates to reduce variation in the outcome when concerned with a treatment effect (Hopkin et al., 2015).

Discussion

This study examined access to core coursework for immigrant EL-classified students who arrived in Oregon secondary schools. I descriptively explored how access differed in comparison with other EL and non-EL students, as well as how access differed among immigrant EL-classified students

who arrived in secondary grades. I documented administrator and educator perceptions of policies and practices that shaped access for this population and estimated the effect of newcomer programs on course access. The findings bring to light potentially concerning issues regarding access to core content. Enrollment levels in core courses were low in comparison with peers. This suggests that districts and schools failed to structure educational pathways for immigrant EL-classified students that included full access to core content upon arrival. Enrollment levels were disproportionately low for Black and Hispanic/Latinx immigrant EL-classified students, as well as students eligible for free/reduced price lunch, those with low ELP levels, and SIFE. Once modeled, differences in access was consistently predicted by both initial ELP and SIFE identification.

Interview responses suggested that course placement was primarily shaped by individual discretion and decision-making within policy frameworks that prioritized ELD and math access. Many expressed that course placement was done in reaction to perceptions of students' lack of academic readiness and ELP, seeing students as likely to not be successful in the classroom until they were more adjusted and had developed stronger English proficiency. Newcomer programs, while designed as a support, did not lead to fuller course enrollment in any year observed. In contrast, coefficients were consistently negative, with estimates suggesting negative estimated effects on ELA enrollment in the year of arrival, and social studies and math in the third year after arrival. While initial exclusion from ELA may not be as surprising, given that schools may be prioritizing ELD and orientation supports through newcomer programs, exclusion later is more surprising. While I do not observe the exact mechanisms for this observed pattern, it may be that students are not transitioning out of newcomer programs and over time there is more displacement of core content, or it could be that newcomer programs may be leading to students being tracked away from these courses.

Access to core content is a critical measure of equitable access to learning (Callahan & Shrifer, 2016). Further, access to rich instructional settings where students are exposed to academic English

language is a key support for developing language proficiency (Alvarez et al., 2022; Bunch, 2013). In this discussion section, I situate my findings within the broader research and policy space, exploring how my work aligns with other research to contextualize the patterns and impacts observed while also emphasizing how this work extends our knowledge and signals important areas for intervention.

Constricted Access to Core Content

While there were numerous descriptive findings, a key insight from this work is that enrollment levels were much lower for secondary-age arriving immigrant EL-classified students in their year of arrival compared to their peers for all core content subjects other than math. These findings align with prior research. Johnson (2019) finds that recently arrived immigrant EL-classified students were enrolled, on average, in fewer credits in high school than other EL-classified students. Estrada (2014) documents how middle school curricular streams developed for EL-classified students with lower measured ELP consisted of very little core content coursework. Additionally, low levels of enrollment in courses other than math enrollment align with work that finds middle school EL-classified students are more likely to experience exclusionary tracking in ELA than math (Umansky, 2016a).

Exclusion from core content is likely driven by a constellation of factors. However, prior work suggests that exclusion is driven, in part, by the belief or assumption that students cannot meaningfully participate in core content learning until they are proficient, or at least 'more proficient', in English (Estrada, 2014; Hopkins et al., 2013; Valdés, 2001). This aligns with the qualitative findings from this study. Some interviewees described not wanting to place students in core classes because they felt students would not be successful because of low assessed ELP. Additionally, course placement policies focused on ensuring ELD enrollment, with a limited focus on core content other than math. These findings also align with models predicting access by student characteristics, which found that students above the minimum ELP level were more likely to be enrolled in all core content areas other than math. The privileging of ELD over content for EL-classified students has been codified through

policy in other contexts—such as Arizona's now-overturned policy that required EL-classified students be in four hours of ELD a day (Rios-Aguilar et al., 2012) as well as other EL tracking policies (Estrada, 2014). The exception of math enrollment may be explained by the belief that math is the most accessible core subject for EL-classified students (Hansen-Thomas & Cavagnetto, 2010). This was echoed by interviewees who used the common phrasing of math as a universal language. This belief, however, fails to account for the complex language demands of math (Barrow, 2014; Lee & Lee, 2017).

Constricted enrollment upon arrival is not strictly at odds with EL policy, which can be interpreted such that constricted access to core content is as an allowable practice if sufficient compensatory services and supports are provided to address earlier under-enrollment. However, troubling gaps persist beyond the year of arrival for some students. Any exclusion at all, regardless of whether allowable through policy, is a lost opportunity for students to engage in rich learning opportunities, both for their English language proficiency development and their core content learning (Bunch, 2013). Exclusion is also a mechanism for potential segregation. Ensuring access to core content is a critical issue in EL education broadly (Callahan, 2005; Johnson, 2019, Umansky, 2016). However, short timelines to graduation and age-out policies make identifying pathways to graduation that include ELD access as well as rich content learning an important step towards more equitable schooling for immigrant students (Umansky et al., 2018).

Interviewees also linked course exclusion among immigrant EL-classified students to a lack of capacity to meet students' linguistic needs in the core content classroom. This goes hand-in-hand with the belief that students will not be successful in core content classes until their English proficiency is stronger. Together, these insights speak to an overall sense that, at the secondary level, core content classes are being designed and delivered by educators who are comfortable with supporting EL-classified students with a wide range of linguistic proficiencies in the classroom. This characterization

is supported by research, both in teachers' own perceptions of their ability to support EL-classified students in the classroom (Penfield, 1987; Reeves, 2006) and observed professional development content and teachers' abilities (de Jong & Harper, 2007; Harper & de Jong, 2009). Teacher survey results have also revealed that the most desired source of support for working with EL-classified students is content instruction support (Hansen-Thomas & Cavagnetto, 2010).

Understanding Differences in Course Access by Student Characteristics

Another key insight is that there were descriptive differences in course enrollment by student race/ethnicity and socioeconomic status. Raw differences were striking, and often consistent across subject areas and over time. In particular, Black and Hispanic/Latinx students were, with few exceptions, enrolled in core content coursework at lower rates, as were students who were eligible for free/reduced price lunch. These findings, while specific to arriving immigrant EL-classified students, align with other research that finds students from historically marginalized communities are often enrolled in core courses at lower rates than students from other racial/ethnic groups (Asim et al., 2019; Tyson et al., 2007; Yoon & Strobel, 2017). This also aligns with work on how students experiencing economic disadvantage are tracked into less demanding coursework (Ansalone, 2003; Moller & Stearns, 2012).

Differences in enrollment by race/ethnicity or socioeconomic status were not significant once modeled in a regression framework, however this does not mean that we can fully discount the raw differences observed. Descriptively disaggregating outcomes within small subgroups is a critical way to move forward our foundational understanding of differential experiences among underserved, often structurally vulnerable populations (Etz & Arroyo, 2015). Additionally, "controlling away" variation attributable to students' characteristics, schools, or districts can be a form of discounting the racialized systems that contribute to inequities (Castillo & Gillborn, 2022). For example, students from certain racial/ethnic groups may be concentrated in under-resourced districts as a result of

discriminatory, racialized policies or practices (Castillo & Gillborn, 2022). Including fixed effects that account for district-level variation may explain away this variation, while the reason students from certain racial/ethnic groups are in those districts itself may still be very much so tied to racism. As such the raw differences are important to recognize as inequitable and tied to structural racism alongside non-significant estimates on student race/ethnicity as modeled (Castillo & Gillborn, 2022). These descriptive findings point to a need for more action around addressing inequities in access to learning opportunities, especially as experienced by Black, Hispanic/Latinx, and socioeconomically disadvantaged immigrant students. This also highlights a need for deeper work on the intersection of race, class, immigrant status, and opportunity to learn among EL-classified students.

Modeled differences reveal that SIFE students were significantly less likely to be enrolled in math and social studies, as compared with other immigrant EL-classified students. This also aligns with prior work that finds lower prior schooling levels can negatively impact course access and academic performance (Callahan, 2005; Suarez-Orozco et al., 2010). Similar to how teachers often report they feel unprepared to support an array of linguistic backgrounds (Reeves, 2006), there are also reports that teachers feel that they are unable to provide the appropriate levels of support to address significant gaps in students' academic knowledge and skills (Ruiz-de-Velasco & Fix, 2001). This can create tension within districts and schools as teachers and administrators navigate course placement decisions in recognition that students should be in core content coursework, but also with a sense of uncertainty around how well students will be supported in grade-level core content classes (Hopkins et al., 2022; Sugarman, 2017; Umansky et al., 2018).

Newcomer Programs and Course Access

Newcomer programs were the second most common language instruction program used to support secondary-age immigrant EL-classified students in Oregon from 2015/16 to 2018/19. I found no evidence that newcomer programs supported fuller course access, with some estimates of

constricted access. I interpret these findings as cautious estimates, yet there are important takeaways even if not definitive causal findings.

Newcomer programs are likely providing important supplemental services not captured in this study. Socioemotional supports, cultural orientation, targeted academic instruction, and instruction on the English language are all common features of newcomer programs that are important and may support students' well-being in ways that other instructional models do not (Hos, 2016; MAEC, 2019; Short & Boyson, 2012). It may also be that newcomer programs are including primary language supports or instruction, which could be important supports for students. However, it is important these supports be provided alongside core content access, not at the expense of it. Again, shortened timelines underscore a need for thoughtful schedule-building, with critical attention towards what supports are provided and how, if provided, a constellation of supports are functioning together to ensure students have access to core content learning, viable pathways to graduation, instruction on the English language, and culturally-responsive supports (Umansky et al., 2018)

Implications

There are numerous implications for policy and practice that come from this work. In this section, I discuss a need for guidance, both on course placement and instruction. I also advocate for supports for local data use to support action to address constricted course access.

The first implication is that there is room for thoughtful guidance on course placement to be developed by policymakers, practitioners, and other key stakeholders. Such guidance focused on course access can help practitioners as they integrate their perceptions of student readiness with the understanding that students' ELP and prior academic preparedness should not exclude them from the opportunity to engage in core content learning. Those making placement decisions should be provided with information on students' right to access core content, as well as potential flow charts or other aids that outline steps to take to make informed course placement decisions. One potential way to

support more comprehensive course placement in secondary school may be implementing a policy of developing, in partnership with arriving students, individualized graduation plans. These plans can outline proposed trajectories through secondary school while also allowing for individualization depending on students' incoming academic needs, school resources, and student preference. In addition to outlining the trajectory, this would be an opportunity to explicitly identify supports that will be provided, especially in early years. This may be critical to avoid continuous gaps in enrollment observed over time. Some states and agencies have developed templates that can be adapted for local education agencies (Greenberg Motamedi et al., 2021; Rumpf, 2019).

Another area where guidance is critical is supporting teachers in integrating students from diverse linguistic and academic backgrounds into core content classrooms. Once placement decisions are made, it is critical that students and teachers alike do not find themselves in classrooms where they cannot be successful. There is evidence that universal enrollment policies may do harm, rather than good, if they are not delivered in ways that support student success (Domina et al., 2015). Thus, there is a need for more attention towards professional development and supports for core content teachers on how to support a range of student linguistic and academic profiles in the classroom. Professional development and supports can focus on leveraging the assets of arriving students and encouraging student participation (Pacheco & Brown, 2022), as well as developing teachers' understanding of pedagogical language and how to support students' authentic, embedded language development (Bunch, 2013). It may also be a space where co-teaching is prioritized.

Findings on newcomer programs also highlight a need for state and local education agency leaders to provide guidance on how different instructional programs, including newcomer programs, can be delivered so that all the rich supports that come with them are leveraged alongside core content access. This work speaks to an opportunity to re-examine policies around course placement within newcomer

programs to ensure that program timing or structures are not mechanically crowding out core content, as well as identify opportunities to provide core content access in students' home language.

Many important patterns and inequities can be revealed through descriptive data analysis. Districts and schools should be encouraged and supported to analyze their own course enrollment data to reveal inequities in access. Doing so may be an important way to highlight patterns in access among immigrant EL-classified students within schools or districts, as well as encourage action. Descriptive statistics and regression analyses can reveal important patterns and encourage conversations among those who are embedded in the important work of designing course placement policies, implementing course placement policies, and supporting students in the classroom. The descriptive analyses and visualizations in this study are examples of ways in which simple data analysis can reveal important differences in access. This may entail investing in data infrastructure, providing analytic support, and providing the resources to ensure there is time allocated to such tasks. States can also work with local education agencies to consider whether introducing accountability indicators around access to learning opportunities may complement indicators such standardized assessment performance.

There are also implications for research. A deeper mixed-method inquiry could include student voice through qualitative data collection, providing invaluable insights into how students themselves experience the course placement process as well as newcomer programs. Future work can integrate information both on secondary school completion and postsecondary outcomes to better understand the implications of exclusion from coursework, as well as the role of newcomer programs. This study represents just one context, the state of Oregon, which has a smaller immigrant and EL-classified student population than many other states. Future work could look towards other contexts.

Conclusion

I found evidence that secondary age-arriving immigrant EL-classified students in Oregon had limited access to core coursework upon arrival, especially in subjects other than math. Differences in

access was primarily driven by initial ELP and identification as SIFE, although raw differences highlight key inequities by race/ethnicity and socioeconomic status. Course placement practices were largely described as a function of individual discretion, with limited policy frameworks that focused on ELD and math access. Newcomer programs were not related to fuller course access.

These findings have implications outlined above—chief among them being a need for more guidance and resources for developing student schedules and supporting students in the classroom. Schools and districts have the responsibility to ensure that students' socioemotional needs are met, they are provided with English language supports as needed, that they have access to meaningful learning opportunities, and that they are on track to pursue their preferred postsecondary goals. Focusing on core content access and viable pathways to graduation may help educators to better acknowledge and leverage the assets that immigrant students EL-classified students arrive with. This, in turn, may support the creation of educational environments that support students' opportunities to learn and thrive from the start.

APPENDIX

Table A1. Within Districts that Offer Newcomer Programs (N=15), Students Characteristics by Whether in a Newcomer Program

	Not in Newcomer Program (N=1,060)	In a Newcomer Program (N=334)	*p*-value
Student demographics	**(%)**	**(%)**	
Grade of arrival			0.2
6	0.12	0.09	
7	0.14	0.13	
8	0.12	0.12	
9	0.25	0.24	
10	0.16	0.15	
11	0.13	0.14	
12	0.08	0.12	
Female	0.47	0.47	>0.9
Race/Ethnicity			<0.001
Multi-Ethnic & AIAN	0.01	0.00	
Asian/Pacific Islander	0.27	0.19	
Black	0.10	0.12	
Hispanic/Latinx	0.42	0.60	
White	0.20	0.09	
Language			<0.001
Spanish	0.38	0.56	
Arabic	0.08	0.02	
Chinese	0.06	0.04	
Vietnamese	0.05	0.04	
Somali	0.03	0.05	
Other	0.41	0.28	
Special education	0.01	0.00	0.3
SIFE	0.19	0.14	0.049
Initial ELP, lowest level	0.65	0.91	<0.001
FRPL	0.614	0.92	<0.001
School and district characteristics (in year of arrival)	***M***	***M***	
School % EL	0.11	0.15	<0.001
School % Ever-EL	0.18	0.28	<0.001
School % Recent Immig.	0.03	0.02	0.20
School % FRPL	0.58	0.77	<0.001
District % EL	0.15	0.16	<0.001
District % Ever-EL	0.21	0.23	<0.001
District % Recent Immig	0.01	0.01	
District % FRPL.	0.57	0.66	<0.001

Note. EL=English learner. FRPL=Free/reduced price lunch. SIFE= Student with interrupted formal education. ELP=English language proficiency. Recent immigrants are students who've been in the country three years or fewer.

Table A2. Within Schools that Offer Newcomer Programs (N=28), Students Characteristics by Whether in a Newcomer Program

	Not in Newcomer Program (N=313)	In a Newcomer Program (N=334)	*p*-value
Student demographics	**(%)**	**(%)**	
Grade of arrival			<0.001
6	0.06	0.09	
7	0.06	0.13	
8	0.09	0.12	
9	0.38	0.24	
10	0.18	0.15	
11	0.15	0.14	
12	0.09	0.12	
Female	0.41	0.47	0.12
Race/Ethnicity			0.012
Multi-Ethnic & AIAN	0.02	0.00	
Asian/Pacific Islander	0.24	0.19	
Black	0.10	0.12	
Hispanic/Latinx	0.50	0.60	
White	0.14	0.09	
Language			<0.001
Spanish	0.43	0.56	
Arabic	0.08	0.02	
Chinese	0.04	0.04	
Vietnamese	0.03	0.04	
Somali	0.02	0.05	
Other	0.39	0.28	
Special education	0.02	0.00	0.11
SIFE	0.22	0.14	0.011
Initial ELP, lowest	0.73	0.91	<0.001
FRPL	0.72	0.92	<0.001
School and district characteristics (in year of arrival)	***M***	***M***	
School % EL	0.11	0.15	<0.001
School % Ever-EL	0.20	0.28	<0.001
School % Recent Immig.	0.01	0.02	<0.001
School % FRPL	0.67	0.77	<0.001
District % EL	0.19	0.16	0.008
District % Ever-EL	0.26	0.23	0.002
District % Recent Immig	0.01	0.01	<0.001
District % FRPL	0.67	0.66	0.007

Note. EL=English learner. FRPL=Free/reduced price lunch. SIFE= Student with interrupted formal education. ELP=English language proficiency. Recent immigrants are students who've been in the country three years or fewer.

Table A3. Relationship between Student Characteristics and Probability of ELA and Math Course Enrollment, Multilevel Models

	ELA						Math					
	I		II		III		I		II		III	
Predictors	Estimates	95% CI	Estimates	95% CI	Estimates	95% CI	Estimates	95% CI	Estimates	95% CI	Estimates	95% CI
Year in school	0.109***	(0.095; 0.123)	0.093***	(0.081; 0.104)	0.092***	(0.081; 0.104)	0.027***	0.018; 0.036	0.021***	0.013; 0.029	0.021***	0.013; 0.029
SIFE	-0.031	(-0.064; 0.002)	-0.039**	(-0.072; -0.005)	-0.041**	(-0.074; -0.008)	-0.062***	-0.084; -0.039	-0.055***	-0.080; -0.031	-0.055***	-0.080; -0.031
Multi-Ethnic/AIAN	0.032	(-0.071; 0.135)	0	(-0.090; 0.089)	0	(-0.090; 0.089)	0.016	-0.054; 0.085	0.012	-0.055; 0.079	0.013	-0.054; 0.080
Asian/PI	0.019	(-0.016; 0.055)	0.016	-(0.013; 0.046)	0.017	(-0.013; 0.046)	0.014	-0.010; 0.038	0.013	-0.009; 0.035	0.013	-0.009; 0.035
Hispanic/Latinx	-0.005	-(0.052; 0.042)	-0.043**	(-0.082; -0.004)	-0.041*	-(0.080; -0.001)	-0.016	-0.047; 0.016	-0.003	-0.033; 0.026	-0.003	-0.032; 0.027
Black	-0.009	(-0.043; 0.024)	-0.005	(-0.035; 0.025)	-0.006	(-0.036; 0.023)	-0.047***	-0.069; -0.024	-0.02	-0.042; 0.002	-0.022	-0.044; 0.001
Initial ELP	0.131***	(0.104; 0.158)	0.138***	(0.115; 0.161)	0.138***	(0.114; 0.161)	0.026**	0.008; 0.044	0.02*	0.003; 0.038	0.02*	0.003; 0.038
FRPL eligible	-0.062***	(-0.090; -0.034)	-0.03**	(-0.056; -0.004)	-0.032	(-0.058; -0.006)	0.029**	0.010; 0.048	0.021*	0.002; 0.041	0.02*	0.001; 0.040
Identified for special education	-0.04	(-0.153; 0.074)	-0.048	(-0.142; 0.046)	-0.046	(-0.140; 0.048)	-0.019	-0.095; 0.058	0.015	-0.055; 0.085	0.016	-0.054; 0.087
Non-male	0.006	(-0.016; 0.029)	0.009	(-0.009; 0.028)	0.01	(-0.009; 0.028)	-0.01	-0.025; 0.005	-0.015*	-0.029; -0.001	-0.015*	-0.029; -0.001
Random Effects												
σ2			0.1		0.1				0.06		0.06	
School τ00			0.05		0.06				0.03		0.03	
District τ00					0.02						0.01	
School ICC			0.37		0.27				0.36		0.30	
District ICC					0.11						0.07	
Observations	5156		5156		5156		5156		5156		5156	
R2 / R2 adjusted	0.085 / 0.081		0.075 / 0.426		0.076 / 0.419		0.058 / 0.054		0.032 / 0.389		0.034 / 0.380	
AIC	5465.35		3541.626		3571.27		1345.145		556.651		569.847	

Note. Models also include fixed effects for grade of arrival. There were 98 districts and 349 schools represented in the sample. AIAN=American Indian/Alaska Native, PI=Pacific Islander, ELP=English language proficiency, FRPL=Free/reduced price lunch

Table A4. Relationship between Student Characteristics and Probability of Science and Social Studies Course Enrollment, Multilevel Models

	Science						Social Studies					
	I		II		III		I		II		III	
Predictors	Estimates	95% CI	Estimates	95% CI	Estimates	95% CI	Estimates	95% CI	Estimates	95% CI	Estimates	95% CI
Year in school	0.124***	0.110; 0.138	0.114***	0.101; 0.127	0.114***	0.101; 0.127	0.099***	0.084; 0.114	0.084***	0.071; 0.097	0.084***	0.071; 0.098
SIFE	-0.03	-0.064; 0.004	-0.085***	-0.123; -0.047	-0.084***	-0.122; -0.047	-0.015	-0.051; 0.022	-0.095***	-0.134; -0.057	-0.094***	-0.133; -0.056
Multi-Ethnic/AIAN	0.035	-0.070; 0.140	-0.032	-0.134; 0.070	-0.03	-0.132; 0.071	0.024	-0.089; 0.137	0	-0.104; 0.104	0.005	-0.099; 0.109
Asian/PI	0.03	-0.006; 0.066	0.028	-0.006; 0.062	0.028	-0.006; 0.062	0.019	-0.020; 0.057	0.025	-0.009; 0.060	0.025	-0.010; 0.059
Hispanic/Latinx	-0.038	-0.085; 0.010	-0.039	-0.084; 0.007	-0.039	-0.084; 0.007	-0.037	-0.089; 0.014	-0.017	-0.063; 0.029	-0.017	-0.063; 0.030
Black	-0.012	-0.046; 0.022	-0.025	-0.059; 0.009	-0.028	-0.062; 0.006	-0.008	-0.045; 0.029	0.014	-0.021; 0.048	0.009	-0.025; 0.044
Initial ELP	0.139***	0.112; 0.167	0.12***	0.093; 0.146	0.12***	0.093; 0.147	0.14***	0.110; 0.169	0.119***	0.092; 0.146	0.119***	0.092; 0.147
FRPL eligible	-0.04**	-0.069; -0.011	0.001	-0.028; 0.031	0	-0.029; 0.030	-0.092***	-0.123; -0.061	-0.012	-0.043; 0.018	-0.015	-0.045; 0.015
Identified for special education	-0.008	-0.123; 0.108	-0.032	-0.140; 0.076	-0.029	-0.138; 0.079	-0.019	-0.143; 0.106	-0.051	-0.161; 0.059	-0.047	-0.157; 0.063
Non-male	-0.005	-0.028; 0.018	-0.02	-0.041; 0.001	-0.02	-0.041; 0.002	0.011	-0.014; 0.035	0.001	-0.021; 0.023	0.002	-0.020; 0.024
Random Effects												
σ2			0.14		0.14				0.14		0.14	
School τ00			0.03		0.03				0.04		0.06	
District τ00					0.004						0.01	
School ICC			0.19		0.17				0.28		0.22	
District ICC					0.02						0.07	
Observations	5156		5156		5156		5156		5156		5156	
R2 / R2 adjusted	0.128 / 0.125		0.100 / 0.276		0.102 / 0.274		0.104 / 0.101		0.089 / 0.353		0.091 / 0.348	
AIC	5669.936		4938.236		4941.059		6400.872		5137.423		5155.26	

Note. Models also include fixed effects for grade of arrival. There were 98 districts and 349 schools represented in the sample.

Table A5. Summary Statistics for Research Question 2 Analytic Sample, Matched

	Year 1			Year 2			Year 3		
	Not in Newcomer Program	Newcomer Program		Not in Newcomer Program	Newcomer Program		Not in Newcomer Program	Newcomer Program	
	N = 682	N = 274	*p*-value	N = 369	N = 158	*p*-value	N = 270	N = 106	*p*-value
Arrival Grade			>0.9			>0.9			>0.9
6	0.08	0.08		0.07	0.07		0.11	0.11	
7	0.13	0.13		0.12	0.12		0.17	0.17	
8	0.11	0.11		0.09	0.09		0.12	0.12	
9	0.27	0.27		0.32	0.32		0.34	0.34	
10	0.15	0.15		0.15	0.15		0.17	0.17	
11	0.13	0.13		0.13	0.13		0.06	0.06	
12	0.14	0.14		0.11	0.11		0.03	0.03	
Female	0.42	0.48		0.43	0.46	0.9	0.41	0.52	0.9
Race/Ethnicity			>0.9			>0.9			>0.9
Multi-Ethnic & AIAN	0.00	0.00		0.00	0.00		0.00	0.00	
Asian/Pacific Islander	0.18	0.18		0.20	0.20		0.17	0.17	
Black	0.09	0.09		0.08	0.08		0.13	0.13	
Hispanic/Latinx	0.64	0.64		0.60	0.60		0.58	0.58	
White	0.08	0.08		0.11	0.11		0.11	0.11	
Language			>0.9			>0.9			0.8
Spanish	0.66	0.66		0.65	0.63		0.51	0.53	
Arabic	0.02	0.02		0.02	0.02		0.06	0.03	
Chinese	0.03	0.03		0.02	0.02		0.03	0.02	
Vietnamese	0.03	0.03		0.04	0.03		0.06	0.06	
Somali	0.02	0.03		0.02	0.03		0.08	0.05	
Other	0.24	0.24		0.26	0.27		0.27	0.32	
Special Education	0.01	0.00	0.2	0.01	0.01	0.5	0.01	0.01	0.8
SIFE	0.09	0.09	>0.9	0.07	0.07	>0.9	0.08	0.08	>0.9
Initial ELP > 0	0.08	0.07		0.08	0.08		0.08	0.08	
FRPL Eligible	0.92	0.92	>0.9	0.95	0.94	0.70	0.92	0.92	0.6
School & district chars (yr. arrival)									
School % EL	0.08	0.14	<0.001	0.08	0.12	<0.001	0.11	0.11	0.30
School % Ever-EL	0.16	0.26	<0.001	0.15	0.25	<0.001	0.24	0.26	0.20
School % Immig.	0.01	0.02	<0.001	0.01	0.01	0.40	0.01	0.01	<0.001
School % FRPL	0.68	0.76	<0.001	0.67	0.74	0.20	0.66	0.72	0.40
District % EL	0.12	0.16	<0.001	0.11	0.15	<0.001	0.16	0.14	0.08
District % Ever-EL	0.18	0.23	<0.001	0.17	0.23	<0.001	0.26	0.23	0.02
District % Immig.	0.01	0.01	0.13	0.01	0.01	0.60	0.01	0.01	<0.001
District % FRPL	0.65	0.66	0.20	0.65	0.64	0.70	0.64	0.62	0.20

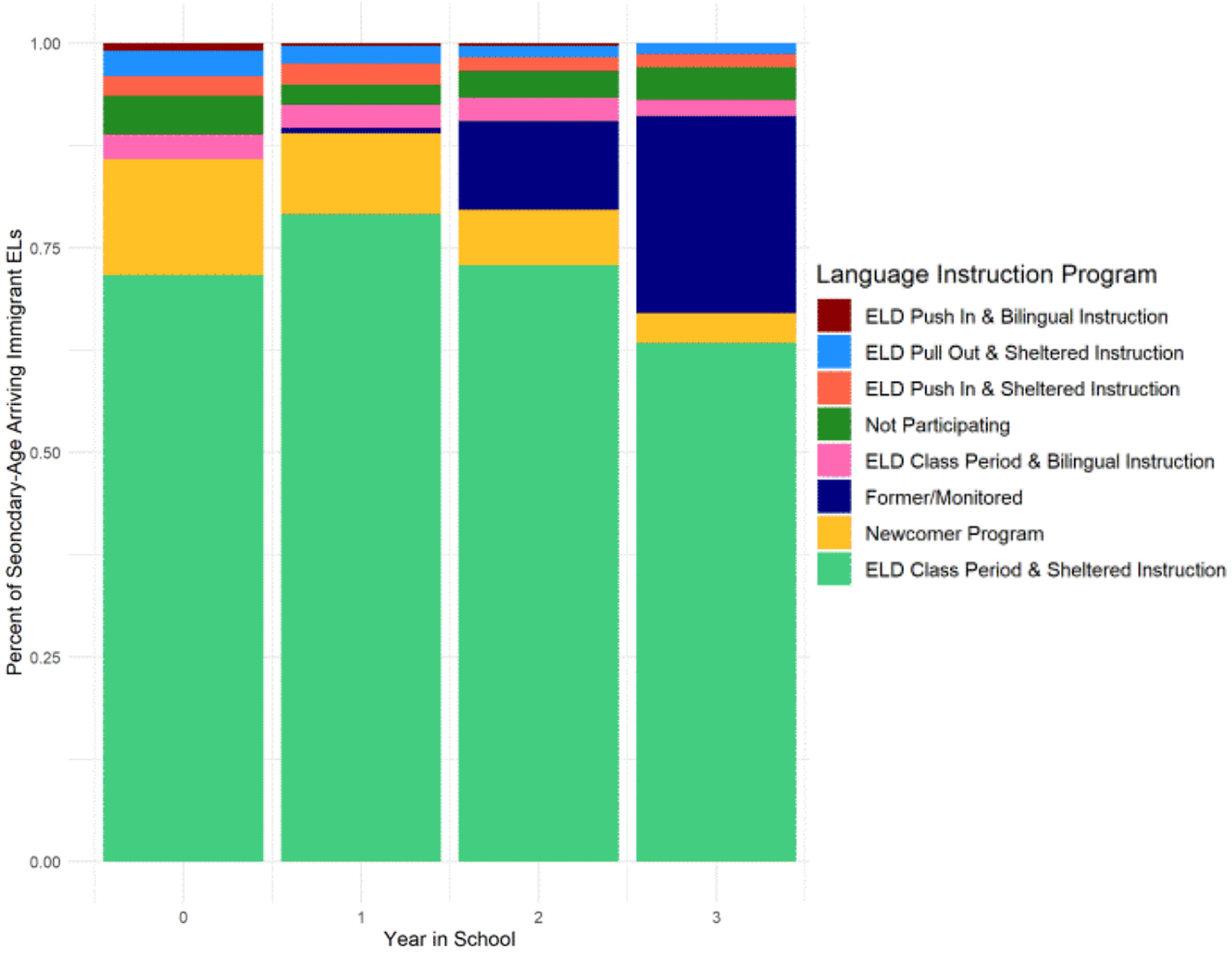
Note. EL=English learner. FRPL=Free/reduced price lunch. Immig.=Immigrant students who've been in the country three years or fewer.

Table A6. Estimated Relationship between Newcomer Program Participation and Course Enrollment, Unmatched Sample

		Model I: Cov. Adj				Model II: Cov. Adj. w/Chrt FE				Model III: Cov. Adj. w/Chrt & Dist. FE			
		Year 1	Year 2	Year 3	Year 4	Year 1	Year 2	Year 3	Year 4	Year 1	Year 2	Year 3	Year 4
ELA													
	Newcomer Program [95 % CI]	-0.284* [-0.515; -0.053]	-0.148. [-0.306; 0.011]	-0.124* [-0.225; -0.023]	-0.019 [-0.108; 0.070]	-0.274* [-0.502; -0.045]	-0.153. [-0.312; 0.006]	-0.124* [-0.225; -0.023]	-0.019 [-0.108; 0.070]	-0.227. [-0.482; 0.027]	-0.082 [-0.195; 0.030]	-0.071 [-0.163; 0.021]	-0.031 [-0.103; 0.041]
Math													
	Newcomer Program [95 % CI]	0.045 [-0.063; 0.154]	0.027 [-0.019; 0.073]	0.015 [-0.023; 0.053]	0.002 [-0.084; 0.088]	0.046 [-0.063; 0.155]	0.025 [-0.020; 0.070]	0.015 [-0.024; 0.053]	0.002 [-0.084; 0.088]	0.061. [-0.009; 0.131]	-0.003 [-0.038; 0.032]	-0.032* [-0.060; -0.004]	0.008 [-0.101; 0.117]
Science													
	Newcomer Program [95 % CI]	-0.123 [-0.337; 0.090]	-0.105 [-0.263; 0.052]	-0.042 [-0.114; 0.029]	0.115* [0.021; 0.208]	-0.120 [-0.333; 0.094]	-0.107 [-0.267; 0.052]	-0.042 [-0.114; 0.030]	0.115* [0.021; 0.208]	-0.110 [-0.361; 0.142]	-0.094 [-0.277; 0.090]	-0.075. [-0.154; 0.004]	-0.014 [-0.138; 0.110]
Social Studies													
	Newcomer Program [95 % CI]	-0.180 [-0.405; 0.045]	-0.111. [-0.227; 0.004]	-0.065 [-0.154; 0.024]	0.108. [-0.012; 0.227]	-0.178 [-0.400; 0.044]	-0.109. [-0.221; 0.003]	-0.065 [-0.155; 0.025]	0.108. [-0.012; 0.227]	-0.173 [-0.487; 0.141]	-0.083 [-0.188; 0.022]	-0.171** [-0.279; -0.063]	0.150. [-0.019; 0.318]
	Cov.	X	X	X	X	X	X	X	X	X	X	X	X
	Chrt. FE					X	X	X	X	X	X	X	X
	Dist. FE									X	X	X	X
	N	2,370	1,446	943	303	2,370	1,446	943	303	2,370	1,446	943	303

Note. .*p*<.01, **p*<.05, ***p*<.01, ****p*<.001. N represents the total student observations in the model. Estimates are accompanied by 95% confidence intervals. Results are presented for three model specifications, each estimated for each separate subject and year in the data. Cov.= mode that includes student, school, and district-level covariates. Chrt FE=Cohort fixed effects, Dist. FE=District fixed effects.

Figure A1. Percent of Students in a Given Language Instruction Program Combination Over Time, 2015/16-2018/19



Note. Figure represents the combined language instruction programs used to provide English language development (ELD) and support core content access for immigrant EL-classified students who arrived in grades 6-12 from 2015/16-2018/19. While the initial year (0) reports data for 2,370 students, this number shrinks over time due to the combination of cohorts into one sample and students exiting the data set.

Figure A2. Science Course Access for Arriving Immigrant EL-Classified Students, EL-Classified Students, and Students who are neither EL-Classified, nor Arriving Immigrant Students

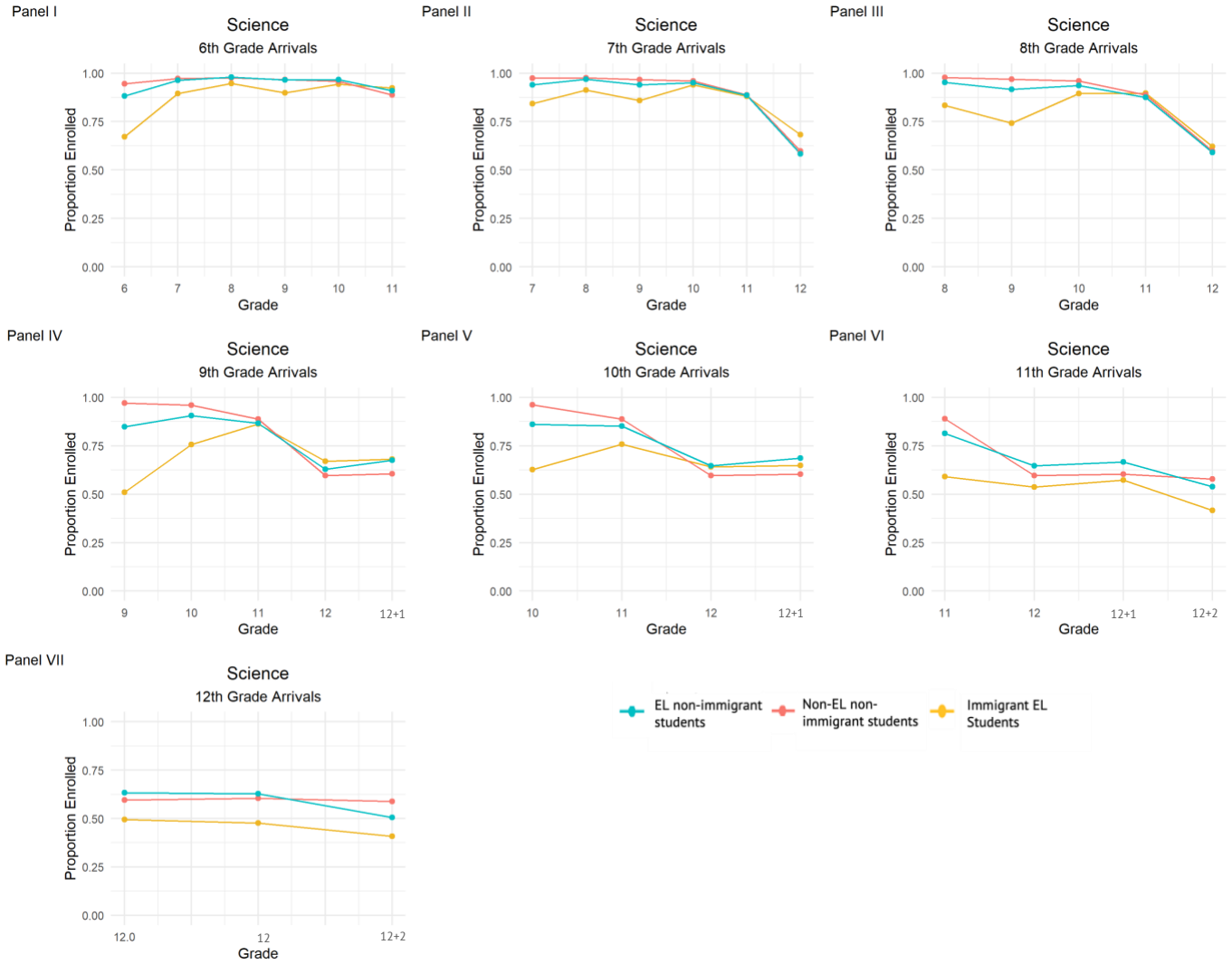


Figure A3. Social Studies Course Access for Arriving Immigrant EL-Classified Students, EL-Classified Students, and Students who are neither EL-Classified, nor Arriving Immigrant Students

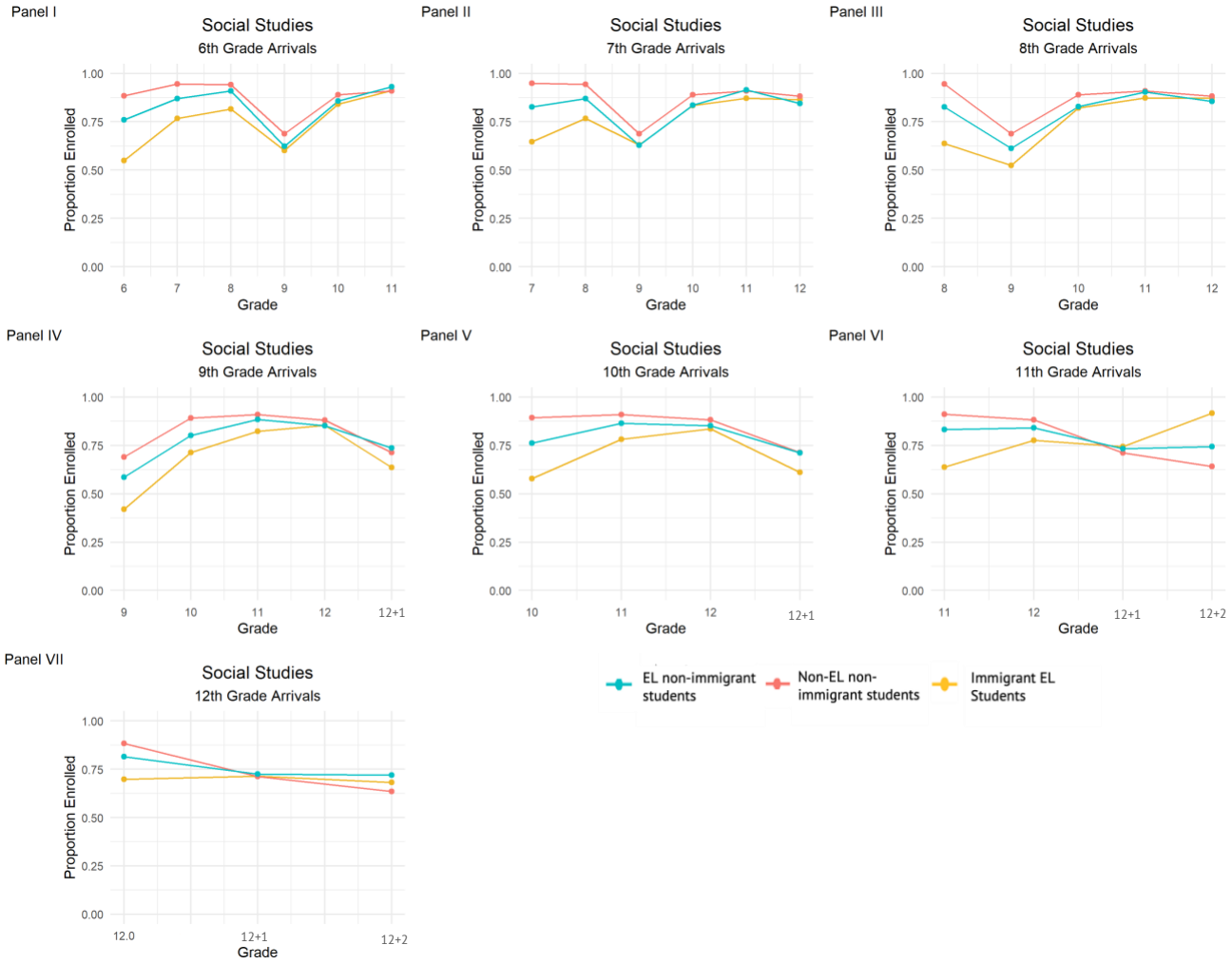
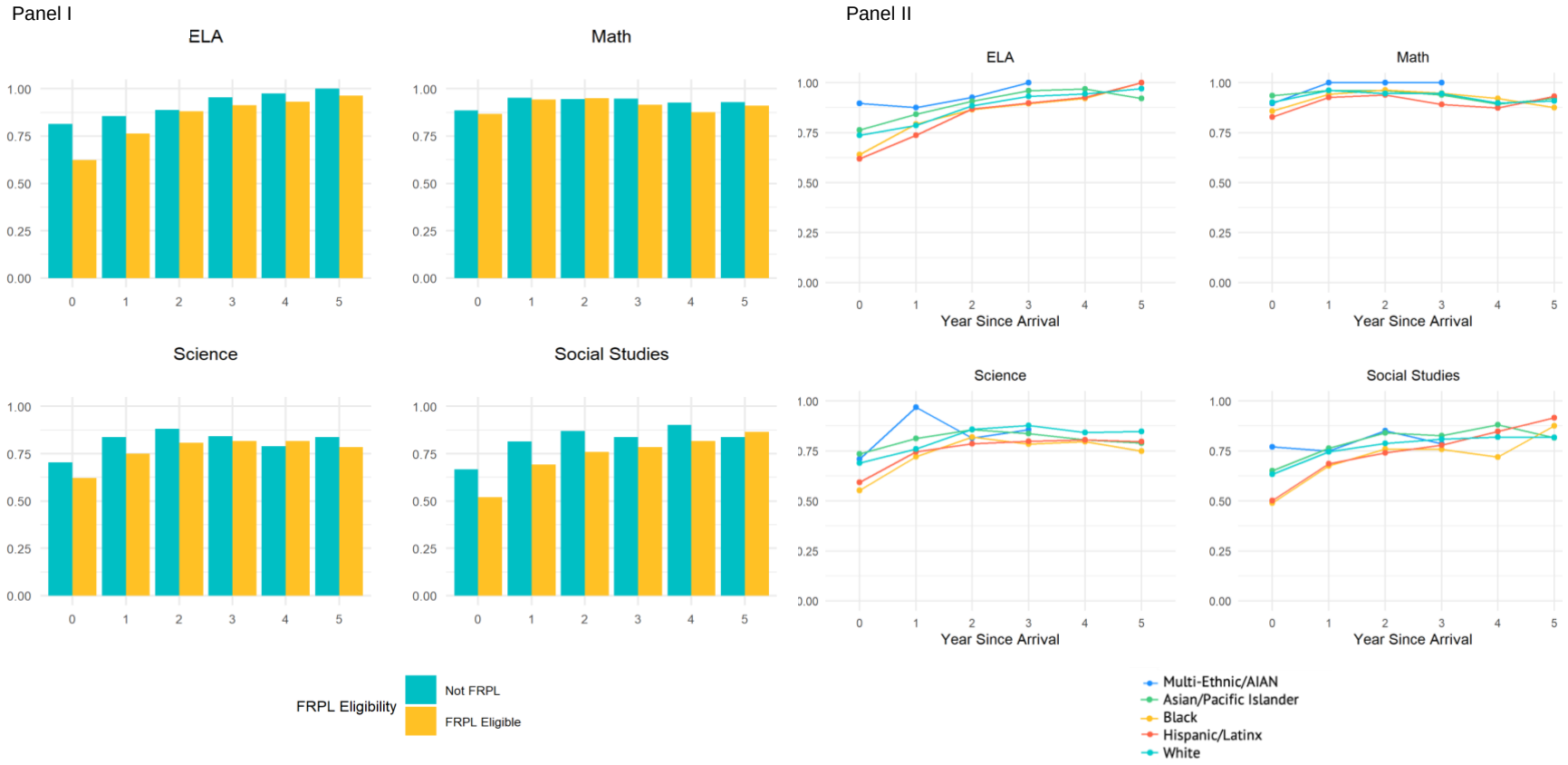
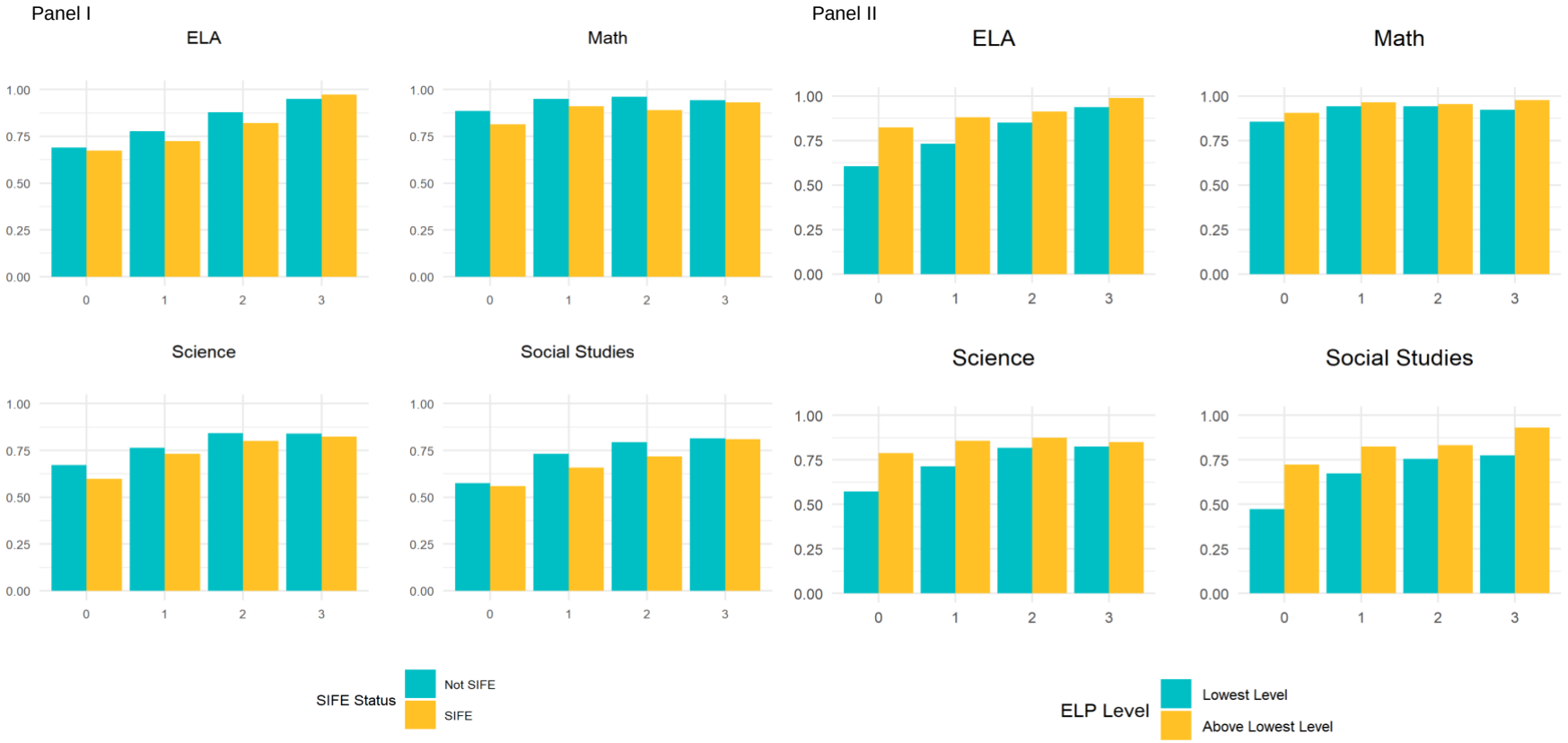


Figure A4. Percent of Secondary-Age Arriving Immigrant EL-Classified Students Enrolled in a Given Subject Area by Whether or Not they are Eligible for Free/Reduced Price Lunch (Panel I) and Race/Ethnicity (Panel II)



Note. AIAN=American Indian/Alaska Native, FRPL=Free/reduced price lunch. Figures in Panel 1 and 2 represents the proportion of students enrolled in a given core content area over time, for 4,645 students who arrived in grades 6-12 from 2013/14-2018/19.

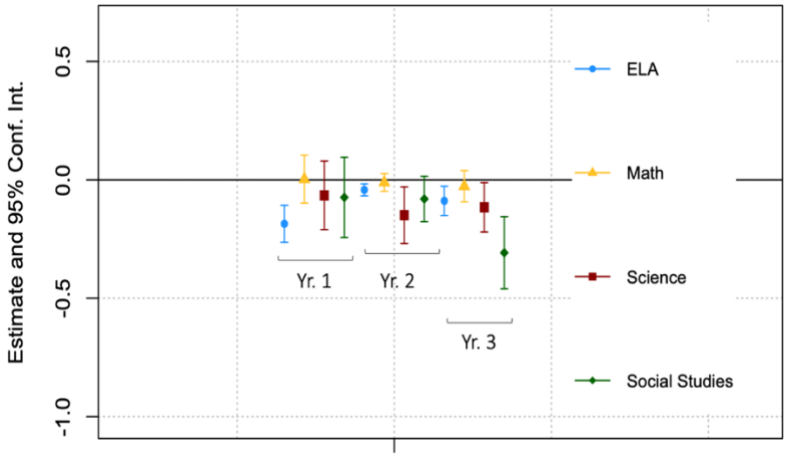
Figure A5. Percent of Secondary-Age Arriving Immigrant EL-Classified Students Enrolled in a Given Subject Area by Whether or Not they are Identified as Having Interrupted Formal Education (Panel I), or Initial ELP Level (Panel II)



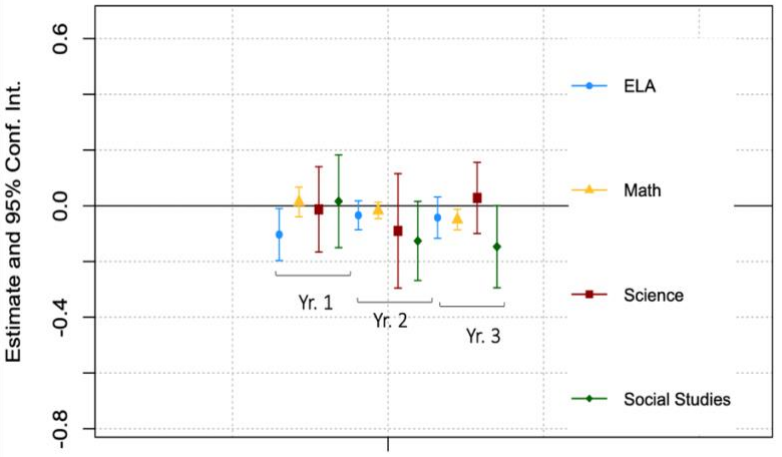
Note. ELP=English language proficiency, SIFE=Student with interrupted formal education. Panel 1 represents the proportion of students enrolled in a given core content area over time, by whether they were identified as having interrupted formal education, for 3,185 students who arrived in grades 6-12 from 2015/16-2018/19. Panel 2 represents the proportion of students enrolled in a given core content area over time, by whether their initial English proficiency was assessed at the lowest level, or above the lowest level, for 2,398 students who arrived in grades 6-12 from 2015/16-2018/19

Figure A6. Relationship between Newcomer Program Participation and Course Access, Alternative Matching Specifications and Unmatched Sample, District Fixed Effects Specification

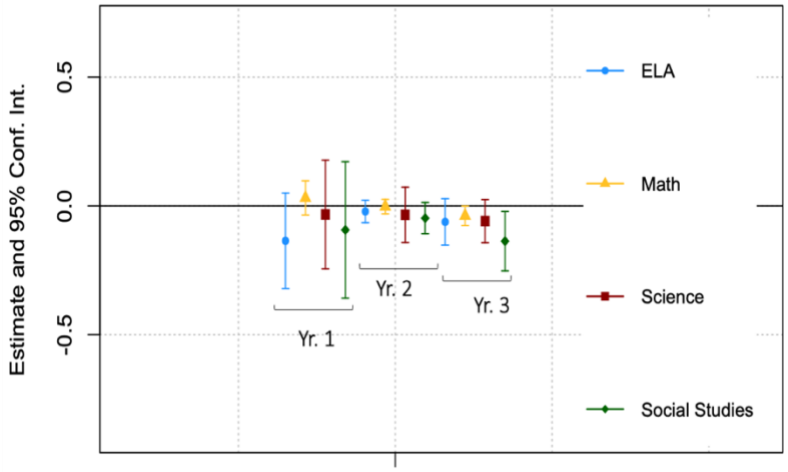
Panel I. Exact District Match



Panel II. Match on School Covariates



Panel III. Match on Student Covariates



Panel IV. Unmatched Estimates

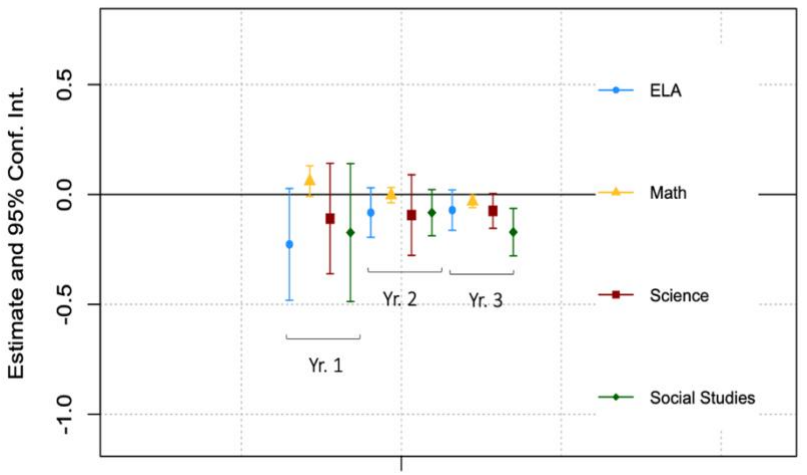
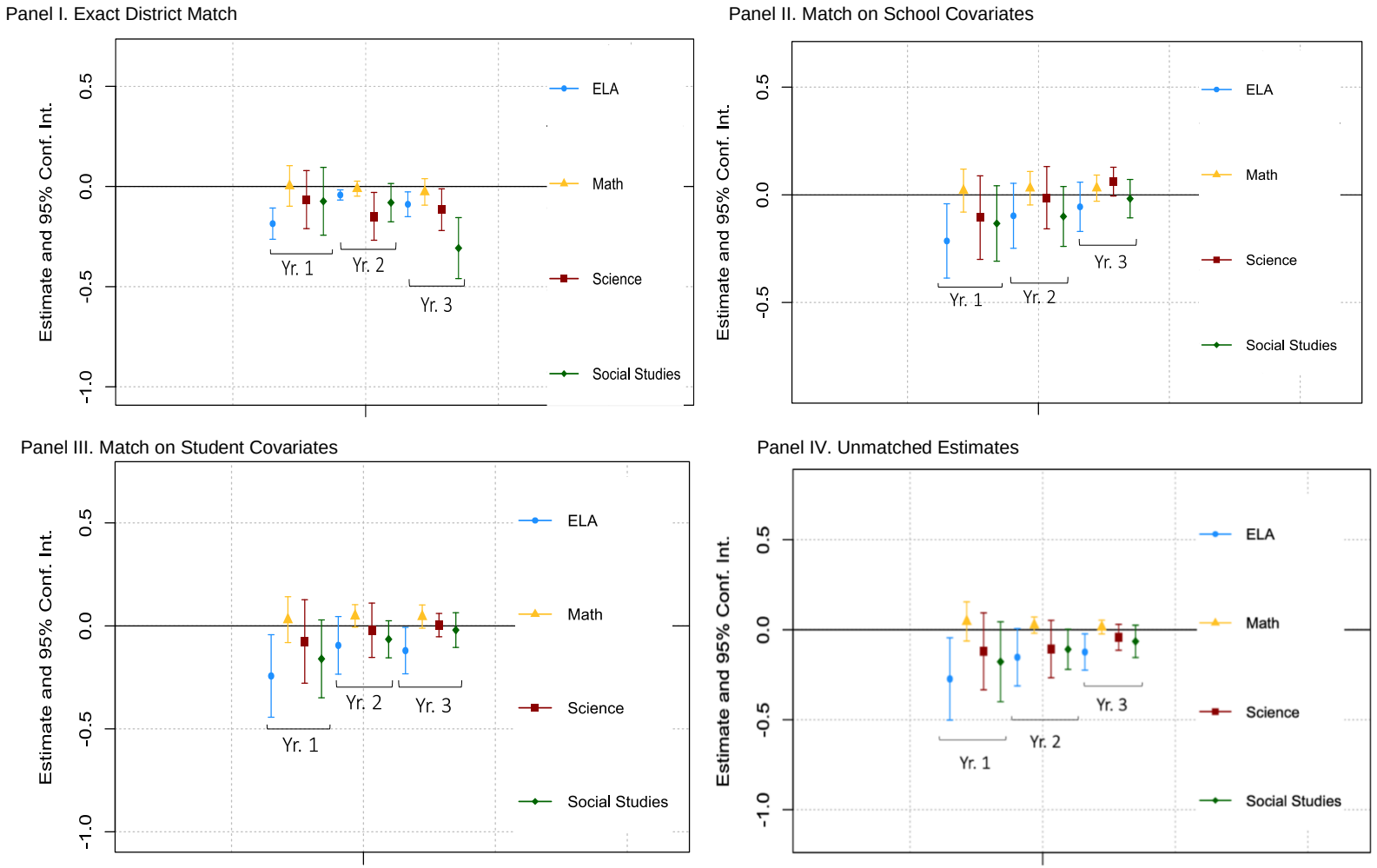


Figure A7. Relationship between Newcomer Program Participation and Course Access, Alternative Specifications, Cohort Fixed Effects Specification



REFERENCES CITED

Abedi, J., & Herman, J. (2010). Assessing English language learners' opportunity to learn mathematics: Issues and limitations. *Teachers College Record, 112*(3), 723-746.

Ansalone, G. (2003). Poverty, tracking, and the social construction of failure: International perspectives on tracking. *Journal of Children and Poverty, 9*(1), 3-20.

Alvarez, L., De Loney, M., Capitelli, S., Valdés, G., & Julia Biernacki, P. (2022). Expanding participation: supporting newcomer students' language development through disciplinary practices. *Language and Education*, *32*(1) 1-21.

Asim, M., Kurlaender, M., & Reed, S. (2019). *12th Grade Course-Taking and the Distribution of Opportunity for College Readiness in Mathematics*. Policy Analysis for California Education.

Attewell, P., & Domina, T. (2008). Raising the bar: Curricular intensity and academic performance. *Educational Evaluation and Policy Analysis, 30*(1), 51-71.

Barrow, M. A. (2014). Even math requires learning academic language. *Phi Delta Kappan*, *95*(6), 35-38.

Bray, L. E., & Russell, J. L. (2016). Going off script: Structure and agency in individualized education program meetings. *American Journal of Education*, *122*(3), 367-398.

Browder, C. T. (2014). English learners with limited or interrupted formal education: Risk and resilience in educational outcomes. *University of Maryland*.

Brown, C. S., & Chu, H. (2012). Discrimination, ethnic identity, and academic outcomes of Mexican immigrant children: The importance of school context. *Child Development, 83*(5), 1477-1485.

Bunch, G. C. (2013). Pedagogical language knowledge: Preparing mainstream teachers for English learners in the new standards era. *Review of Research in Education*, *37*(1), 298-341.

Callahan, R., & Shifrer, D. (2016). Equitable access for secondary English learner students: Course taking as evidence of EL program effectiveness. *Educational Administration Quarterly, 52*(3), 463-496.

Callahan, R., Wilkinson, L., & Muller, C. (2010). Academic achievement and course taking among language minority youth in U.S. schools: Effects of ESL placement. *Educational Evaluation and Policy Analysis, 32*(1), 84-117.

Callahan, R. M. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal, 42*(2), 305-328.

Callahan, R. M., & Shifrer, D. (2016). Equitable access for secondary English learner students: Course taking as evidence of EL program effectiveness. *Educational Administration Quarterly, 52*(3), 463-496.

Carhill, A., Suárez-Orozco, C., & Páez, M. (2008). Explaining English language proficiency among adolescent immigrant students. *American Educational Research Journal, 45*(4), 1155-1179.

Castillo, Wendy, and David Gillborn. (2022). How to "QuantCrit:" Practices and Questions for Education Data Researchers and Users. (EdWorkingPaper: 22-546).

Chu, S.-Y. (2009). Implementation of supportive school programs for immigrant students in the United States. *Preventing School Failure: Alternative Education for Children and Youth, 53*(2), 67-72.

Custodio, B., & O'Loughlin, J. B. (2017). *Students with interrupted formal education: Bridging where they are and what they need*. Corwin Press.

Dabach, D. B. (2014). "I am not a shelter!": Stigma and social boundaries in teachers' accounts of students' experience in separate "sheltered" English learner classrooms. *Journal of Education for Students Placed at Risk 19*(2), 98-124.

Dear Colleague Letter: English Learner Students and Limited English Proficient Parents (2015). U.S. Department of Education.

Domina, T., McEachin, A., Penner, A., & Penner, E. (2015). Aiming high and falling short: California's eighth-grade algebra-for-all effort. *Educational Evaluation and Policy Analysis, 37*(3), 275-295.

Dryden-Peterson, S. (2017). Refugee education: Education for an unknowable future. *Curriculum Inquiry, 47*(1), 14-24.

Ee, J., & Gándara, P. (2020). The impact of immigration enforcement on the nation's schools. *American Educational Research Journal, 57*(2), 840-871.

Estrada, P. (2014). English learner curricular streams in four middle schools: Triage in the trenches. *The Urban Review, 46*(4), 535-573.

Etz, K. E., & Arroyo, J. A. (2015). Small sample research: Considerations beyond statistical power. *Prevention Science*, *16*(7), 1033-1036.

Faltis, C., & Coulter, C. (2008). *Teaching English learners and immigrant students in secondary schools*. Pearson/Merrill Prentice Hall.

Feinberg, R. C. (2000). Newcomer schools: Salvation or segregated oblivion for immigrant students? *Theory into Practice, 39*(4), 220-227.

Fine, M., Stoudt, B., & Futch, V. (2005). *The Internationals Network for public schools: A quantitative and qualitative cohort analysis of graduation and dropout rates*. The Graduate Center, City University of New York, 1-87.

Fry, R. (2005). *The higher dropout rate of foreign-born teens: The role of schooling abroad*. Pew Research Center.

Gamoran, A., & Hannigan, E. C. (2000). Algebra for everyone? Benefits of college-preparatory mathematics for students with diverse abilities in early secondary school. *Educational Evaluation and Policy Analysis, 22*(3), 241-254.

Gándara, P., & Orfield, G. (2012). Segregating Arizona's English Learners: A Return to the" Mexican Room"? *Teachers College Record, 114*(9), 1-27.

Gándara, P., Rumberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English Learners in California Schools: Unequal resources,'Unequal outcomes. *education policy analysis archives, 11*(36), 1-24.

Garver, R., & Hopkins, M. (2020). Segregation and integration in the education of English learners: Leadership and policy dilemmas. *Leadership and Policy in Schools, 19*(1), 1-5.

Gibbons, C. E., Serrato, J. C. S., & Urbancic, M. B. (2019). Broken or fixed effects?. *Journal of Econometric Methods*, *8*(1), 1-33.

Giersch, J. (2018). Academic tracking, high-stakes tests, and preparing students for college: How inequality persists within schools. *Educational Policy, 32*(7), 907-935.

Gonzales, R. G., Suárez-Orozco, C., & Dedios-Sanguineti, M. C. (2013). No place to belong: Contextualizing concepts of mental health among undocumented immigrant youth in the United States. *American Behavioral Scientist, 57*(8), 1174-1199.

Greenberg Motamedi, J., Porter, L., Taylor, S., Leong, M., Martinez-Wenzl, M., & Serrano, D. (2021). *Welcoming, Registering, and Supporting Newcomer Students: A Toolkit for Educators of Immigrant and Refugee Students in Secondary Schools*. REL 2021-064. Regional Educational Laboratory Northwest.

Hakuta, K. (2020). A policy history of leadership dilemmas in English learner education. *Leadership and Policy in Schools, 19*(1), 6-9.

Hansen-Thomas, H., & Cavagnetto, A. (2010). What do mainstream middle school teachers think about their English language learners? A tri-state case study. *Bilingual Research Journal*, *33*(2), 249-266.

Harklau, L. (1994). Tracking and linguistic minority students: Consequences of ability grouping for second language learners. *Linguistics and Education, 6*(3), 217-244.

Hersi, A. A., & Watkinson, J. S. (2012). Supporting immigrant students in a newcomer high school: A case study. *Bilingual Research Journal, 35*(1), 98-111.

Ho, D., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, *42*(8), 1–28.

Hopkin, C. R., Hoyle, R. H., & Gottfredson, N. C. (2015). Maximizing the yield of small samples in prevention research: A review of general strategies and best practices. *Prevention Science*, *16*(7), 950-955.

Hopkins, M., Lowenhaupt, R., & Sweet, T. M. (2015). Organizing English learner instruction in new immigrant destinations: District infrastructure and subject-specific school practice. *American Educational Research Journal, 52*(3), 408-439.

Hopkins, M., Martinez-Wenzl, M., Aldana, U. S., & Gándara, P. (2013). Cultivating capital: Latino newcomer young men in a U.S. urban high school. *Anthropology & Education Quarterly, 44*(3), 286-303.

Hopkins, M., Weddle, H., Bjorklund Jr, P., Umansky, I. M., & Blanca Dabach, D. (2021). "It's created by a community": Local context mediating districts' approaches to serving immigrant and refugee newcomers. *AERA Open, 7*. Online first.

Hopkins, M., Weddle, H., Lavadenz, M., Murillo, M. A., & Vahdani, T. (2022). Examining the English learner policy ecology: how educators navigated the provision of designated English Language Development (ELD) support at the secondary level. *Peabody Journal of Education*, *97*(1), 47-61.

Hos, R. (2016). Caring is not enough: Teachers' enactment of ethical care for adolescent students with limited or interrupted formal education (SLIFE) in a newcomer classroom. *Education and Urban Society, 48*(5), 479-503.

Houston Independent School District (2017). *Newcomer Immigrant Program Evaluation Report, 2016-2017*. Houston Independent School District.

Houston Independent School District (2018). *Newcomer Immigrant Program Evaluation Report, 2017-2018*. Houston Independent School District.

Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, 1-24.

Jaffe-Walter, R., & Miranda, C. P. (2020). Segregation or sanctuary? Examining the educational possibilities of counterpublics for immigrant English learners. *Leadership and Policy in Schools, 19*(1), 104-122.

Joensen, J. S., & Nielsen, H. S. (2009). Is there a causal effect of high school math on labor market outcomes? *Journal of Human Resources, 44*(1), 171-198.

Johnson, A. (2019). A matter of time: Variations in high school course-taking by years-as-EL subgroup. *Educational Evaluation and Policy Analysis, 41*(4), 461-482.

Kanno, Y., & Kangas, S. E. (2014). "I'm not going to be, like, for the AP" English language learners' limited access to advanced college-preparatory courses in high school. *American Educational Research Journal, 51*(5), 848-878.

Kelly, S. (2009). The black-white gap in mathematics course taking. *Sociology of Education*, *82*(1), 47-69.

Lee, W., & Lee, J. S. (2017). Math instruction is not universal: Language-specific pedagogical knowledge in Korean/English two-way immersion programs. *Bilingual Research Journal*, *40*(4), 353-371.

Lee, S. J., Park, E., & Wong, J. H. S. (2017). Racialization, schooling, and becoming American: Asian American experiences. *Educational Studies*, *53*(5), 492-510.

Lee, M. G., & Soland, J. G. (2022). Does reclassification change How English learners feel about school and themselves? Evidence from a regression discontinuity design. *Educational Evaluation and Policy Analysis*, Online first.

Lewis, L., & Gray, L. (2016). *Programs and Services for High School English Learners in Public School Districts: 2015-16. First Look*. National Center for Education Statistics.

Lillie, K. E., Markos, A., Estrella, A., Nguyen, T., Trifiro, A., Arias, M. B., ... & Pérez, K. (2010). *Policy in Practice: The Implementation of Structured English Immersion in Arizona.* Civil Rights Project/Proyecto Derechos Civiles.

Long, M. C., Conger, D., & Iatarola, P. (2012). Effects of high school course-taking on secondary and postsecondary success. *American Educational Research Journal, 49*(2), 285-322.

Lowenhaupt, R. (2016). Immigrant acculturation in suburban schools serving the new Latino diaspora. *Peabody Journal of Education, 91*(3), 348-365.

Lowenhaupt, R., Bradley, S., & Dallas, J. (2020). The (re) classification of English learners: A district case study of identification, integration, and the design of services. *Leadership and Policy in Schools, 19*(1), 60-80.

Lowenhaupt, R., & Reeves, T. (2015). Toward a theory of school capacity in new immigrant destinations: Instructional and organizational considerations. *Leadership and Policy in Schools, 14*(3), 308-340.

Lowenhaupt, R., & Scanlan, M. (2020). Lessons from the primary and secondary school context: Building capacity to support migrant and refugee students. In *Refugees and Higher Education* (pp. 275-294). Brill Sense.

Massey, D. S. (Ed.). (2008). *New faces in new places: The changing geography of American immigration.* Russell Sage Foundation.

Mavrogordato, M., & White, R. S. (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.

McBrien, J. L. (2005). Educational needs and barriers for refugee students in the United States: A review of the literature. *Review of Educational Research, 75*(3), 329-364.

McDonnell, L. M. (1995). Opportunity to learn as a research concept and a policy instrument. *Educational Evaluation and Policy Analysis, 17*(3), 305-322.

Menken, K. (2013). Emergent bilingual students in secondary school: Along the academic language and literacy continuum. *Language Teaching, 46*(4), 438-476.

Mid-Atlantic Equity Consortium (2019). *Creating New Futures for Newcomers: Lessons from Five Schools that Serve K-12 Immigrants, Refugees, and Asylees*. Mid-Atlantic Equity Consortium.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2018). *Qualitative data analysis: A methods sourcebook*. Sage publications.

Moller, S., & Stearns, E. (2012). Tracking success: High school curricula and labor market outcomes by race and gender. *Urban Education, 47*(6), 1025-1054.

Morita-Mullaney, T., Renn, J., & Chiu, M. M. (2021). Contesting math as the universal language: a longitudinal study of dual language bilingual education language allocation. *International Multilingual Research Journal*, *15*(1), 43-60.

Mosqueda, E. (2010). Compounding inequalities: English proficiency and tracking and their relation to mathematics performance among Latina/o secondary school youth. *Journal of Urban Mathematics Education*, *3*(1), 57-81.

Oakes, J. (1986). Tracking, inequality, and the rhetoric of reform: Why schools don't change. *Journal of Education*, *168*(1), 60-80.

Oakes, J., & Guiton, G. (1995). Matchmaking: The dynamics of high school tracking decisions. *American Educational Research Journal*, *32*(1), 3-33.

Pacheco, M. B., & Brown, J. C. (2022). Newcomer Emergent Bilingual Students' Meaning-Making in Urban Biology Classrooms: A Communities of Practice Perspective. *Urban Education 21*(3). 1-37.

Penn, D. M. (2021). Beyond receptivity: Exploring the role of identity in educators' orientation toward newcomers in a new immigrant destination. *AERA Open, 7*, Online first.

Portes, A., & Rumbaut, R. G. (2014). *Immigrant America*. University of California Press.

Potochnick, S. (2018). The academic adaptation of immigrant students with interrupted schooling. *American Educational Research Journal, 55*(4), 859-892.

Quinn, D. M. (2020). Experimental evidence on teachers' racial bias in student evaluation: The role of grading scales. *Educational Evaluation and Policy Analysis, 42*(3), 375-392.

Rana, M., Baolian Qin, D., Bates, L., Luster, T., & Saltarelli, A. (2011). Factors related to educational resilience among Sudanese unaccompanied minors. *Teachers College Record, 113*(9), 2080-2114.

Reeves, J. R. (2006). Secondary teacher attitudes toward including English Language Learners in mainstream classrooms. *Journal of Educational Research 99*(3) 131–143

Rigby, J. G., Woulfin, S. L., & März, V. (2016). Understanding how structure and agency influence education policy implementation and organizational change. *American Journal of Education*, *122*(3), 295-302.

Rios-Aguilar, C., Canche, M. S. G., & Moll, L. C. (2012). Implementing structured English immersion in Arizona: Benefits, challenges, and opportunities. *Teachers College Record*, *114*(9), 1-18.

Rodriguez, S. (2020). "I was born at the border, like the 'Wrong' side of it": Undocumented Latinx Youth experiences of Racialization in the US South. *Anthropology & Education Quarterly*, *51*(4), 496-526.

Rosa, J., & Flores, N. (2017). Unsettling race and language: Toward a raciolinguistic perspective. *Language in Society, 46*(5), 621-647.

Ruiz-de-Velasco, J., & Fix, M. (2000). *Overlooked & Underserved: Immigrant Students in US Secondary Schools*. The Urban Institute.

Rumpf, R. (2019). The importance of enrollment and graduation planning for adolescent newcomer English learners. In *Educational Leadership of Immigrants* (pp. 187-194). Routledge.

Sáenz, R., & Manges Douglas, K. (2015). A call for the racialization of immigration studies: On the transition of ethnic immigrants to racialized immigrants. *Sociology of Race and Ethnicity*, *1*(1), 166-180.

Short, D. J. (2002). Newcomer programs: An educational alternative for secondary immigrant students. *Education and Urban Society, 34*(2), 173-198.

Short, D. J., & Boyson, B. A. (2012). *Helping newcomer students succeed in secondary schools and beyond* (Vol. 78). Center for Applied Linguistics.

Short, D. J., & Fitzsimmons, S. (2007). *Double Work*: *Challenges and Solutions to Acquiring Language and Academic Literacy for Adolescent English Language Learners. A Report to Carnegie Corporation of New York,* Carnegie Corporation.

Starck, J. G., Riddle, T., Sinclair, S., & Warikoo, N. (2020). Teachers are people too: Examining the racial bias of teachers compared to other American adults. *Educational Researcher, 49*(4), 273-284.

Suárez-Orozco, C., Gaytán, F. X., Bang, H. J., Pakes, J., O'Connor, E., & Rhodes, J. (2010). Academic trajectories of newcomer immigrant youth. *Developmental Psychology, 46*(3), 602-623.

Sugarman, J. (2017). *Beyond Teaching English: Supporting High School Completion by Immigrant and Refugee Students*. Migration Policy Institute.

Sugarman, J. (2019*). The unintended consequences for English learners of using the four-year graduation rate for school accountability*. Migration Policy Institute.

Thompson, K. D., Umansky, I. M., & Porter, L. (2020). Examining contexts of reception for newcomer students. *Leadership and Policy in Schools, 19*(1), 10-35.

Tyson, W., Lee, R., Borman, K. M., & Hanson, M. A. (2007). Science, technology, engineering, and mathematics (STEM) pathways: High school science and math coursework and postsecondary degree attainment. *Journal of Education for Students Placed at Risk, 12*(3), 243-270.

Umansky, I. M. (2016a). Leveled and exclusionary tracking: English learners' access to academic content in middle school. *American Educational Research Journal, 53*(6), 1792-1833.

Umansky, I. M. (2016b). To be or not to be EL: An examination of the impact of classifying students as English learners. *Educational Evaluation and Policy Analysis, 38*(4), 714-737.

Umansky, I. M., & Dumont, H. (2021). English learner labeling: How English learner classification in kindergarten shapes teacher perceptions of student skills and the moderating role of bilingual instructional settings. *American Educational Research Journal*, *58*(5), 993-1031.

Umansky, I. M., Hopkins, M., & Dabach, D. B. (2020). Ideals and realities: An examination of the factors shaping newcomer programming in six U.S. school districts. *Leadership and Policy in Schools, 19*(1), 36-59.

Umansky, I., Hopkins, M., Dabach, D. B., Porter, L., Thompson, K., & Pompa, D. (2018). *Understanding and Supporting the Educational Needs of Recently Arrived Immigrant English Learner Students: Lessons for State and Local Education Agencies*. Council of Chief State School Officers.

Valenzuela, A. (1999). Subtractive Schooling: US-Mexican youth and the politics of caring. SUNY Series. *The Social Context of Education*.

Vazquez Cano, M., Umansky, I., Thompson, K. (2021). *How state, district, and school levers can improve the course access of students classified as English learners in secondary schools*. National Research and Development Center to Improve Education for Secondary English Learners. Retrieved from: https://www.elrdcenter.wested.org/improve-course-access-of-els.

Walker, A., Shafer, J., & Iiams, M. (2004). "Not in my classroom": Teacher attitudes towards English language learners in the mainstream classroom. *NABE Journal of Research and Practice, 2*(1), 130-160.

Walqui, A. (2000). *Access and Engagement: Program Design and Instructional Approaches for Immigrant Students in Secondary School*. Center for Applied Linguistics

White, R. S., & Mavrogordato, M. (2019). Educators' use of policy resources to understand English-learner policies. *Leadership and Policy in Schools*, *18*(4), 560-590.

Yonezawa, S., & Jones, M. (2006). Students perspectives on tracking and detracking. *Theory into Practice*, *45*(1), 15-23.

Yoon, S. Y., & Strobel, J. (2017). Trends in Texas high school student enrollment in mathematics, science, and CTE-STEM courses. *International Journal of STEM education, 4*(1), 1-23.

CHAPTER III: THE ESTIMATED IMPACT OF A STATE ENGLISH LEARNER ACCOUNTABILITY AND SUPPORT POLICY ON DISTRICT AND STUDENT OUTCOMES IN OREGON

Across education agencies there is a shared responsibility to ensure that students classified as English learners (EL) have equitable access to meaningful learning opportunities (Equal Educational Opportunities Act, 1974). Towards this responsibility, federal, state, and local education agencies all shape EL education through policy development and implementation. The state holds a large role in this process (Mitra, 2017). There are many pathways through which the state can shape EL education at the local level (Umansky & Porter, 2020). This includes setting accountability metrics, providing professional development, allocating funds, and developing supports and laws (Martínez & Spikes, 2020; Mitra, 2017; Tanenbaum et al., 2012; Umansky & Porter, 2020). However, we have limited evidence on the impacts of many of these policy levers—with a need for stronger evidence on ways state education agencies can provide effective, efficient supports that result in meaningful changes for EL-classified students. The EL-classified student population is growing and inequities in both access to opportunities and outcomes persist. Additionally, districts and schools continue to report that they feel under-supported in much of their EL work, with limited resources and guidance (Gándara et al., 2005; Sugarman, 2016; Tanenbaum et al., 2012). While state education agencies are positioned to play a large role in addressing inequities through EL policy decisions, they also operate within a resource-constrained environment (Tannenbaum, et al., 2012). Thus, research on effective, efficient state polices to support districts and schools in their EL work is a critical area for study.

In Oregon, the passage of a state-level policy, House Bill (HB) 3499, created the conditions to study how a state support and accountability policy impacted the outcomes of EL-classified students in struggling, high-need districts. As HB 3499 was signed into law in 2015, Oregon governor Kate Brown (2015) described the policy as

> …a law that I anticipate will be seen as a watershed moment for education in our state; one where we gave new focus to the needs of Oregon students who are becoming bilingual…demonstrating how cooperation can provide the opportunity for each student to reach her or his full potential.

The policy had multiple elements, including the development of a statewide strategic plan. However, a key element of the policy was focused on identifying and supporting districts struggling to meet their EL-classified students' needs. Through HB 3499, 40 districts were identified as both struggling to meet their EL-classified students' needs and facing high levels of need given their EL-classified student population. These 40 districts received a multi-pronged intervention from 2016/17 to 2019/20. The intervention consisted of additional funding and technical assistance, as well as increased accountability (Oregon Revised Statutes [ORS] § 604). Within the intervention, 25 districts were designated Target districts, a lower-intensity intervention, and 15 districts were designated Transformation districts, a higher-intensity intervention. At the end of the four-year cycle, the Oregon Department of Education (ODE) evaluated whether each district had met a series of state-established accountability indicators, with districts either exiting out of the intervention, entering monitored status, or entering a stage of directed funding, depending on their outcomes on the indicators. For districts under directed funding, ODE assumed control of their EL-specific expenditures for up to three years, or until the district met their accountability indicators.

EL-classified students have incredible assets and strengths too often constricted by structural and social barriers in U.S. schools (Valenzuela & Rubio, 2018). It is critical that education policy, including state education policy, move towards a better understanding of how to support local education agencies in appropriately serving EL-classified students. In this study, I examine the impact of HB 3499 on a set of proximal and distal outcomes—with the goal of capturing a range of ways in which the policy may have translated into change within identified districts in the first three years following policy implementation. By looking at proximal outcomes, such as EL expenditures per EL-classified student and the ratio of teachers who are ESOL-endorsed to EL-classified students, I aim

to capture effects that may not immediately present themselves as changes in student outcomes but represent shifts that may eventually impact EL-classified student outcomes. I also look at distal outcomes, including the likelihood that an EL-classified student receives bilingual education services and students' English language arts (ELA) and math outcomes. These outcomes are potential evidence the policy impacted instructional environments and learning outcomes of EL-classified students.

An analysis of HB 3499 can yield important information for state agencies nationwide interested in policy approaches to strengthen EL services and outcomes. These analyses will also provide important information for Oregon policymakers interested in understanding the policy impacts of HB 3499, both to evaluate the policy implementation as well as inform work with future HB 3499 cohorts. Additionally, HB 3499 was not a uniform intervention—some districts received more funding and technical assistance than others. As such, it may be of interest to see whether effects descriptively differed by whether they were estimated for districts that received a higher-intensity treatment or a lower-intensity treatment. In this study, I ask the following research questions regarding policy effects in the first three years of the policy:

1. What is the estimated effect of HB 3499 identification on:
    a. district EL expenditures per EL-classified student?
    b. the ratio of teachers who hold an ESOL endorsement to EL-classified students?
    c. the probability that an EL-classified student receives bilingual services?
    d. EL-classified students' ELA and math standardized assessment scores?
2. Do estimated effects of HB 3499 identification differ across Transformation and Target districts?

In the following section, I provide a brief overview of research on EL funding, technical assistance, and accountability—the three intervention elements of HB 3499. I also highlight research on a policy of similar design but focused on whole-school performance—School Improvement

Grants (SIG). I then describe the HB 3499 policy in more detail before outlining a theoretical framing for why it is expected that the policy would impact the outcomes studied. I move to a description of the data and analytic approach before presenting results and concluding with a discussion.

Background on State English Learner Policy

There are many ways in which state governments shape EL education (Umansky & Porter, 2020). HB 3499 integrates three types of policies into one overarching intervention—the provision of additional funding specific for EL expenses, the provision of technical assistance focused on EL education, and the increased accountability regarding EL-classified students. In this section, I outline research on each of these policy levers in EL education, contextualized within the current policy landscape. I close the section with a brief review of research on whole-school improvement grants, an intervention model with important parallels to HB 3499.

Funding

Funding is an important policy lever that facilitates the legally obligated supports and services that come with EL classification (Jimenez-Castellanos & Topper, 2012; Sugarman, 2016). There is strong evidence that funding impacts outcomes for all students, with additional funding linked to improved academic and labor market outcomes for students (Jackson, 2020). While there is no causal study of the impact of additional funds for EL-classified students specifically, there are multiple studies that examine the impact of additional targeted funds for students facing challenges such as poverty or educational disadvantage. For example, evidence from North Carolina's Disadvantaged Student Supplemental Fund suggests that additional targeted funds for districts facing high need regarding student poverty, low academic achievement, and challenges in retaining teachers resulted in academic improvement for all students, including those who had entered high school not proficient in math or ELA (Henry et al., 2010). Similar work has examined the impact of Title I eligibility, where funds are allocated to districts serving large populations of economically disadvantaged students and families

(Jackson, 2020). There is mixed evidence of the impact of Title I, as the impact of being eligible for these additional funds did not lead to improvements for low-income districts in New York (van der Klaauw, 2008), but there is suggestive evidence in other states that the funds resulted in improved educational outcomes (Jackson, 2020).

There is reason to believe that additional funds may be particularly critical for schools facing high-need regarding their EL-classified student population. With educational goals and rights that, at a minimum, encompass both grade-level content standards and the development of the English language, there is consensus that additional funds are required above and beyond those allocated per pupil to provide the level of supports and services necessary for EL-classified students' core rights to be met (Jimenez-Castellanos & Topper, 2012; Sugarman, 2016). Key services that have been linked to beneficial outcomes for EL-classified students and require additional funds include bilingual services (Steele et al., 2018), effective teachers and teacher professional development (Sugarman, 2016), co-teaching models (Chitiyo, 2017), and specialized curricula (Okhremtchouk, 2017). As research finds positive impacts of additional funding for students facing economic or education disadvantage (Henry et al., 2010), such benefits may extend to EL-classified students who also often are experiencing overlapping challenges.

Both the federal and state government play a role in providing additional funds for EL expenditures. The majority of federal funds specific to EL-classified students come through Title III (Sugarman, 2016). Title III funds are allocated to each state based on their demographics; states then distribute funds to local education agencies based on their EL-classified student and immigrant student (arrived within the last three years) populations (Sugarman, 2016). Title III funds must be used for a set of allowable expenses, and "recipients may not use those [Title III] funds to pay for services that, in the absence of Title III funds, would be necessary to be provided by other Federal, or State, or local funds" (U.S. Department of Education, 2016, p. 2). Despite growing EL-classified student

populations, Title III funds have largely remained flat over the last years (Williams, 2020). Specifically, Williams (2020) reports that Title III was appropriated 737.4 million dollars in 2016, roughly $150 per EL-classified student, down from the estimated $175 per EL-classified student in the early 2000s.

State policies also influence the amount of EL-specific funds that districts receive. The three most common state-level approaches for determining the amount of funding to provide per EL-classified student include: a) a weighted funding formula that provides additional funds weighted by the number of EL-classified students, b) allocating additional categorical funds; and c) reimbursement methods (Education Commission of the States [ECS], 2021; Sugarman, 2016). Weighted formulas are the most common, with varied weights (Sugarman, 2016). Some weights are the same for all EL-classified students. In Oregon, for example, EL-classified students are counted at an additional 0.5 weight in the state funding formula (ORS § 327.013). In Florida the weight is 0.199, in Rhode Island 0.1, and in Maryland 1.0 (ECS, 2021). Weights may vary by students' English proficiency level, as in Hawaii, or by the number of EL-classified students in each school, as in Maine (ECS, 2021). Other states, such as West Virginia, use a categorical grant or resource allocation approach, while states such as Wyoming and Virginia use a model where funds are allocated for EL teachers following a determined EL-classified student/teacher ratio (ECS, 2021).

Variation in state EL funding allocation is further compounded by differences in overall state and local funding levels (Sugarman, 2017). There are also other sources of funding that may support programs and services that EL-classified students participate in or benefit from, such as state-provided bilingual education grants or Refugee School Impact Grants (Office of Refugee Resettlement, 2017), although these are not tied to EL-classification specifically. Across this variation in funding sources and approaches, many argue that EL-specific funds are typically lower than what is necessary to provide the necessary services and supports (Jimenez-Castellanos & Topper, 2012; Okhremtchouk,

2017). Additional funds, especially if targeting high-need districts, may be a critical source of supports for districts serving EL-classified students.

Technical Assistance

Another important state-provided intervention to support EL education is technical assistance. Technical assistance in education broadly refers to capacity-building efforts provided to districts and schools (Strunk & McEachin, 2014; Strunk et al., 2014). States are often the primary source for district and school technical assistance, either through directly providing assistance or contracting external providers to provide the support (Strunk & McEachin, 2014). EL-focused technical assistance is an important source of support for districts and schools (Linquanti, 2014). A Title III evaluation from the early 2000s found that 45 states reported that they provide technical assistance focused on EL-classified students (Taylor et al., 2010), while an updated version reported all fifty states and Washington DC provided technical assistance (Tanenbaum et al. 2012). Common topics included standards and assessments, testing accommodations, instruction, finances, data use, and professional development. While important, there are gaps between the amount of technical assistance schools report they need and what is provided (Tanenbaum et al., 2012; Taylor et al., 2010).

There is limited research on the impacts of technical assistance. As an exception, state-provided intensive technical assistance in California positively impacted students' math scores as well as closed select achievement gaps between marginalized students and their peers (Strunk et al., 2014; Strunk & McEachin, 2014). Districts and schools report that they need more support and guidance in order to provide the necessary services for EL-classified students (Gándara et al., 2005), with the state well-positioned to fill this role through the provision of technical assistance (Linquanti, 2014).

Accountability

A third policy lever in EL education is accountability. In education, accountability systems broadly encompass a set of policy parameters that identify a set of educational objectives and put into place either rewards or sanctions for performance on these objectives (Levin, 1974). The study of accountability policy impacts on student outcomes has yielded mixed findings, with effects that range from negative, to null, to positive. (Chiang, 2009; Figlio & Ladd, 2008; Holbein & Ladd, 2017). In EL education specifically, accountability policy research has focused largely on the potential for negative effects of whole-school accountability measures on EL-classified students' learning opportunities and outcomes (Menken & Solorza, 2014; Palmer & Snodgrass Rangel, 2011). In particular, the introduction of high stakes testing and accountability policies may lead to a focus on test preparation, crowding out opportunities for rich instruction on the English language or bilingual instruction focused on the target language (Palmer & Snodgrass Rangel, 2011). In New York, for example, testing and accountability pressures led to movement away from bilingual programs to English-only programs amidst test score concerns (Menken & Solorza, 2014). However, these described accountability systems have focused on whole-school performance, not EL-classified students' performance, specifically. Additionally, the inclusion of EL-classified students in accountability systems, either through required benchmarks for English language proficiency development or disaggregated reporting for other measures, brings important institutional attention towards an underserved population (Morita-Mullaney & Singh, 2021).

EL-classified students are present in federal accountability policy in important ways. Under the most recent Elementary and Secondary Education Act (1965), the Every Student Succeeds Act ([ESSA]; 2016), federal accountability for EL-classified students' linguistic growth shifted from Title III to Title I, meaning that the accountability parameters no longer just apply to districts receiving Title III funds, but would fall under the same accountability process as Title I districts (Goldschmidt & Hakuta, 2017). The accountability parameters that apply to EL-classified students focus primarily

on English language proficiency growth and status (ESSA, 2016). As an additional accountability parameter, however, states must include in their statewide accountability systems performance outcomes for different student groups, including EL-classified students. These outcomes include academic performance, high school completion, and school quality or student success (Sugarman, 2020).

While ESSA (2016) outlines the broad strokes of accountability for EL-classified students, many specifics are determined at the state level (Goldschmidt & Hakuta, 2017). The accountability framework decisions held at the state level include the type of model used to quantify English language proficiency growth, how to incorporate recently arrived immigrant students, the minimum sample size at which reporting is required, and the goals for student progress (Goldschmidt & Hakuta, 2017). Additionally, states may contribute to accountability frameworks outside of ESSA, such as school turnaround efforts (Morando Rhim & Redding, 2014). State accountability policies can shape the narrative around what is important in schools, as well as introduce strong incentives to shape local behavior (Morando Rhim & Redding, 2014).

School Improvement Grants

The integration of these three policy levers described above—funding, technical assistance, and increased accountability—is similar in design and approach to policies implemented through the School Improvement Grant (SIG) program (U.S. Department of Education, 2016). The SIG program is a federal program aimed at identifying the lowest performing schools in the country for intervention—including funding, technical assistance, and increased accountability pressures (Morando Rhim & Redding, 2014). Importantly, while the intervention components are similar, HB 3499 is focused on EL-classified students specifically, while SIG is focused on whole-school improvement efforts (Title 1 § 1003(g)). Additionally, the SIG interventions are much larger in scope (Morando Rhim & Redding, 2014). Despite these differences, understanding the research on the

impacts of SIG identification can help to contextualize the study of HB 3499 as a policy of similar design, but narrower focus and smaller scope.

Research on the impact of being identified for the SIG program yields mixed findings. Positive effects of SIG identification were found in California (Dee, 2012; Sun et al., 2017), Kentucky (Bonilla & Dee, 2020), and Ohio (Carlson & Lavertu, 2018), and either null, negative, or mixed effects in Texas (Dickey-Griffith, 2013), Louisiana (Dee & Dizon-Ross, 2019), Rhode Island (Dougherty & Weiner, 2019), and North Carolina (Heissel & Ladd, 2016; Henry et al., 2015). Improvement efforts may not be observed in student outcomes until multiple years after identification (Sun et al., 2017), and positive impacts were concentrated among the districts that implemented the most intensive school turnaround models (Dee, 2012). Key takeaways from this work are that school context likely plays a role in determining whether or not the intervention yields meaningful change, and that impacts may not show up until years after intervention.

House Bill 3499

Oregon has experienced a significant demographic shift, seeing increased racial, ethnic, and linguistic diversity. This includes an increase in the proportion of students who are current and former EL-classified students (ODE, 2019). While noting academic gains over the past years, EL-classified students in Oregon face disparities in academic outcomes in comparison to never EL-classified students (ODE, 2019). In response to a growing EL-classified student population and persistent disparities in outcomes, HB 3499 was passed in 2015. A policy component required that 12.5 million dollars be allocated to the Statewide English Language Learner Program Account each biennium. Funds from this account would then be allocated to cohorts of districts identified as HB 3499 districts. Some districts would be Transformation districts, set to receive a more intensive intervention, and some would be Target districts, set to receive a lighter touch intervention.

HB 3499 Identification

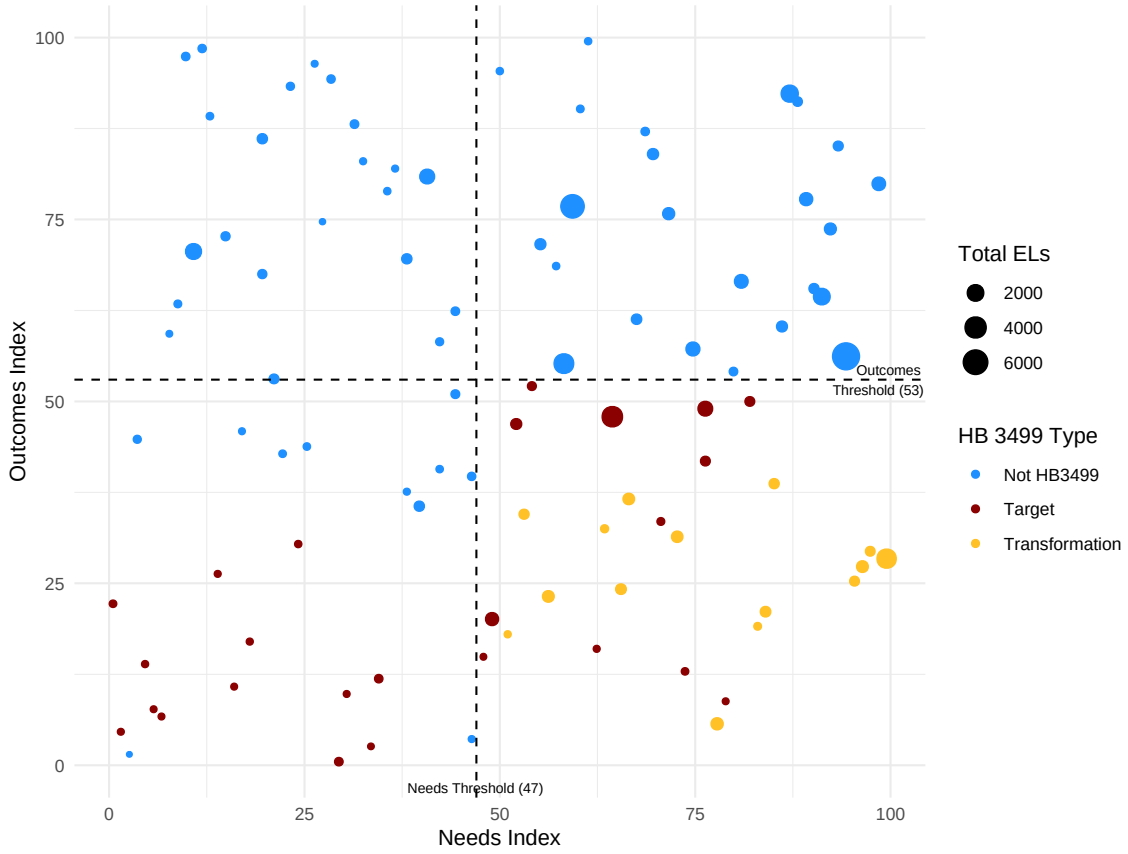
With guidance from an external stakeholder group, ODE constructed and applied a needs index and outcomes index using data from 2013/14 and 2014/15 to identify the first cohort of districts. The needs index was an unweighted combination[11] of districts' percent of ever EL-classified students, the percentages of ever EL-classified students who were economically disadvantaged, homeless, migrant, recent arrivers (arrived within the last three years), and mobile (attended more than one school within the school year, exited without earning a diploma, or had a significant enrollment gap), and the district small area income and poverty estimate (a measure that combines total population, number of children ages 5 to 17, and number of related children ages 5 to 17 in families in poverty into a singular value). The outcomes index was a weighted combination[12] of current EL-classified students' English language proficiency growth (excluding kindergarten; weighted 0.45), the five-year ever-EL adjusted cohort graduation rate (weighted 0.35), growth on the standardized math assessment for ever-EL-classified students in grades 6-8[13] (weighted 0.15), and post-secondary ever-EL enrollment (weighted 0.05). Both indices were scaled 0 to 100, with 0 as low and 100 as high.

These indices guided identification. All districts that had fewer than 20 EL-classified students were not eligible for identification, regardless of their combined needs/outcomes index. Of the districts that did have 20 or more EL-classified students and were eligible for identification (n=97), those above the 47th percentile in need and below the 53rd percentile in outcomes were all identified as the first cohort of HB 3499 districts, although ODE was able to adjust selection beyond the intersection of the two indices (ORS § 339.079). The end result was that the threshold guided

[11] To create a single value, a district's value for each of the above elements was transformed into a percentile rank in relation to other districts; these percentile rank values were then summed, then again converted into a percentile rank value, creating a score on a scale of 0-100, with higher values indicating higher need (Personal Communication, 2020).
[12] To calculate a single measure for each district, each variable was multiplied by its respective weight, those values were summed, then divided by the total possible points to produce the percentage of points earned. This value was then converted into a percentile.
[13] Excluding current EL-classified students whose EL start date occurred on or after July 1, 2013 or whose current English language proficiency performance levels were '1' or '2' and did not use specific accommodations on the mathematics assessment

identification, but identification was not perfectly aligned with these two criteria. Figure 2.1 maps out identification by the needs and outcomes indices, with the shaded grey area denoting the area where treatment was intended. All districts that were above the 47th percentile in need and below the 53rd in outcome were identified; 12 were Target districts and 15 were Transformation districts. There is no clear threshold used to distinguish whether a district would be Target or Transformation, although it the policy states, "The Department may use the demonstrated commitment level of a district's superintendent and board as a factor in determining whether the district is an ELL transformation or ELL target district" (OAR § 581-020-0613). Additionally, 13 districts that were below the outcomes threshold (53rd percentile) and below the needs index threshold (not above the 47th percentile) were identified as Target districts, while 12 other districts in the same quadrant were not.

Figure 2.1 District Needs and Outcomes Indices with Intended Treatment Area Shaded, Weighted by Number of EL-classified Students in Analytic Sample



HB 3499 Intervention

In total, 40 districts were identified as the first cohort in the 2016/17 academic year, with treatment continuing through the 2019/2020 academic year. Fifteen were Transformation districts, set to receive $180,000 per year and 25 were Target districts, set to receive up to $90,000 per year (Carnock, 2017). HB 3499 required that districts clearly track their expenditures (ORS § 39.079). Broadly, HB 3499 funds were intended to support professional development, increasing the number of ESOL or bilingual endorsed teachers, monitoring, instructional resources, hiring of additional staff, and extended learning time opportunities, although there were not specific requirements other than on EL-classified student expenditures (Jimenez-Castellano et al., 2020).

As funds were awarded, ODE was positioned to provide on-going technical assistance for the years of identification, including support in identifying and evaluating district needs, providing further individualized support, and overseeing annual professional development (ORS § 39.079). Specifically, for Transformation districts this included an evidence-based needs assessment, identifying relevant interventions, guiding implementation and monitoring efforts, and reflecting on the intervention, while engaging with administrators, staff, families, and community interest holders (OAR § 581-020-0613). Target districts were also tasked with evaluating their EL programs through a needs assessment process with support from ODE, although the policy outlined that individual technical assistance would be provided to Target districts only if resources were available after providing technical assistance to Transformation districts (OAR § 581-020-0615).

HB 3499 also included an accountability element. Based on districts' progress on a set of indicators (described below), after four years districts would fall into one of three categorizations—directed funding, monitored, or reclassified (ODE, 2021a). Directed funding entailed a process by which ODE assumed control of funds received under ORS 327.013(1)(c)(A)(ii), which allocates an additional 0.5 per pupil funding allocation for each current EL-classified student. This spending

control would last up to three years if districts continued to not meet their expected targets (ORS § 39.079). Monitored status also would last up to three years—districts could be transitioned into reclassification status if the district demonstrated significant progress or could be transitioned into directed funding status if they do not demonstrate significant progress (ODE, 2021b). Reclassified status entails having completed the intervention.

The process by which a district's final status at the end of four years was determined shifted during the intervention period in response to community and district feedback (ODE, 2021b). In the final model, ODE put into place a set of indicators that districts would be evaluated on regarding EL-classified student outcomes, English language proficiency growth, regular attendance, exclusionary discipline, ELA achievement, ELA growth, math achievement, and math growth. Districts received points for their values on each of these indicators at both the secondary and elementary level in relation to a set threshold, and these points were then multiplied by weights. Districts could also earn an additional bonus point if ever-EL outcomes were equal to or better than never EL outcomes on that indicator, where applicable. Districts then were rated as either making *notable progress* if their point value was 75% or more of the weighted points available, *some progress* if their point value was 50-74.9% of available points, or *limited progress* if their point value was less than 50%. Districts received a combined point value for their elementary indicators as well as their secondary indicators and an overall rating for both. If districts were rated as limited progress on both, they were to be placed in the directed funding category. If districts meet one of the two, they were to be placed in monitored status. If they met both, they were reclassified (ODE, 2021b).

A preliminary summary of areas in which districts concentrated their efforts revealed that the most prevalent focus area was professional development (83% of HB 3499 districts), followed by parent/family engagement (75%), student supports (55%), and instructional staff (33%), while other areas included extended learning, teacher certifications, school climate, and immersion programs

(ODE, 2020). One in-depth report focused on a Target district found that the district primarily focused on professional development, hiring consultants to work with teachers on instructional strategies for academic language development (Pacific Research and Evaluation, 2020).

As of May 2022, three districts were in the directed funding category, 15 in monitored status, and 22 reclassified (ODE, 2022). By these measures, the majority of HB 3499 districts met many important goals for their EL-classified students and were successful in improving their practices and outcomes. This current study is not meant to take away from this evaluation, but rather add another lens through which to evaluate the impact of HB 3499 on outcomes, while accounting for other trends in the state by using non-identified districts as the counterfactual. This study also includes outcomes outside those considered in the evaluation—specifically including three process-based measures (expenditures, proportion of teachers who are EOSL endorsed, and likelihood of receiving bilingual services) that may include important insights into changes among HB 3499 districts.

Theoretical Framework

There are numerous pathways through which HB 3499 identification could be theorized to impact the outcomes studied. Here I outline how theory suggests each HB 3499 policy lever may impact the outcomes examined in the study.

Funding

The provision of additional funding specifically earmarked for EL services with required budgeting stipulations may lead to changes in EL spending and the instructional environment, as well as academic outcomes. The link between additional funds and EL expenditures is clear, with the expected outcome that EL-specific expenditures increase. However, given that the amount is not as large as other targeted funding interventions such as SIG or Title I (Yatsko et al., 2012), it remains to be seen if the additional funds will result in a significant increase in EL expenditures. For example, schools identified for SIG in 2010 through 2012 often saw their budgets double, receiving between

$50,000 and $2,000,000 per year (Sun et al., 2021), while Title I funds allocate around an additional $1,000 per student (U.S. Department of Education, 2019). As further context, the additional 0.5 funding weight for EL-classified students translates to roughly an additional $4,200 per EL-classified student per year. (Jimenez-Castellanos et al., 2020). In comparison, $180,000 or $90,000 per district may not be enough to see significant spending shifts, especially in districts facing resource constraints. Additionally, if districts use funds in a manner where they are supplanting, rather than supplementing, EL expenditures, there may not be a significant increase in EL expenditures per EL-classified student.

I also examine an outcome related to teacher training and certification—the ratio of teachers who hold ESOL endorsements to EL-classified students. Given that professional development and teacher certifications were named as potential areas for investment through HB 3499 (Jimenez-Castellanos et al., 2020; ODE, 2021a), it could be expected that additional funds were directed in ways to support teachers in accessing ESOL endorsements, or to hire more endorsed teachers. A third outcome that I examine is the probability that an EL-classified student receives bilingual services. As bilingual programs benefit EL-classified students' academic and linguistic outcomes but require funds above and beyond other program models to implement (Steele et al., 2018), HB 3499 identification may be an important source of funding to support the expansion of current bilingual programs or provision of new programs. However, in relation to other funding sources for bilingual expansion, such as Oregon's dual language/two-way bilingual grant program (OAR § 581-018-0215), HB 3499 may not have represented an intensive enough intervention to see meaningful shifts in bilingual service provision once changes are differenced out from non-identified districts.

There are also reasons to believe additional funds may impact EL-classified students' academic outcomes. Additional funding broadly has been linked to improved student academic outcomes (Jackson, 2020) and targeted funding for underserved students has been found to positively impact student academic outcomes (Henry et al., 2010). Additional funds allocated through HB 3499 may

provide districts with the funds to invest in resources such as hiring of staff focused on EL services or instruction, multilingual resources, afterschool tutoring for EL-classified students, and materials for engagement with families of EL-classified students. If such expenses are spent on evidence-based costs (Zarate & Gàndara, 2019) and sufficient to support EL-classified students across the district, they are likely to support improved instruction and outcomes, potentially evidenced through improved performance on standardized ELA and math assessments.

Technical Assistance

In addition to funding, HB 3499 identification includes state-provided technical assistance, which theory suggests could strengthen both instruction and outcomes. Technical assistance is an important pathway through which state education agencies support local EL work (Linquanti, 2014; Tannenbaum et al., 2012; Taylor et al., 2010). As described above, the state worked with districts to conduct needs assessments, develop strategic plans, and set goals for improvement, while also providing annual supports and building relationships focused on supporting local EL work. If such technical assistance is focused on identifying opportunities to implement evidence-based supports, such as training educators, there may be an increase in the personnel who are endorsed in important areas, such as ESOL education, given the wide body of evidence on the importance of effective teachers (Chetty et al., 2014). If technical assistance is focused on supporting local contexts to provide evidence-based language instruction programs, we may see that bilingual education, a program model with some of the most consistent, robust evidence on the positive effects for EL-classified students (Porter et al., in preparation), becomes more widely implemented. We also might expect that state-provided technical assistance leads to improved academic outcomes due to improvements in instruction and overall services. This would complement work on the impact of state-provided, intensive technical assistance in California, which had positive impacts on student outcomes (Strunk et al., 2014; Strunk & McEachin, 2014).

Accountability

Beyond funding and technical assistance, the additional level of accountability that comes with HB 3499 identification may impact instructional environments and student outcomes. In response to the threat of losing discretion of their EL funding, districts may use expenditures in ways they feel can support student outcomes, such as investing in ESOL-endorsed teachers, professional development, and instructional services such as bilingual education to support student outcomes.

Many federal and state education policies have high-stakes accountability elements, with mixed results as to the impacts of accountability on school-level performance (Chiang, 2009; Figlio & Ladd, 2008; Holbein & Ladd, 2017). Qualitative work has highlighted that accountability pressures can alter teaching practices (Hamilton et al., 2008; Louis et al., 2005); in the context of HB 3499 accountability pressures may drive improvement in outcomes for EL-classified students if districts and schools adapt or change their practices in ways that support EL-classified students' academic outcomes, such as broader access to bilingual education and improved instructional approaches and materials. Conversely, accountability pressures could lead to a constriction of bilingual education opportunities and a loss of authentic English language development opportunities in exchange for "teaching to the test" if administrators perceive bilingual programs to be inhibiting early academic or English proficiency development (Menken & Solorza, 2014; Pandya, 2011).

Hypotheses

I hypothesize that HB 3499 identification will result in a small, but significant increase in the ratio of district EL expenditures to EL-classified students. While the funding amount is relatively low, it still represents an additional boost to a limited pool of funds. Districts are required to clearly track expenditures, which would limit, in theory, the risk that districts are redistributing other, non-HB 3499 funds used for EL expenses towards non-EL costs with the influx of HB 3499 funds.

I also hypothesize that HB 3499 districts will see, by the third year of identification, an increase in the ratio of ESOL-endorsed teachers to EL-classified students. Given that this was a focus area of the policy (Jimenez-Castellanos et al., 2020), increasing the number of ESOL-endorsed teachers seems like a malleable instructional policy lever that HB 3499 could address. However, earning an endorsement takes time, and I anticipate effects will not be observable until later year estimates.

I anticipate a small increase in the probability that an EL-classified student receives bilingual services, although given the scope of the intervention is relatively small, and there is likely heterogeneity across treated districts in how they chose to invest in EL services, I am not confident it will be statistically distinguishable from zero. Implementing new programs represents a significant undertaking for districts without existing programs. It is more likely that, if programs are already in place prior to identification, HB 3499 identification may be better positioned to support the expansion of existing programs, rather than the implementation of new programs. Only eight HB 3499 districts had any EL-classified students receiving bilingual services prior to identification, thus it may be difficult to identify a significant increase in enrollment if the majority of identified districts are facing the task of developing and implementing full new programs within a three year timeframe.

I anticipate null effects of HB 3499 identification on EL-classified students' academic outcomes in the first two years, but, as with ESOL teacher endorsements, by the third year I hypothesize that the event study will show positive effects. I anticipate that results using ever-EL students will show positive effects in earlier years. Improvements in teachers' instruction may take multiple years to result in improved academic outcomes, especially for currently EL-classified students. Given the nature of the EL student categorization, positive effects may be more likely to be observed if former EL-classified students are included in the analysis, as in a sensitivity analysis I conduct.

Finally, while an intervention larger in scope should, in theory, see more precise or larger effects of HB 3499 across the outcomes examined, I theorize that there will be no evidence that effects differ

across Target and Transformation districts. The key differences between the two treatments are the amount of funding and the prioritization of technical assistance. Given that funding amounts are relatively small in relation to the scale of education funding, I theorize that technical assistance and accountability will play a larger role in shaping district improvement than funding. I anticipate that ODE will have provided technical assistance in a way that did not vary widely by whether a district was a Target or Transformation district. If this is the case, then I anticipate Target or Transformation designation will not be associated with differential effects.

Data

I draw on a set of Oregon statewide records to conduct this study. For expenditure, teacher endorsement, and bilingual program outcomes, I use ODE student and district-level data from 2013/14 to 2018/19, with years 2013/14-2015/16 establishing pretreatment trends and years 2016/17-2018/19 establishing post-treatment trends. I only use data from Oregon student-level records from 2014/15 to 2018/19 for academic outcomes because of a shift in the test from the Oregon Assessment of Knowledge and Skills (which had different testing requirements) to the Smarter Balanced Assessment. While the intervention runs through 2019/20 for the cohort, I did not have access to 2019/20 data because of pandemic related disruptions. I only include data from districts that had 20 or more EL students, a baseline criterion for eligibility for HB 3499 identification, a total of 97 school districts.

For student-level data I dropped all observations that were missing data on eligibility for free/reduced price lunch, race/ethnicity, gender, or language for EL-classified students after imputing for missing values using modal values from students' full available data in the dataset. I also drop observations that are missing data on whether a student was identified for special education services

or EL-classified. I drop any student who, by way of transferring districts within the state, moved from treatment to control, or control to treatment post-identification.[14]

Table 2.1 reports descriptive statistics from the year prior to identification at the district level by whether districts were non-HB 3499 districts, Target districts, or Transformation districts. On average, non-HB 3499 districts had larger EL-classified student populations. Target districts tended to have smaller EL-classified student populations, as well as smaller student populations overall, while Transformation districts, on average, had a higher proportion of their students who were EL-classified, as well as larger EL populations than Target districts.

Table 2.1 District-Level Descriptive Statistics by Identification Type, 2015/16

	Not HB 3499	**Target**	**Transformation**
	M	***M***	***M***
EL expenditures	$2,007,932.42	$1,032,511.40	$1,536,099.54
EL exp/EL	$3,041.13	$3,213.91	$3,087.16
EL exp/Total expenditures	0.02	0.01	0.03
Total EL	637.75	323.52	567.07
District % EL	0.09	0.06	0.13
District % Ever-EL	0.18	0.13	0.24
District % special education	0.14	0.14	0.14
District % nonwhite	0.35	0.28	0.44
District Size	5,802.86	4,334.16	5,313.40
ESOL endorsed teachers	0.14	0.19	0.10
Needs index	48.10	40.26	76.47
Outcomes index	69.55	22.33	26.36
N	57	25	15

Note. All school districts in Oregon were eligible for potential HB 3499 identification if they enrolled 20 or more EL students in the 2014/15 school year, a total of 97 school districts. EL=English learner. HB=House bill. FRPL=Free/reduced price lunch. Statistics report the average value of student or district-level characteristics in the year prior to HB 3499 identification by whether the district was identified in the following year.

District EL Expenditures

[14] In total, 4,013 EL-classified student-year observations (1,207 EL-classified students) were dropped due to transferring in the data.

A primary outcome in this study is district EL expenditures per EL-classified student. Districts provide expenditure data to the state, detailing expenditures by function, which identifies the overall spending category, and object, which captures the service or commodity bought. I used the ODE (2019) public budgeting manual to identify, by a combination of function and object codes, expenditures specific to EL education. This was primarily denoted through the overall function code being 1291[15] or 1295[16] although other special cases outlined in the manual identified further EL-specific expenditures, typically those in translation and interpretation services.[17] The HB 3499 policy states that Transformation districts were set to receive $180,000 per year, and Target districts $90,000. When dividing the projected amount by the total EL-classified students in a given year for post-treatment district/year observations across years, the average additional per-EL funding amount for Transformation districts is about $880/EL-classified student a year. For Target districts the amount is higher, about $2,000/EL-classified student a year.

I summed EL-specific expenditures within district and year, then divided the total by the number of EL-classified students for each district-year observation. After dropping five district/year observations due to missing data, the main analytic sample includes 577 district/year observations from 97 unique districts. Trends over time in EL-expenditures per EL-classified student are plotted in Figure A1, Panel I, by HB 3499 districts and non-HB 3499 districts. Overall, EL expenditures per EL-classified student increased over time across districts. HB 3499 districts spent more on EL expenditures per EL-classified students across the time series, although the difference visibly increases after identification. On average, HB 3499 districts spent between $2,678.53 and $4,832.32 a year on EL expenditures per EL-classified student, while non-HB 3499 districts spent between $2,431.14 and

[15] As per ORS 336.079, refers to instructional activities for EL students used in acquisition of the English language.

[16] Instructional Activities not related to ORS 336.079 for EL students used in acquisition of the English language.

[17] This included expenditures with function codes other than 1291 and 1295, but an area of responsibility code of 280, or a function code of 2680 with an object code of 319 or 389 and no area of responsibility code.

$3,802.66. In the year prior to identification, HB 3499 districts spent roughly $125 more in EL expenditures per EL-classified student. For reference, the overall national average K-12 per-pupil expenditure rate was $13,187 in 2018/19, while in Oregon the average was $12,450 per pupil (U.S. Census, 2021). When HB 3499 districts are broken down into Target and Transformation districts (Figure A1, Panel II) there is a larger increase in spending among Target districts.

Looking just at HB 3499 districts, descriptive information on EL expenditures across the panel years is displayed in Figure A2, Panel 1. In this figure, I plot average district EL expenditures by the expenditure object. Importantly, the plotted average expenditures represent all EL expenditures, not those exclusive to HB 3499 funds. As seen, EL expenditures are predominantly salary and payroll costs, with costs in this category increasing over time. A distant second is retirement, benefits, and social security-related expenditures, followed by administrator costs. In Panel II I remove salary and retirement expenditures to allow for a more detailed look at other EL-specific expenditures. Among these lower-expenditure categories, the highest average expenditures were concentrated among professional, instruction, and technical services, temporary staff/substitutes, classified managerial expenses, and supplies and technology. On average, after identification HB 3499 districts spent more on EL-specific professional, instructional, and technical services, supplies and technology, classified positions, textbooks, and other expenditures than in the years prior to identification.

Teacher Endorsement

I also examined the district ratio of teachers who held an ESOL endorsement to EL-classified students. In ODE data, teachers are flagged as holding an active ESOL endorsement, which is focused on providing teachers with instruction on how to support students' English language development (ORS § 584-420-0360). I summed the number of teachers holding these endorsements at the district level in a given year, then divided the total by the number of EL-classified students in the district in that year. The total sample is 582 district/year observations and 97 unique districts, with no missing

data points. Trends in the ratio are plotted in Figure A3. As with expenditures, the first panel represents the main sample, while the second breaks out treated districts into Target and Transformation districts. Overall, there is an increase in the ratio prior to identification among both HB 3499 districts and non-HB 3499 districts, increasing from below one endorsed teacher per ten EL-classified students in 2014 to up to around one endorsed teacher per seven or so EL-classified students in 2019. As seen in Panel I, HB 3499 and non-HB 3499 districts were on different trajectories prior to identification. The increase in endorsed teachers/EL-classified students was steeper among HB 3499 districts than non-HB 3499 districts. As seen in Panel II, the steeper increase is concentrated among Target districts, which had the highest ratio among the three groups of districts, while Transformation districts and non-HB 3499 districts were on parallel trends prior to identification.

Bilingual Services

Another set of analyses focus on the probability that an EL-classified student received bilingual education services to support core content access. Districts report the language instruction programs used to support core content access—with options that include different bilingual program models. I collapse the different types of bilingual education (transitional, developmental, two-way immersion, maintenance, and other) into an overall bilingual services code, coded 1 for students who received bilingual services to support core content access, and 0 for those who did not. In total, the main sample includes 291,860 student/year observations, 94,950 EL-classified students. Table 2.2 reports descriptive statistics for the different student samples represented in the study, providing average characteristics for the EL-classified student population observed the year prior to identification, by whether they were in an HB 3499 district or a non-HB 3499 district. On average, across the bilingual services sample, a higher proportion of EL-classified students in HB 3499 districts were eligible for free/reduced price lunch, while a lower proportion were Latinx/Hispanic and White and a higher proportion Black and American Indian/Alaska Native. EL-classified students in HB 3499

districts tended to be in districts where the overall student population was comprised of a slightly lower proportion of ever-EL students and slightly higher proportion of students FRPL eligible.

Trends in the proportion of EL-classified students in bilingual services over time are plotted in Figure A4. Overall, a higher proportion of EL-classified students in non-HB 3499 districts received bilingual services across the years examined. From 2013/14 to 2018/19, more than 20 percent of EL-classified students in non-HB 3499 districts received bilingual services, compared to less than ten percent in HB 3499 districts. From 2013/14 to 2015/16, there appear to be slightly different pre-identification trends by HB 3499 identification, with a small average increase observed among HB 3499 districts and a small average decrease among non-HB 3499 districts. On average, in both HB 3499 and non-HB 3499 districts there was a drop in the proportion of students receiving bilingual services in 2016/17, the first year of identification. This is followed by modest increases over the next two years, although the raw gap between HB 3499 and non-HB 3499 districts remains roughly constant. As seen in Panel II, Target and Transformation districts were on different trends prior to identification. There was a decline in the proportion of EL-classified students receiving bilingual services among Transformation districts, while there was an increase among Target districts. Additionally, while there was an observed drop in both non-HB 3499 districts and Transformation districts in the first year of identification, there was no drop among Target districts.

Academic Outcomes

The academic outcomes analytic sample is constructed of EL-classified students in grades 3-8[18] who took the spring-administered standardized academic assessment. For the years of data examined, Oregon administered the Smarter Balance Assessment Consortium (SBAC) assessment; I use math and ELA scale scores to measure impacts on students' academic achievement. Results are vertically scaled, with possible scores that range from 2000 to 3000. I drop any student in grades 3-8

[18] I excluded 11th grade test scores due to high levels of missingness and students who took an alternate assessment.

who is missing a test score in a given year. Scores are standardized using the full population of students in the state with assessment data within grade and year. In total, the main ELA analytic sample includes 117,580 EL-classified student/year observations and the math analytic sample 119,505 EL-classified student/year observations. Table 2.2 presents descriptive statistics for EL-classified students in the year prior to HB 3499 intervention in both the ELA and math analytic samples. There were similar differences as observed in the bilingual services sample, with differences in student racial/ethnic make-up and eligibility for free/reduced price lunch.

Table 2.2 Descriptive Statistics for EL-classified Students in Districts Eligible for HB 3499 Identification, by Outcome Examined, 2015/16

	Bilingual Services (K-12)		English Language Arts (3-8)		Math (3-8)	
	Non-HB 3499	HB 3499	Non-HB 3499	HB 3499	Non-HB 3499	HB 3499
Female	0.46	0.46	0.45	0.44	0.46	0.44
Identified for FRPL	0.89	0.93	0.91	0.96	0.91	0.96
Identified for special education	0.20	0.20	0.23	0.22	0.22	0.22
Race/Ethnicity						
AIAN	0.00	0.03	0.00	0.04	0.00	0.04
Asian/Pacific Islander	0.10	0.09	0.09	0.08	0.09	0.08
Black	0.02	0.04	0.02	0.03	0.02	0.03
Hispanic	0.78	0.77	0.81	0.78	0.80	0.78
Multi-Ethnic	0.01	0.01	0.01	0.01	0.01	0.01
White	0.10	0.06	0.08	0.06	0.08	0.06
% in bilingual services	0.24	0.09	--	--	--	--
ELA score	--	--	-0.94	-1.06	--	--
Math score	--	--	--	--	-0.84	-0.98
District characteristics						
% ELs	0.16	0.15	0.16	0.15	0.16	0.15
% Ever-EL	0.30	0.27	0.31	0.27	0.31	0.27
% Non-White	0.48	0.47	0.49	0.47	0.49	0.47
% FRPL	0.61	0.64	0.62	0.65	0.62	0.65
% Special Education	0.14	0.15	0.14	0.14	0.14	0.14
N	32,944	14,912	15,405	6,931	15,566	7,003

Note. All school districts in Oregon were eligible for potential HB 3499 identification if they enrolled 20 or more EL students in the 2014/15 school year, a total of 97 school districts. EL=English learner. HB=House bill. FRPL=Free/reduced price lunch. AIAN=American Indian or Alaska Native. ELA=English language arts. Statistics report the average value of student or district-level characteristics in the year prior to HB 3499 identification by whether the district was identified in the following year.

Figure A5 presents the line of best fit through the average math and ELA scores of EL-classified students over time, by whether districts were identified for HB 3499 or not (Panel I) and by whether districts were Target districts, Transformation districts, or not identified (Panel II). Additionally, Panels III and IV plot the same trends, but drawing on average scores for students who were ever EL-classified in the panel. On average, scores were higher in non-HB 3499 districts across the years observed, with parallel trends observed across the two pre-intervention and post-intervention data points. This holds for both EL-classified and ever-EL student samples. Transformation districts have lower scores than Target districts across all time points. For the sample of EL-classified students, scores jump in the year after identification among both HB 3499 and non-HB 3499 districts, then decline over the next two years. For ever-EL students, however, the scores trend upward over time.

For the academic outcomes, I primarily focus my analyses on students who were EL-classified, however, as I describe in the Sensitivity Analyses section, I re-run the main academic analyses using different student subsamples. The main alternative samples are students who were ever EL-classified (ever-EL) during the panel of data examined, and outcomes for current EL-classified students, as well as students who retained their EL classification in the first year of identification, regardless of whether they reclassified during the intervention period.

Analytic Strategy

Research Question 1

To estimate the impact of HB 3499 identification on district and student-level outcomes I first fit a dynamic two-way fixed effects event study, wherein I estimate a yearly treatment effect for identified districts in comparison to non-identified districts, relative to their pre-intervention year outcomes. In the event study specification, estimates are generated for each individual year without imposing a functional form and without pooling effects pre- and post-treatment (Angrist & Pischke, 2008). I also fit a difference-in-differences specification, which is similar to the event study

specification, but individual dynamic effects are no longer estimated for each year. Instead, effects are pooled pre- and post-treatment, yielding an overall average treatment effect across the post-intervention period. The event study estimation approach allows me to estimate the effect in each year, providing information on how the policy impact differs by year, or whether effects are detectable only later in the treatment period (Sun et al., 2021). The difference-in-differences estimates provide information on whether there was an overall average impact of HB 3499 on a set of outcomes.

The intuition behind the event study and difference-in-difference specification is that, in situations where treatment is assigned to one group but not another, the treatment effect can be recovered by assuming that, if treatment had not been assigned to the treated group, their outcome levels would have changed at the same rate as untreated units. More specifically, in contexts where treatment is not as good as random, the causal effect of treatment can be identified by differencing any observed changes in the outcome of interest post-identification among non-treated units from the observed changes in the outcome of interest post-identification among treated units, if key assumptions hold. To obtain event study estimates for each treatment year, I fit the following model for EL-classified students[19] in a given year:

$$Y_{it} = \sum_{r=-3}^{2} 1(t = t^*_s + r)\,\beta_r + \boldsymbol{X}_{it} + \boldsymbol{V}_{st} + \boldsymbol{Z}_{jt} + \rho_j + \sigma_g + \Gamma_t + \varepsilon_j \quad (1)$$

In this general student-level model (see footnote 9 for description of district level model with additional outcomes), Y_{it} is the outcome of interest—either probability of receiving bilingual services, standardized math score, or standardized ELA score, for student *i* in year *t*. In this approach, the parameters of interest are each β_r estimate, which capture the effect of HB 3499 identification on a given outcome of interest in year *r,* relative to the year prior to identification. This is represented in

[19] This is the student-level model. For expenditures and teacher endorsements I focus on district-level outcomes, with the district as the unit of analysis. I chose to focus on the district, rather than school, level given that expenditures are reported to the state at the district level, and for teachers there are instances where teachers may be supporting students across multiple schools within a district. Given that the district is the treated unit, this also aligns with the policy question of interest. For bilingual services and academic outcomes, however, I include students as the unit of analysis. This allows me to include student-level characteristics in the model, further accounting for variation in the outcome of interest not attributable to HB 3499 identification.

the model, where *r* takes on a set of values from -3 to 2, spanning three years prior to identification and two years post-identification. I include a set of plausibly exogenous individual covariates ($\boldsymbol{X}_{it}$): students' race/ethnicity, receipt of special education, eligibility for free/reduced price lunch, and gender. I also include time-varying district- ($\boldsymbol{Z}_{jt}$) and school-level ($\boldsymbol{V}_{st}$) covariates: percent of students who are ever EL-classified, non-white, identified for special education, and eligible for free/reduced price lunch, to improve model precision. I include grade (σ_g), district (ρ_j) and year (Γ_t) fixed effects.

While I include results from models that do not include the full suite of fixed effects presented above, the above two-way fixed effects model above is my preferred specification and I focus my discussion on results from this model. This is my preferred specification given that district fixed effects account for time-invariant variation attributable to the district a student is in and baseline differences in the outcome of interest. Grade fixed effects account for variation attributable to what may be differences in performance by grade level, important because districts may be serving different grade levels at different rates. Finally, year fixed effects account for yearly shocks, or unobserved factors specific to each year that have the same effect on all districts. Essentially, the two-way fixed effects approach is isolating the treatment effect through not only comparing changes in outcomes among treated and untreated units, but also further accounting for unobservable factors that may impact the outcome of interest specific to the district or time In all models, I cluster standard errors at the district level, the unit of assignment (Bertrand et al., 2004). This constrains power, as standard errors will reflect how the error terms are correlated within districts.

In addition to the event study approach, for each outcome of interest I also fit a difference-in-differences model. The model is almost identical to the event study modeled above, but effects are pooled pre- and then post-treatment by estimating a coefficient on the interaction between being in an HB 3499-identified district and being in a post-identification time period, rather than separate estimates for each year. This yields an estimated, average overall treatment effect of HB 3499

identification on the outcomes examined. This approach answers a slightly different question than the event study, moving away from individual yearly effects and estimating an overall, average treatment effect that captures the combined effect across years in relation to all pre-treatment years in the panel, rather than the time period just before identification. While losing information on the dynamic effects, the approach requires the estimation of fewer parameters, providing more power for the analysis and potentially more precise effect estimations. The two approaches together provide important insights into the nature of HB 3499 impacts, if observed.

Research Question 1: Robustness Checks

For the event study and difference-in-differences approaches to yield an unbiased estimate of the policy impact, several assumptions must hold. As modeled above, treated and untreated units must follow parallel trends prior to intervention, otherwise the approach may pick up significant differences in the outcome that are not due to the policy implementation, but rather to unobservable factors driving differential rates of change in the outcomes examined prior to identification for HB 3499. I examine this assumption in multiple ways. First, I examine the raw data, as well as pre-treatment estimates from event study plots for any evidence of significant pre-trends, conditional upon the model specification. However, given concerns about the inability of the visual test to detect pre-trends in underpowered contexts (Roth, 2019), I also implement various checks using the HonestDiD package to explore how estimates would change depending on the projected severity of post-identification trend violations as a function of pre-trend violations (Rambachan & Roth, 2022).

As another concern, estimates can be biased if simultaneous policy shifts occur that may impact the outcomes of interest. To explore this issue, I fit a triple difference model that, in addition to differencing out changes in EL-classified student outcomes among untreated districts, differences out changes in outcomes for students who were never EL-classified (either initially fluent English proficient or English-only students). To do so, I include data from both current and never-EL

students, then interact indicators for whether a student is EL-classified, in an HB 3499 district, and is in the post-identification period. This approach is designed to account for any other policy change that may have impacted the instructional environment for students broadly and confounded the estimate. I conduct this check for academic outcomes, the only outcome for which I have data on for never-EL students. I also conduct placebo analyses which are very similar to the triple difference models, fitting the main event study and difference-in-differences models using never-EL academic outcomes as the outcome of interest, as well as non-EL district expenditures. Any significant effects detected through the placebo analyses may signal other policy or contextual changes that are influencing instruction or expenditures outside of HB 3499.

There should also be no concern of anticipatory effects (Roth et al., 2022). Given that the indices used to guide treatment were complex, the values used to create the indices spanned two years, and districts were not notified in advance of identification (Carnock, 2017), I argue that it is unlikely that identified districts made changes in anticipation of identification.

In addition to the placebo and triple-difference specifications described above, I conduct a further set of sensitivity analyses for academic outcomes focused on accounting for the changing nature of the EL-classified student category. As students reach English proficiency, they leave the category and any changes in their performance after reclassification are not represented in average EL-classified student performance levels. Using an ever-EL framework can account for the changing category by retaining reclassified students in the analytic sample (Thompson et al., 2022). To address these concerns, I run the same event study analyses for academic outcomes for a set of different populations. For one check, I include outcomes for any student who was EL-classified from 2014/15-2018/19, regardless of whether they reclassified during the time period. This would include students who were recently reclassified prior to the intervention, and as such may benefit from the additional services and supports that come through HB 3499. As another check, I only include current EL-

classified students and those who were EL-classified or reclassified during the HB 3499 intervention period. Given that language instruction programs are not required to be provided or reported for reclassified students, I am unable to conduct the same sensitivity analyses for bilingual services.

Research Question 1: Supplemental Analysis

As a secondary analytic approach, I integrated a regression discontinuity (RD) approach with the main difference-in-differences specification to estimate a local average treatment effect for a subset of districts identified as more comparable groups through an RD approach. The intuition behind an RD design is that, when a given value on known continuous index is used to assign treatment, the assignment mechanism can be used to create two comparison groups—those who just barely missed treatment because of their index value, and those who were just barely treated because of their index value. The assumption is that falling just above or below the threshold on this 'forcing' variable is quasi-random and the treatment effect can be estimated almost as in a random experiment within a bandwidth around that threshold.

In this context, identification for treatment was guided by the combined needs and outcomes indices. Theorizing that effects may be different at each threshold, I conducted a frontier RD, which subsets the data to conduct an RD at each threshold, accounting for meeting the other threshold criteria (Reardon et al., 2012). The first RD is conducted at the outcomes threshold among high-need districts, with the data reduced to districts above the 47^{th} percentile in need, and the forcing variable being districts' outcomes index percentile point value. The second RD is conducted at the needs threshold among low-outcome districts, with the data reduced to districts below the 53rd percentile in outcomes, and the forcing variable being districts' needs index value. To utilize the panel nature of the data and threshold assignment, I integrated RD sample selection methods with difference-in-differences models to run the same models described above, but for two smaller, theoretically more comparable sets of districts, one at each identification threshold (needs and outcomes).

I limit my discussion of the methods and findings for this analysis in the main text, although they are discussed in more detail in Appendix B. I do so given concerns about the validity of the method in this context, described next. First, two thresholds are used to assign treatment (outcomes and needs), but only one threshold predicts treatment well (outcomes; see Figure B1). There is substantial deviation from the threshold assignment policy at the needs threshold. As such, a key RD assumption is violated—that the threshold predict treatment—for the needs threshold. The optimal bandwidth of data on the forcing variable, calculated while accounting for the clustered nature of the data, was calculated to be around 7.5 percentage points on either side of the outcomes threshold, which included a very small subset of districts (Calonico et al., 2020). Even manually widening the bandwidth to 10 percentage points, the sample sizes remained very small. At the outcomes threshold, the sample is just 11 districts (treatment and control combined), and at the needs threshold the sample is 13 districts (treatment and control combined). While the student populations are large, treatment was assigned at the district level and the clustering of standard errors and relatively few time points render the analysis under-powered. There were also significant differences in the treated and control group populations at the baseline year (Tables B1 and B2) and discontinuities in select student characteristics at both the outcomes and needs threshold (see Figure B6). This is evidence to suggest the necessary assumption in RD of continuity of pretreatment covariates is not met. Together, these checks provide evidence that assignment was not as good as random at the threshold conditional out the districts' needs and outcomes' index values. While I conduct the analyses and present the results from a difference-in-differences specification that draws on the RD subsamples in Appendix B, I only make a brief note of how the findings align or diverge from the main results in the Results section.

Research Question 2

Given that HB 3499 treatment varied in intensity, I also explored if the overall average effect of HB 3499 descriptively differed for Transformation and Target districts. I first ran a set of separate

models for each type of treated district (Target or Transformation). In these models, I only retain non-identified districts in the analysis for the comparison group. I also explored the same question, but instead used a multi-value treatment indicator rather than a binary indicator for the treatment value, running the same models above, but with a treatment indicator that differed for non-identified districts (0), Target districts (1), and Transformation (2) districts.

Results

Research Question 1

RQ1: EL Expenditures per EL-Classified Student

Overall, there was a positive effect of HB 3499 identification on EL expenditures per EL-classified student, although not in the first year of treatment. As shown in Figure 2.2, Panel I, and reported in Table 2.3, there was not a significant increase in EL expenditures per EL-classified students among HB 3499-identified districts in the first year of identification. There was a significant increase in the second and third years following identification. In the first year of identification there was an estimated increase of $421.0 [95% CI, -7.30; 849.3] in EL-specific expenditures per EL-classified students that was not precise enough to reject the null. In the second and third years after identification the estimated increase was larger and significantly different from zero, $911.0 in year two and $853.4 in year three, although confidence intervals spanned from an estimated increase of $125.8 to $1,696.2 in EL spending per EL-classified student in year two and $53.2 to $1,653.7 in year three. The overall estimate from the preferred difference-in-differences specification (Table 2.3, Column IV) aligned with the event study estimates, a significant estimated increase in EL expenditures per EL-classified student of $695.4. In comparison with the descriptive estimates, this is larger than the average per-student allocation for Transformation districts, but smaller than for Target districts.

Table 2.3 Estimated Effect of HB 3499 Identification on District EL Expenditures per EL-Classified Student

	I		II		III		IV	
	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI
HB 3499*2014	-477.5*	[-935.2; -19.8]	255.4	[-214.0; 724.7]	81.1	[-278.1; 440.2]		
HB 3499*2015	-255.7	[-675.2; 163.8]	82.6	[-392.1; 557.4]	1.91	[-298.4; 302.2]		
HB 3499*2017	554.2*	[91.6; 1,016.8]	495.0	[-42.8; 1,032.8]	421.0	[-7.30; 849.3]		
HB 3499*2018	1,633.1***	[957.4; 2,308.9]	1,030.4*	[225.6; 1,835.3]	911.0*	[125.8; 1,696.2]		
HB 3499*2019	1,480.0***	[758.9; 2,201.1]	953.1*	[128.4; 1,777.8]	853.4*	[53.2; 1,653.7]		
HB 3499							695.4*	[101.5; 1,289.2]
Fixed Effects								
Year	No		Yes		Yes		Yes	
District	No		No		Yes		Yes	
Obs.	577		577		577		577	
R2	0.183		0.229		0.683		0.679	
Cond. R2			0.104		0.052		0.039	

Note. *p<.05, **p<.01, ***p<.001. FE=Fixed effects. Est=Estimate, CI=Confidence interval. Years are spring of a given academic year, and 2016 was the year prior to identification. Model I includes district-level covariates, Model II has year fixed-effects added in, and Model III has district-level covariates along with year and district fixed effects. Model IV is a difference-in-difference specification with district covariates, year and district fixed effects.

Ratio of ESOL-Endorsed Teachers to EL-Classified Students

HB 3499 identification did not lead to a significant change in the ratio of ESOL-endorsed teachers to EL-classified students in the panel. As plotted in Figure 2.2, Panel II and reported in Table 2.4, post-identification estimates were small and negative in the preferred specification (Column III), ranging from -0.017 [0.054; 0.019] in the first year post-identification to -0.005 [-0.061; 0.051] in the third year of identification. The difference-in-difference specification (Table 2.4, Column IV) yields a small, null estimate, 0.0002 [-0.046; 0.046]. While the effect is small, the confidence interval ranges from an impact of about one additional or one fewer ESOL-endorsed teacher for every 20 EL-classified students.

Table 2.4 Estimated Effect of HB 3499 Identification on the Ratio of ESOL Endorsed Teachers to EL-classified Students

	I		II		III		IV	
	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI
HB 3499*2014	-0.076***	[-0.119; -0.033]	-0.021	[-0.053; 0.011]	-0.037	[-0.074; 0.001]		
HB 3499*2015	-0.040.	[-0.081; 0.0006]	-0.01	[-0.039; 0.019]	-0.018	[-0.044; 0.007]		
HB 3499*2017	0.004	[-0.039; 0.048]	-0.021	[-0.082; 0.040]	-0.017	[-0.054; 0.019]		
HB 3499*2018	0.01	[-0.034; 0.055]	-0.014	[-0.081; 0.052]	-0.014	[-0.060; 0.032]		
HB 3499*2019	0.023	[-0.028; 0.074]	-0.01	[-0.088; 0.068]	-0.005	[-0.061; 0.051]		
HB 3499							0.0002	[-0.046; 0.046]
Fixed Effects								
Year	No		Yes		Yes		Yes	
District	No		No		Yes		Yes	
Obs.	582		582		582		582	
R2	0.216		0.239		0.814		0.773	
Cond. R2			0.207		0.259		0.092	

Note. *p<.05, **p<.01, ***p<.001. FE=Fixed effects. Est=Estimate, CI=Confidence interval. Years are spring of a given academic year, and 2016 was the year prior to identification. Model I includes district-level covariates, Model II has year fixed-effects added in, and Model III has district-level covariates along with year and district fixed effects. Model IV is a difference-in-difference specification with district covariates, year fixed effects, and district fixed effects.

Probability of Receiving Bilingual Education Services

There was no significant effect of being in an HB 3499-identified district after identification on the probability that an EL-classified student received bilingual education services. As plotted in Figure 2.2, Panel III and reported in Table 2.5, event study estimates were small and not significantly different from zero across all three post-identification years. Confidence intervals were relatively stable across the three years post-treatment. We can rule out effects larger than a decline of about four percentage points in the probability of receiving bilingual services, and an increase of about eight percentage points. Pooling effects in a difference-in-differences approach, as reported in Table 2.5, Column V, yielded an overall average effect that aligns with event study estimates, an increase of 3.2 percentage points [-0.017; 0.082] that was not significantly different from zero.

Table 2.5 Estimated Effect of HB 3499 Identification on EL-Classified Students' Probability of Receiving Bilingual Services

	I		II		III		IV		V	
	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI
HB 3499* 2014	-0.144*	[-0.272; -0.016]	-0.158*	[-0.302; -0.013]	-0.027	[-0.101; 0.047]	-0.016	[-0.087; 0.055]		
HB 3499* 2015	-0.131*	[-0.254; -0.009]	-0.151*	[-0.285; -0.017]	-0.023	[-0.047; 0.001]	-0.01	[-0.030; 0.010]		
HB 3499* 2017	-0.108	[-0.249; 0.033]	-0.105	[-0.245; 0.035]	0.012	[-0.037; 0.061]	0.003	[-0.047; 0.052]		
HB 3499* 2018	-0.097	[-0.224; 0.029]	-0.099	[-0.234; 0.035]	0.018	[-0.030; 0.066]	0.002	[-0.050; 0.055]		
HB 3499* 2019	-0.091	[-0.220; 0.037]	-0.100	[-0.247; 0.047]	0.019	[-0.035; 0.074]	-0.002	[-0.066; 0.063]		
HB 3499									0.032	[-0.017; 0.082]
Fixed Effects										
Grade	Yes		Yes		Yes		Yes		Yes	
Year	No		Yes		Yes		Yes		Yes	
District	No		No		Yes		Yes		Yes	
School	No		No		No		Yes		No	
Obs.	291,860		291,860		291,860		291,860		291,860	
R2	0.197		0.198		0.436		0.587		0.436	
Cond. R2	0.168		0.169		0.083		0.018		0.083	

Note. *p<.05, **p<.01, ***p<.001. Est=Estimate, CI=Confidence interval. Years are spring of a given academic year, and 2016 was the year prior to identification. Model I includes student-, school, and district-level covariates with grade fixed effects, Model II has year fixed-effects added in, Model III has district fixed effects added in, and Model V school fixed effects. Model IV is a difference-in-difference specification with the listed covariates, grade, year, and district fixed effects.

Academic Outcomes

There was no significant effect of being in an HB 3499-identified district after identification on EL-classified students' academic outcomes. Presented in Tables 2.6 and 2.7, confidence intervals rule out effects larger than an increase and decrease of about 0.05 SD. Plotted in Figure 2.2, Panel III, the estimated effect of identification on ELA scores in the first year after identification was -0.006 SD [-0.050; 0.039], the second year 0.010 SD [-0.036; 0.055] and the third year 0.003 SD [-0.049; 0.056]. Math estimates were small and null, -0.002 SD [-0.043; 0.038] in year one, -0.006 SD [-0.060; 0.049] in year two and -0.007 SD [-0.055; 0.040] in year three. The overall estimated effect on ELA scores (Table 2.6, Column IV) and math scores (Table 2.7, Column IV) were also null.

Table 2.6 Estimated Effect of HB 3499 Identification on EL-Classified Students' ELA Outcomes

	ELA											
	I		II		III		IV		V		VI	
	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI
HB 3499* 2015	-0.104***	[-0.164; -0.044]	-0.123***	[-0.188; -0.057]	-0.021	[-0.050; 0.009]	-0.005	[-0.040; 0.029]				
HB 3499* 2017	-0.013	[-0.082; 0.055]	-0.103*	[-0.183; -0.024]	-0.006	[-0.050; 0.039]	0.002	[-0.044; 0.048]				
HB 3499* 2018	-0.128***	[-0.188; -0.067]	-0.088**	[-0.151; -0.024]	0.010	[-0.036; 0.055]	0.021	[-0.022; 0.064]				
HB 3499* 2019	-0.129***	[-0.193; -0.065]	-0.089*	[-0.157; -0.021]	0.003	[-0.049; 0.056]	0.022	[-0.025; 0.070]				
HB 3499									0.016	[-0.022; 0.055]	0.018	[-0.012; 0.049]
HB 3499*EL											-0.046	[-0.133; 0.042]
Fixed Effects												
Grade	Yes		Yes		Yes		Yes		Yes		Yes	
Year	No		Yes		Yes		Yes		Yes		Yes	
District	No		No		Yes		Yes		Yes		Yes	
School	No		No		No		Yes		No		No	
Observations	114,132		114,132		114,132		114,132		114,132		1,024,583	
R2	0.149		0.152		0.170		0.196		0.196		0.342	
Cond. R2	0.101		0.099		0.093		0.085		0.085		0.298	

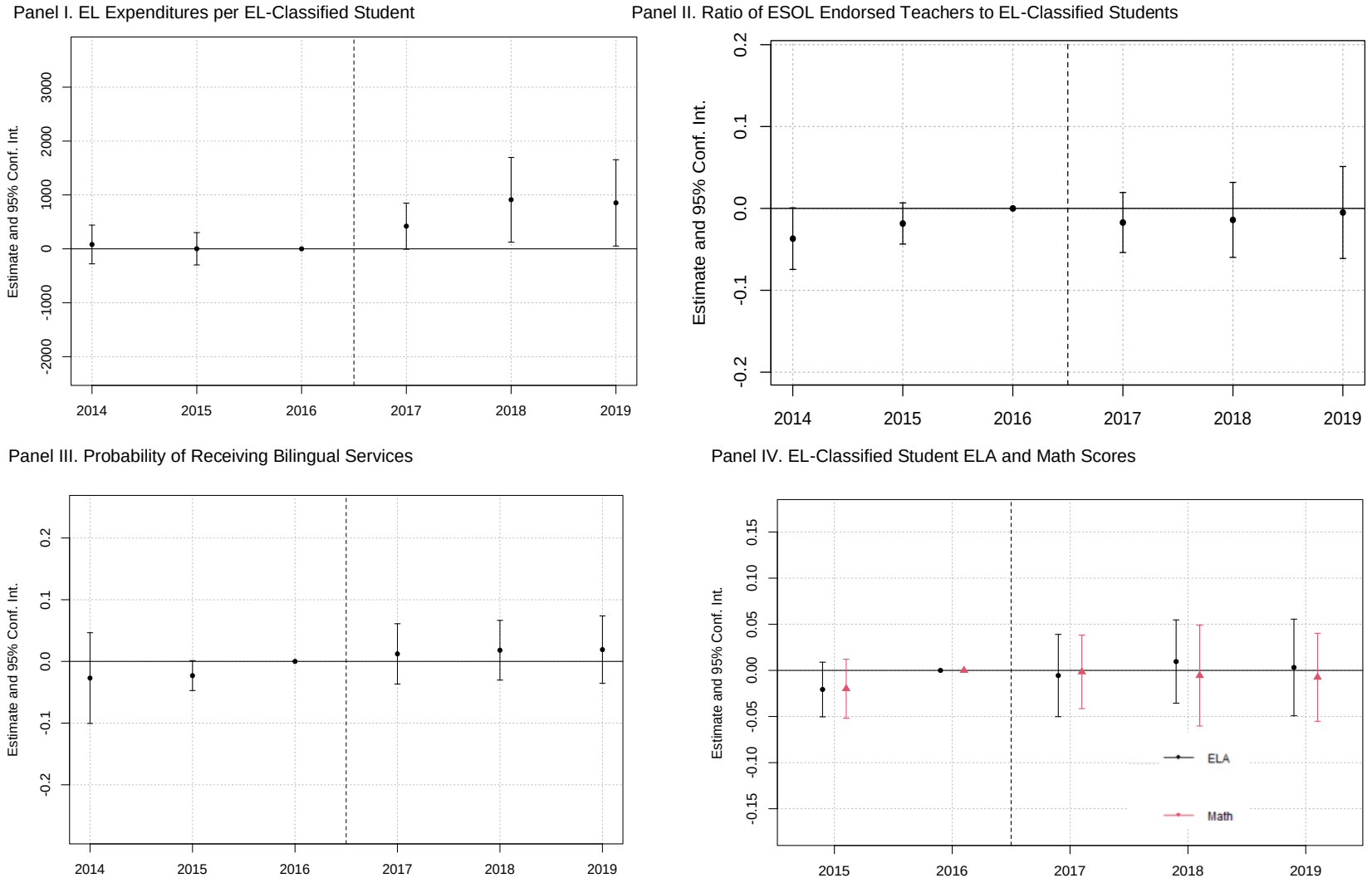
Note. *p<.05, **p<.01, ***p<.001. Est=Estimate, CI=Confidence interval. Years are spring of a given academic year, and 2016 was the year prior to identification. Model I includes student, school, and district-level covariates, as well as grade fixed effects. Model II has year fixed effects added in, Model III adds in district fixed effects, and Model IV school fixed effects. Model V is a difference-in-difference specification with student, school, and district covariates, grade, year, and district fixed effects. Model VI is a triple difference model, but the same specification as Model III.

Table 2.7 Estimated Effect of HB 3499 Identification on EL-Classified Students' Math Outcomes

	Math											
	I		II		III		IV		V		VI	
	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI
HB 3499* 2015	-0.127**	[-0.205; -0.048]	-0.128**	[-0.219; -0.038]	-0.02	[-0.052; 0.012]	-0.003	[-0.031; 0.026]				
HB 3499* 2017	-0.026	[-0.112; 0.059]	-0.108*	[-0.201; -0.014]	-0.002	[-0.042; 0.038]	0.007	[-0.033; 0.048]				
HB 3499* 2018	-0.131**	[-0.213; -0.050]	-0.112*	[-0.198; -0.027]	-0.006	[-0.060; 0.049]	0.013	[-0.039; 0.065]				
HB 3499* 2019	-0.128**	[-0.210; -0.046]	-0.111**	[-0.193; -0.029]	-0.007	[-0.055; 0.040]	0.017	[-0.029; 0.063]				
HB 3499									0.013	[-0.027; 0.054]	0.013	[-0.021; 0.047]
HB 3499*EL											-0.030	[-0.091; 0.031]
Fixed Effects												
Grade	Yes		Yes		Yes		Yes		Yes		Yes	
Year	No		Yes		Yes		Yes		Yes		Yes	
District	No		No		Yes		Yes		Yes		Yes	
School	No		No		No		Yes		No		No	
Observations	115,941		115,941		115,941		115,941		115,941		1,022,434	
R2	0.163		0.165		0.184		0.224		0.224		0.326	
Cond. R2	0.127		0.126		0.119		0.10		0.10		0.276	

Note. *p<.05, **p<.01, ***p<.001. Est=Estimate, CI=Confidence interval. Years are spring of a given academic year, and 2016 was the year prior to identification. Model I includes student, school, and district-level covariates, as well as grade fixed effects. Model II has year fixed effects added in, Model III adds in district fixed effects, and Model IV school fixed effects. Model V is a difference-in-difference specification with student, school, and district covariates, grade, year, and district fixed effects. Model VI is a triple difference model, but the same specification as Model III.

Figure 2.2. Event Study Estimates of the Effect of HB 3499 Identification on District and EL-Classified Student Outcomes



Note. All models include district covariates and district and year fixed effects, while student-level models also include student covariates, school covariates, and grade fixed effects.

Robustness Checks

Parallel Trends. A key assumption for the approaches as modeled above is that treated and untreated units were on parallel trends prior to identification (Roth, 2022). As seen across Panels I through IV in Figure 2.2, overall, there were no point estimates pre-identification significantly different from zero at the .05 level. However, for both teacher endorsements and bilingual services, there were point estimates in the pre-treatment years significant at the .01 level. This means that, when conditioned on the set of fixed effects and covariates included in the main model, there is suggestive evidence that there may have been differential trends between HB 3499 districts and non-HB 3499 districts. When paired with raw data that show different trends in the average linear trend prior to treatment (see Figures A3 and A4), I am concerned that districts were on different trajectories for both teacher endorsements and bilingual services estimates prior to intervention, which could bias the estimates of effects of HB 3499 in treatment districts. This concern is further magnified by concerns that, especially in contexts with few pre-treatment coefficients, there may not be sufficient power to detect true pre-trends (Roth, 2019).

To examine the potential magnitude of bias and how that may impact the point estimates found I used the HonestDiD (Roth, 2019) package to model how violations of the parallel trend assumption of varying magnitudes may shift the point estimates in the years after identification for both bilingual services and teacher endorsements. There are two ways to do so, with one being bounding the relative magnitude of post-treatment trends in relation to those observed in the pre-treatment years, where a value of 1 "bounds the worst-case post-treatment difference in trends by the equivalent maximum in the pre-treatment period" (Rambachan & Roth, 2022, p. 12). The other way is by imposing a smoothness restriction, where one plots the point estimates by varying degrees of how much the slope changes between consecutive periods in the pre-treatment period. In this case, a

value of 0 means that the slope changes are perfectly linear. I chose to implement the relative magnitude check for this main analysis.

In Figure A7, I plot how the point estimates for teacher endorsements (Panel I) and bilingual services (Panel II) would change in each year post-identification under pre-trend violations of varying relative magnitudes ($\bar{M}$). For each yearly point estimate, the original point estimate is plotted, followed by a series of point estimates and confidence intervals as they would be estimated under increasing severity in the violations of pre-trends, from 0 to 1 in increments of 0.1 (Rambachan & Roth, 2019) As seen, the potential violations decrease point estimate precision across outcomes. This does not substantively change the interpretation that there was no overall effect of identification. However, these robustness checks provide important caveats for the results. The main estimates, in the presence of what may be differential pre-trends, may not be precise enough to rule out much larger negative or positive effects of HB 3499 on both teacher endorsements and bilingual services.

Alternative Samples. As an additional robustness check, I estimated the effect of HB 3499 on academic outcomes using two alternative samples—students who were ever-EL classified during the panel, and current EL-classified students in addition to any students who were reclassified to fluent English proficient post-identification. Figure A8 presents point estimates from the main event study specification using these two alternative samples. For both samples, estimates are null and estimates from difference-in-difference specifications small and not statistically distinguishable from zero.

Simultaneous Policy Changes and Placebo Checks. To probe the question of whether simultaneous policy shifts may have impacted the academic results, I also fit a triple-difference specification for the academic outcomes. Results for the coefficient that interacts treatment with EL classification in the post-identification time period are reported in the main tables for academic outcomes (Tables 2.5 and 2.6, Column VI). For both ELA and math, the coefficient is null, presenting no evidence of unobserved factors impacting the main results.

I also fit the main event study specification with a set of placebo outcomes for both the expenditure and academic achievement outcomes. These should yield null results—any significant results may introduce concerns that there were simultaneous policy changes or unobserved factors influencing the outcomes. I fit the main event study specification with non-EL expenditures as the outcome of interest in one model and never-EL students' academic outcomes in another. As seen in Figure A9, both placebo models yielded expected null effects, providing no evidence that simultaneous policies or unobserved factors impacted the results.

Alternative Analyses: Regression Discontinuity Evidence on Effects at the Thresholds. I also estimated the same difference-in-differences specifications using two subsamples identified as theoretically more comparable pre-intervention groups through a frontier RD selection process at each threshold (Reardon et al., 2012). The approach and results are discussed in more detail in Appendix B. One subsample was comprised of districts below the outcomes threshold and within ten percentile points of the needs threshold, a sample of 13 districts, with seven identified for HB 3499. The other subsample was comprised of districts above the needs threshold and within 10 percentile points of the outcomes threshold, a sample of eleven school districts, with five identified for HB 3499. There were multiple ways in which both RD and difference-in-difference assumptions were violated. These violations, in conjunction with the very small number of districts, mean that I am not confident in the validity or reliability of the estimates.

With these concerns in mind, estimates are reported in Table B3 for all included outcomes. For high-need districts at the outcomes threshold, there were no significant effects. While this aligns with the main estimates across the majority outcomes, the findings differ from the main findings regarding EL expenditures per EL-classified students, where there was a significant effect. For low-outcome districts at the needs threshold, there was a significant positive effect of identification on EL-expenditures per EL-classified student, as well as the ratio of ESOL-endorsed teachers to EL-

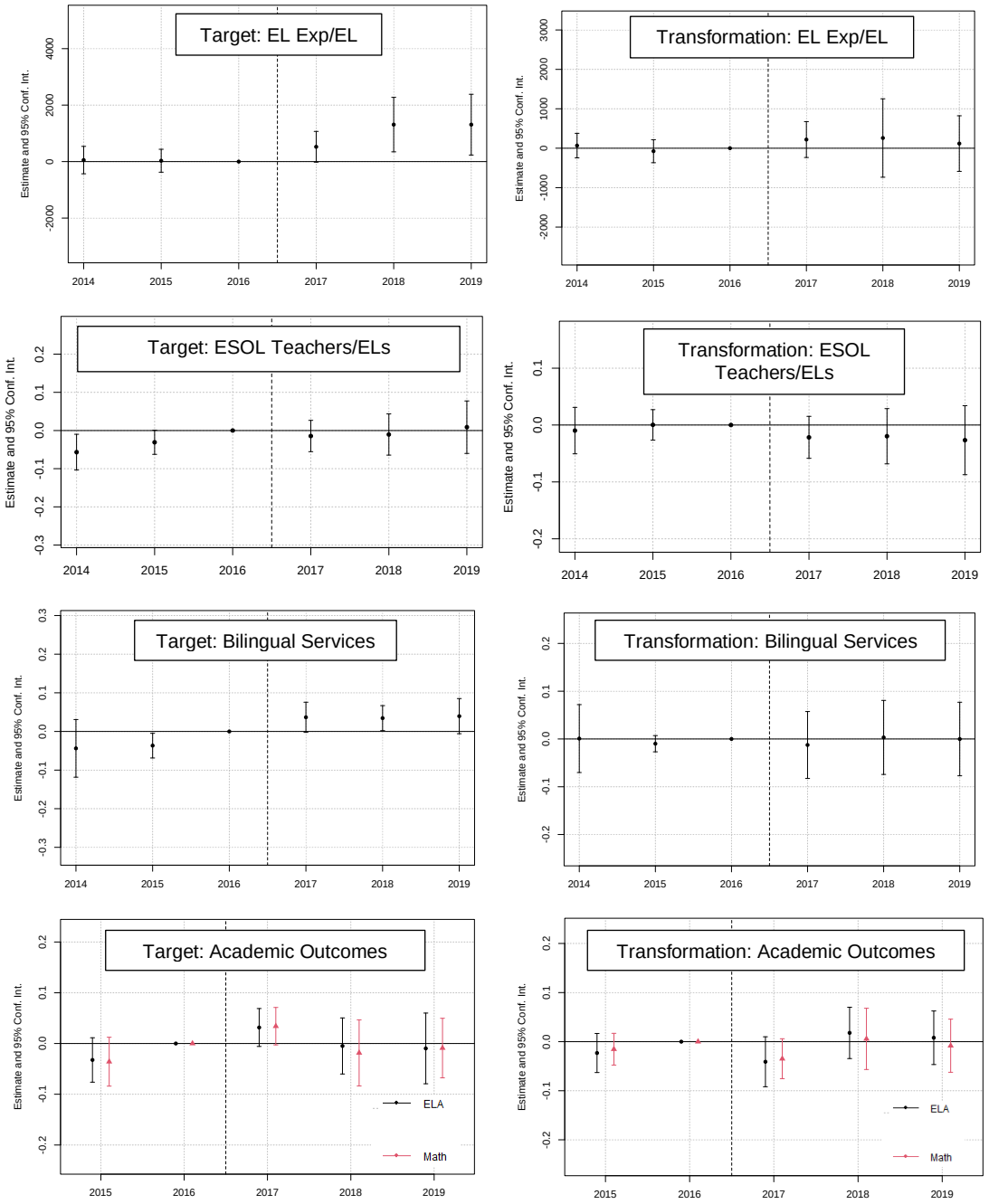
classified students. These noted departures from the main results may be because, in addition to being a very small sample, both the treated and untreated districts are different in composition than the main analytic group.[20] Additionally, trends in the raw data for the subset of low-outcome districts (Figure B4) provide descriptive evidence that there was a significant decline in the teacher ratio among non-HB 3499 districts in the bandwidth, perhaps driving the observed effect among HB 3499 districts. Again, these results are drawn from a small subset of districts and the analyses are provided with strong caveats around the ability to draw meaningful conclusions from the small number of districts amidst the potential violations in both RD and difference-in-difference assumptions, as modeled.

Research Question 2

HB 3499 intervention differed for districts—some districts received a more intensive intervention (Transformation districts) than others (Target districts). To examine potential heterogeneity in the effects of HB 3499 by treatment type, I conducted two additional descriptive analyses. First, I split the samples and estimated the treatment impact for Transformation and Target district separately in comparison with non-identified districts. I then estimated the effect of HB 3499 using a multi-value indicator instead of a binary indicator in the main difference-in-difference specification. Figure 2.3 presents event study point estimates from the separate analyses, and Table 2.8 presents estimates from the difference-in-difference specification. The estimated effect differed in relation to the overall effect estimated in two ways.

[20] Specifically, within the high-need bandwidth, HB 3499 districts' EL-classified student population was comprised of more Asian and Black students, and fewer Hispanic/Latinx than the main analytic sample. The districts also had a smaller proportion of their population EL-classified, and a larger proportion FRPL eligible. Within the low-outcome bandwidth, HB 3499 districts had EL-classified students who were more likely to be Hispanic/Latinx than the overall sample, while the district overall had a lower proportion of the students FRPL eligible and EL-classified.

Figure 2.3. Effect of HB 3499 Identification on EL-Classified Students' Outcomes in Target and Transformation Districts



Note. All models include district covariates and district and year fixed effects, while student-level models also include student covariates, school covariates, and grade fixed effects.

The first instance in which the treatment effect differed from the main results was for EL expenditures per EL-classified students. As seen in Figure 2.3, the estimated yearly effects of identification on Target districts' EL expenditures per EL-classified student aligned with the main estimates, with a null result in the first year and a positive estimate in years two and three. However, there was no estimated increase in spending among Transformation districts, with imprecise estimates. Difference-in-differences results align with the event study estimates (Table 2.8), with an estimated positive effect of identification on EL expenditures per EL-classified student among Target districts, but not Transformation districts.

The second instance was for bilingual services. While the overall estimates found no impact of HB 3499 identification on EL-classified students' probability of receiving bilingual services, heterogeneity analyses reveal differences between Target and Transformation districts. There was a significant estimated increase in the probability of receiving bilingual services for EL-classified students in Target districts in year two, while year one and three estimates were positive, but not statistically significant (Figure 2.3). When estimated in a difference-in-differences specification (Table 2.8), the estimated effect was positive and significantly different from zero—an estimated increase in the probability of receiving bilingual services of 7.3 [0.028; 0.118] percentage points in Target districts. Effects were null for EL-classified students in Transformation districts in both the event study and difference-in-differences specification.

Table 2.8. Estimated Effect of HB 3499 Identification, by Target or Transformation Status

	Expenditures		ESOL Endorsed Teachers		Bilingual Services		ELA		Math	
	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI
Post*HB3499Target	994.3*	[194.4; 1,794.1	0.022	[-0.029; 0.073]	0.073**	[0.028; 0.118]	0.020	[-0.034; 0.074]	0.019	[-0.032; 0.070]
Post*HB3499Transform	202.5	[-369.4; 774.3]	-0.019	[-0.061; 0.022]	0.0101	[-0.070; 0.090	-0.004	[-0.042; 0.050]	-0.009	[-0.052; 0.035]
Fixed Effects										
Grade	No		No		Yes		Yes		Yes	
Year	Yes		Yes		Yes		Yes		Yes	
District	Yes		Yes		Yes		Yes		Yes	
School	No		No		No		No		No	
Obs.	577		582		291,860		114,132		115,941	
R2	0.684		0.815		0.512		0.166		0.1842	
Cond. R2	0.054		0.260		0.091		0.0934		0.119	

Note. *p<.05, **p<.01, ***p<.001. FE=Fixed effects. Est=Estimate, CI=Confidence interval. Years are spring of a given academic year, and 2016 was the year prior to identification.

Robustness Checks

Parallel Trends. The heterogeneity results, importantly, rely on the same assumption that districts were on parallel trends prior to identification. As seen in Figures A3 and A4, there were instances where Target districts were on different trends prior to identification. Specifically, the raw trend data suggest that the ratio of ESOL endorsed teachers to EL-classified students increased at a faster rate, on average, in Target districts than in both Transformation and non-HB 3499 districts in the years prior to identification. Additionally, while Transformation and non-HB 3499 districts saw slight declines in the proportion of students receiving bilingual services prior to identification, Target districts saw an increase.

Thus, as with the main estimates, I conducted the same pre-trend sensitivity analyses for Target districts, plotting each event study point estimate for teacher endorsements and bilingual services to see at which degree of severity the results substantively change in Figure A10. As with the

overall estimates, teacher endorsement (Panel I) estimates remain null, but with increasingly large confidence intervals. For bilingual services (Panel II), the statistically significant result "breaks down" at a value of 0.1. As interpreted, this means that, in order to claim that the significant effect remains robust to differential trends, post-identification trends must not be more than 0.1 times (10%) larger than the maximum pre-trend violation (Rambachan & Roth, 2022). This is a small value and indicates sensitivity. I interpret this to mean that the result is sensitive to observed pre-trends and should be interpreted with caution.

Discussion

HB 3499 is an important policy intervention to support struggling districts in Oregon in better serving their EL-classified students. In this paper, I began by estimating the effect of identification on districts' EL expenditures per EL-classified student, an important step in trying to understand if the intervention had enough funding to see meaningful shifts in spending. Overall, I found that identification led to a significant increase in EL-specific expenditures. Through heterogeneity analyses I found that the significant increase may be concentrated among Target districts. I found no significant effect on the ratio of ESOL endorsed teachers to EL-classified students, nor the probability an EL-classified student received bilingual services, although heterogeneity analyses revealed suggestive evidence that students in Target districts were more likely to receive bilingual services after identification. There was no significant effect of HB 3499 identification on EL-classified students' academic outcomes, overall or in any alternative estimation. In this discussion section, I first note limitations, then contextualize my work within the broader EL and education policy research base before discussing implications.

Limitations

This study comes with important limitations. First, the assignment of treatment at the district level constrained power. As such, despite having large student sample sizes, I may have been limited

in my ability to detect a true significant effect. Additionally, there are relatively few pre-treatment periods observed, which inhibits my ability to identify pre-identification trends. As another caveat, there are many more important outcomes that can, and should be examined that may better capture improvements in districts' EL services and instruction, including linguistic outcomes[21], student perceptions of school belonging, and more mechanistic outcomes, such as changes in teacher skill or curricula used. Such changes may support the outcomes looked at in this study, but the impacts may not be detectable within the relatively short window of this study. Further, there was an additional identification year—2019/20, that I was unable to examine. Given the long timelines of school improvement (Sun et al., 2021), a longer horizon may have revealed different findings.

Contextualizing English Learner Expenditure Findings

Overall, there was an increase in EL-specific expenditures per EL-classified student in HB 3499 districts. While estimates were significant, they were also noisy, suggesting that there was wide variability in the increase of EL-expenditures per EL-classified students among treated districts in the three years observed post-identification. For some districts, the additional funds may be translated into a significant change per EL-classified students, while for others, especially those with larger EL-classified student populations, the amount may not have represented a significant per pupil increase. A descriptive look at average district expenditures by different areas, as discussed in the Data section and plotted in Figure A2, revealed that spending on EL-specific salary and payroll costs represented the vast majority of average district EL expenditures before and after identification, with increases over time. The next most prevalent areas are benefits and administrator costs. Additionally on average after identification, HB 3499 districts spent more on EL-specific professional, instructional, and technical services, supplies and technology, classified positions, textbooks, and other expenditures.

[21] While important, I chose not to look at growth in English language proficiency to center the issue of academic opportunity, instruction, and outcomes rather than a focus on students' English language proficiency growth or timeline to reclassification.

Prior research suggests that the areas where HB 3499 districts invested in EL expenditures are important policy levers in education. For example, Holden (2016) found that additional textbook funds had a positive impact on elementary school academic outcomes. Additionally, teachers have an outsized impact on student academic achievement (Hanushek & Rivkin, 2012), and training focused on EL education can improve teachers' impacts on EL-classified students' academic growth (Master et al., 2016). However, increases in expenditures did not translate into improved academic outcomes relative to un-identified districts. As one potential explanation, additional funding amounts were relatively low—an additional $180,000 per year for Transformation districts, and $90,000 per year for Target districts. The total additional funds represent a fraction of districts' average EL expenditures on salary and payroll costs alone. Additionally, pervasive structural barriers shape EL-classified student experiences, including exclusion from rigorous course content (Dabach, 2014), tracking (Callahan, 2005; Umansky, 2016), racism (Rosa & Flores, 2017), and compounding socioeconomic and segregation inequities (Gándara & Hopkins, 2010). While funding is an important piece of strengthening EL services, pervasive structural barriers are critically important areas to address in the pursuit of more equitable outcomes for EL-classified students. Addressing these barriers will take significant structural and systemic shifts that likely require resources above and beyond the funding and support provided through HB 3499.

The finding that spending increases per EL-classified student appear to be concentrated among Target districts potentially reveals ways in which the policy was insufficient to create substantial change among districts facing the most acute need, Transformation districts. While Transformation districts received a larger overall amount, they tended to have, on average, larger EL student populations. In practice, this means that the actual per EL student amount awarded was actually less, on average, than for Target districts. This likely intersected with the compounding needs Transformation districts faced—having higher proportions of students experienced socioeconomic

disadvantage and higher concentration of EL-classified student need—and potentially was not enough in overall amount to translate into meaningful per-pupil shifts.

Contextualizing Teacher Endorsement Findings

Overall, there was no significant impact of HB 3499 identification on the ratio of ESOL-endorsed teachers to EL-classified students. One potential explanation for the null findings may be the relatively short timeframe of the policy. The ESOL endorsement process requires 18 instructional credits and a passing assessment score (ORS § 584-420-0360). It may be that three years was not a long enough time frame to see an increase among schools. Additionally, endorsements are just one way to "measure" teacher professional development. Teacher professional development efforts through HB 3499 may have been focused more on trainings and supports for a wider base of teachers, rather than focusing on more resource-intensive professional development supports such as endorsements for a smaller number of teachers. This sort of investment is more difficult to capture and is beyond the scope of this study but is important and warrants further research.

Contextualizing Bilingual Service Findings

Overall, I estimated null effects of being in an HB 3499 district on EL-classified students' likelihood of receiving bilingual services to support core content access. The rationale behind positioning bilingual services as an outcome stemmed from documented barriers to bilingual program implementation that could be addressed through HB 3499, such as cost and implementation supports (Steele et al. 2018), as well as concerns that heightened accountability environments can constrict bilingual education (Menken & Solorza, 2014). However, the absence of significant overall effects in the three-year window observed may not be surprising. One reason for overall null results may be the hurdles that come with creating new bilingual programs. This may include the hiring of qualified educators, sourcing materials, and working with the community to develop a program that meets the needs of students and families alike (DeNicolo, 2016). HB 3499 may not represent enough of an

intervention to see new programs be implemented, or existing programs expand substantially, especially within the time frame observed. Another reason for null results may have been the policy design. HB 3499 did not explicitly focus on bilingual services for EL education. Accountability metrics for identified districts focused on academic and linguistic growth, not program models. Bilingual education, while an effective and important support for EL-classified students, has been viewed by some as a risk in the face of accountability pressures given the English-only focus of standardized assessments (Menken & Solorza, 2014).

The potential increase among Target districts, however, suggests that there was a set of districts who leveraged their additional funds and support, in the face of increased accountability, to either create new bilingual programs or expand existing programs. While these are cautionary estimates, the mechanisms behind why there may have been changes in Target districts and not Transformation districts are worth exploring. Transformation districts tended to have higher needs index values and lower outcomes index values than Target districts. It could be that the barriers to implementing bilingual education programs, including cost and finding qualified bilingual educators, were more salient in these higher-need districts. In those contexts, funds and support may have been directed to more foundational supports, such as teacher training or instructional materials, rather than bilingual programs. Target districts, in contrast, may have been more well-positioned to expand or create bilingual education programs.

Contextualizing Academic Outcome Findings

No matter the specification, there was no significant increase in EL-classified students' ELA or math standardized scores, nor an increase in scores when including reclassified student outcomes. There are many reasons why some schools and districts have lower test score outcomes than others (Fullan, 1992). Many are structural and tied to issues of segregation and inequality (Darling-Hammond, 2007). While policies such as HB 3499 may represent important steps towards addressing disparities

in academic services and outcomes in schools, it may be that the scope of the intervention was not significant enough to see improvements in the outcomes examined, especially within a three-year window. Again, given research on how school improvement supported by additional funding and support may take longer time horizons to show improvement in outcomes such as academic performance (Sun et al., 2021), null academic results may not fully capture in-process improvements.

Implications

Null results for many outcomes, found on average for the full set of HB 3499 districts, may be disappointing for those interested in identifying a mechanism for supporting EL-classified students. While there is value in assessing these outcomes, these results should not be used to fully measure whether the policy itself "worked". Instead, these results can inform future iterations of policies to support EL-classified students. As such, an important implication is that ODE should invest in continual evaluation of HB 3499 and the policy impacts, with a focus on multiple evaluation methods, including qualitative work. While null findings are not, I argue, standalone evidence of policy failure, they are, I argue, evidence that many of the issues in EL education that constrict student outcomes may require larger shifts and structural changes than facilitated through HB 3499, especially for districts with acute needs.

As described earlier, at the end of the HB 3499 cohort intervention, only three of the forty districts did not meet one or both of the benchmarks set under the HB 3499 accountability framework and are subject to state directed EL funding. The majority (22) met the necessary benchmarks to be considered "reclassified", with no need for monitoring or direction of EL funding (ODE, 2022). By these measures, the vast majority of districts saw improvement on important measures of student access and outcomes. This measure of impact does not use comparison groups to make conclusions about the policy outcomes, focusing instead on individual benchmarks for districts. Additionally, those benchmarks are set for a host of other outcomes, including discipline, English language proficiency,

graduation and postsecondary access indicators. This is important to note, given that there are countless ways to evaluate policies, and different methods and outcomes likely would result in different conclusions. However, this current study provides important information for how observed improvements in student academic outcomes among HB 3499 districts, while they should be celebrated, were not observed when analyzed through a causal framework. Additionally, there were not significant shifts in some important structural process indicators, such as teacher endorsements and bilingual services, that may be important sources of support for EL-classified students.

HB 3499 is an ongoing policy—with a new cohort of districts beginning their window of support in 2022. It is important to contextualize the support that comes with identification within the current educational environment and provide guidance for districts on how to best utilize the limited, but important, additional funds in ways that are grounded in evidence. It may also be important to work with districts to identify the true costs in Oregon of providing adequate programs, staffing, and services to EL-classified students, and see whether the additional funds provided through HB 3499 come close to helping districts implement new programs, hire new staff, or invest in systemwide changes. It may also be important to work with districts to disentangle their perceptions of how the different policy elements impacted their own work—inquiring as to the impact of the funds, the technical assistance, and the accountability—to see if the different policy levers had different impacts on local practices. Such work can complement high-level policy evaluation to inform future work.

Conclusion

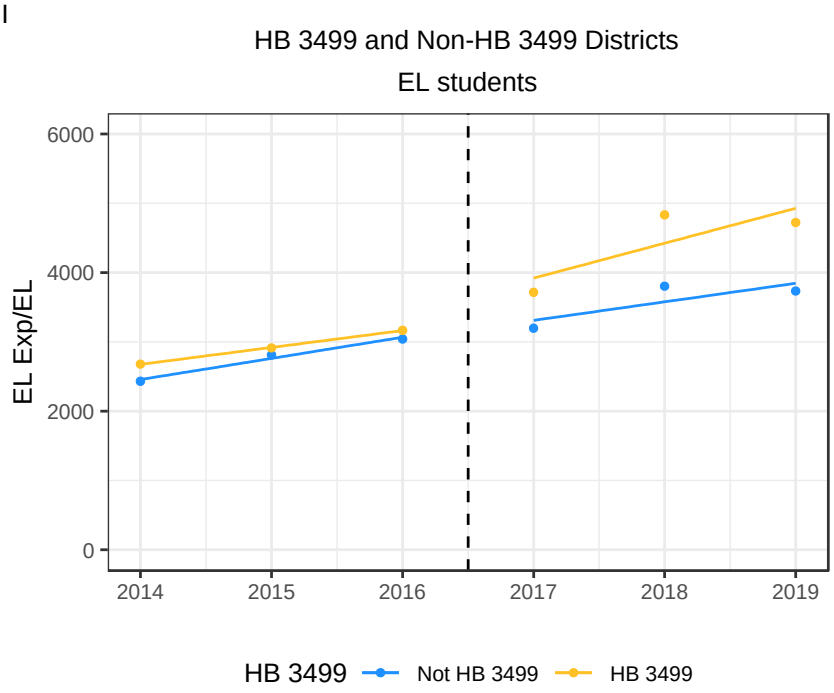
There are many ways through which the state education agency can shape EL education at the local level. HB 3499 is an innovative policy that holds promise as a lever for states interested in efficient ways to direct resources and support. While there was no overall estimated impact of HB 3499 on academic outcomes, there were significant increases in EL spending, and suggestive evidence that some districts expanded bilingual services, a critical support. It is important to continue to

evaluate the impacts of HB 3499 over time to see if these process indicators lead to improved academic outcomes on a longer time horizon. In the meantime, as HB 3499 continues with a new cohort, there is room for conversations around the amount of funding and support in juxtaposition with the costs that come with providing an adequate, appropriate education for EL-classified students and pervasive structural barriers experienced by many EL-classified students. Such conversations may encourage even stronger investments in districts' EL work, seeing HB 3499 as an important mechanism for support that can be strengthened as it continues.

APPENDIX A

Figure A1. EL-Specific Expenditures per EL-classified Students Over Time, by Whether Districts were Identified for HB 3499

Panel I



Panel II

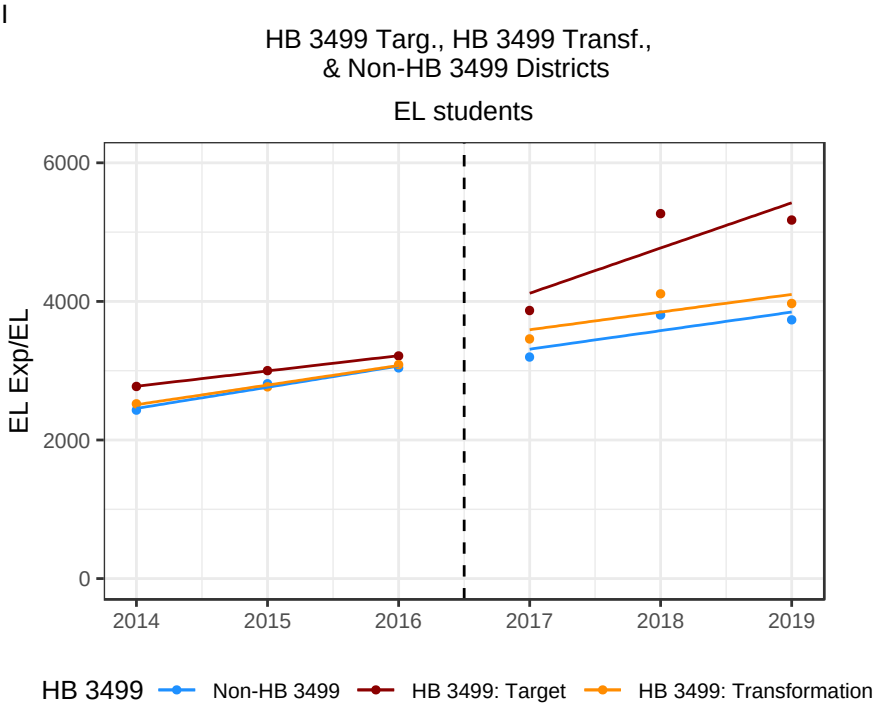


Figure A2. HB 3499 Average Total District EL-Specific Expenditures Over Time

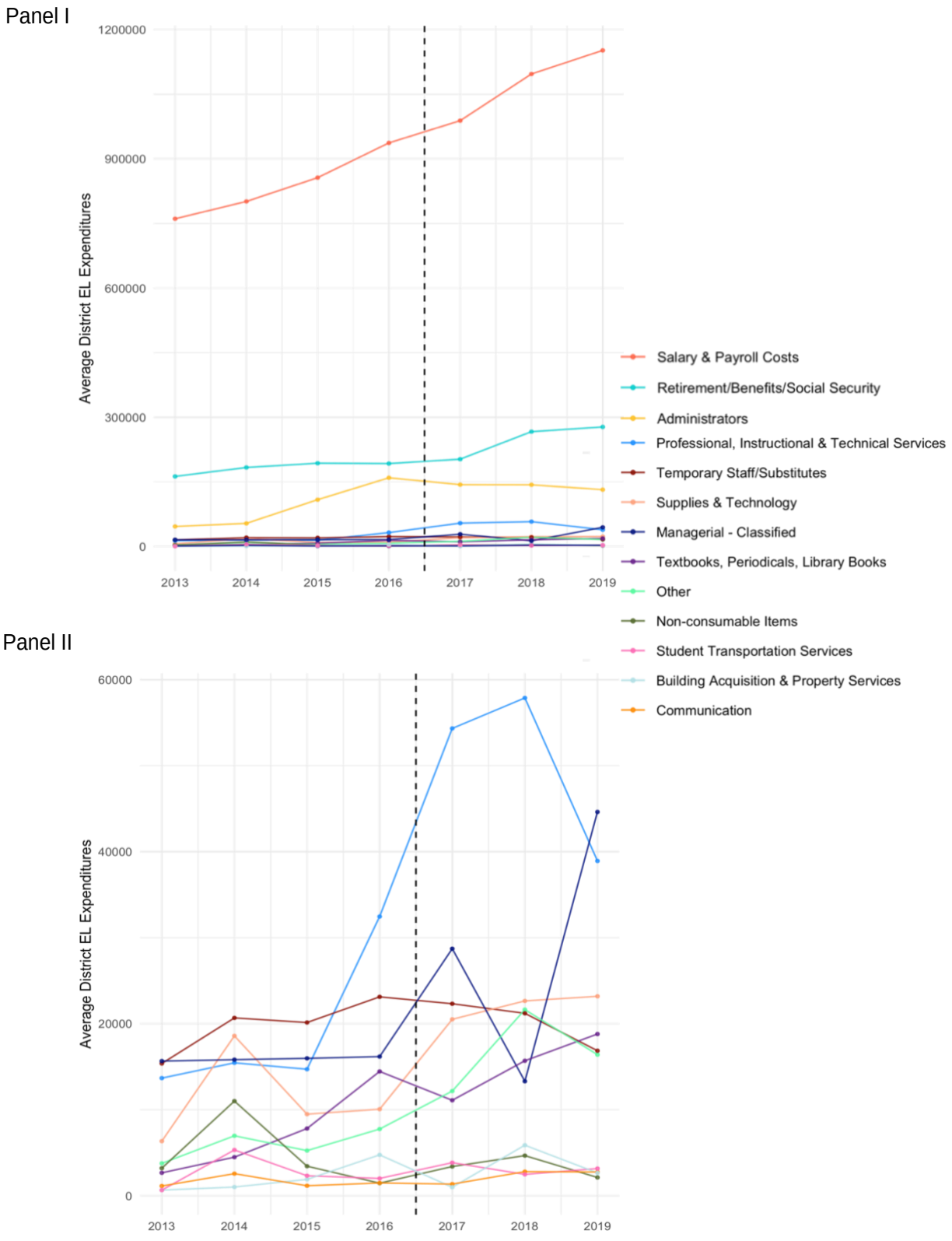
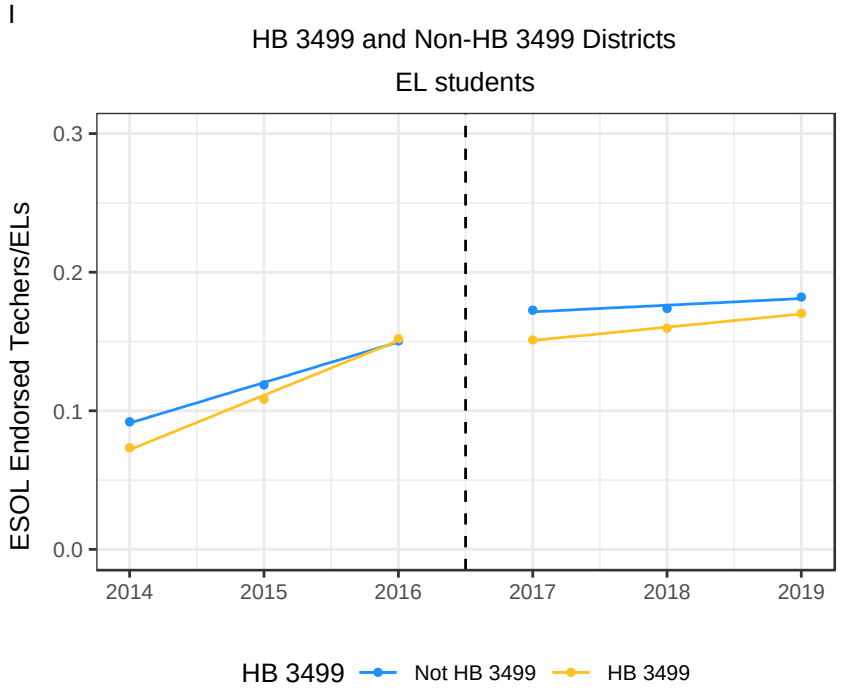


Figure A3. Ratio of ESOL-Endorsed Teachers to EL-Classified Students Over Time, by Whether Districts were Identified for HB 3499

Panel I



Panel II

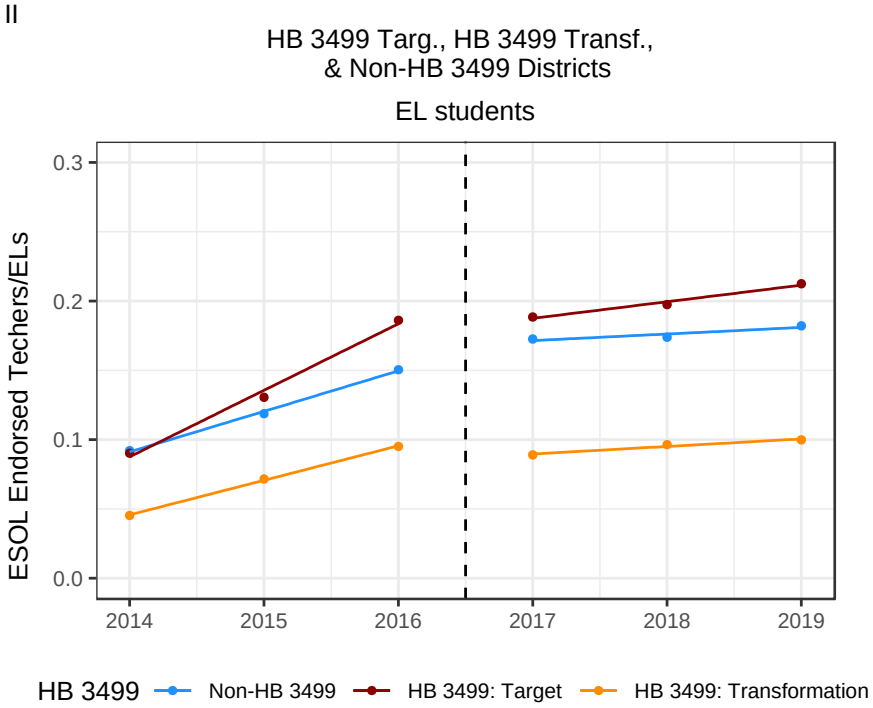
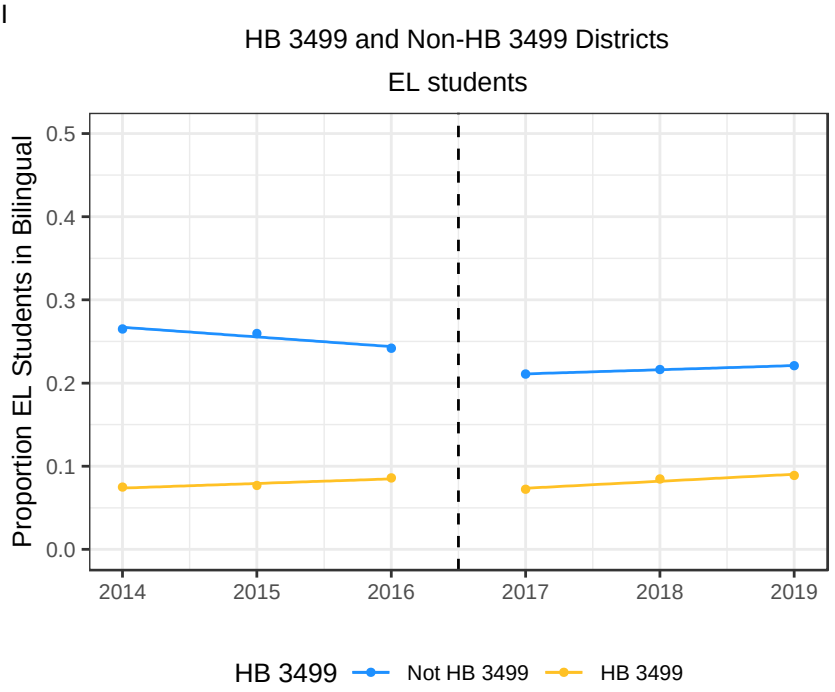


Figure A4. Proportion of EL-Classified Students Receiving Bilingual Services Over Time, by Whether Districts were Identified for HB 3499

Panel I



Panel II

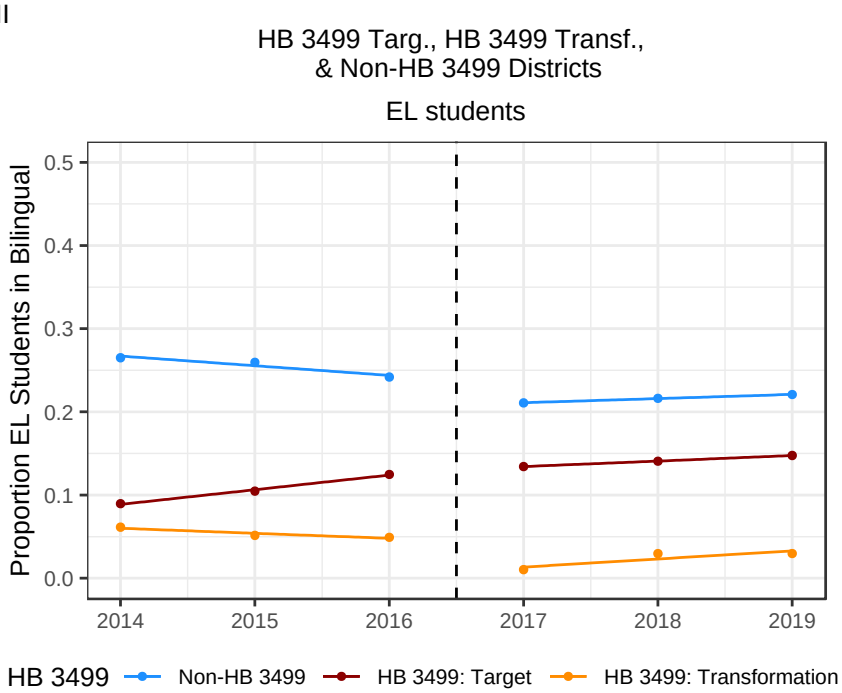
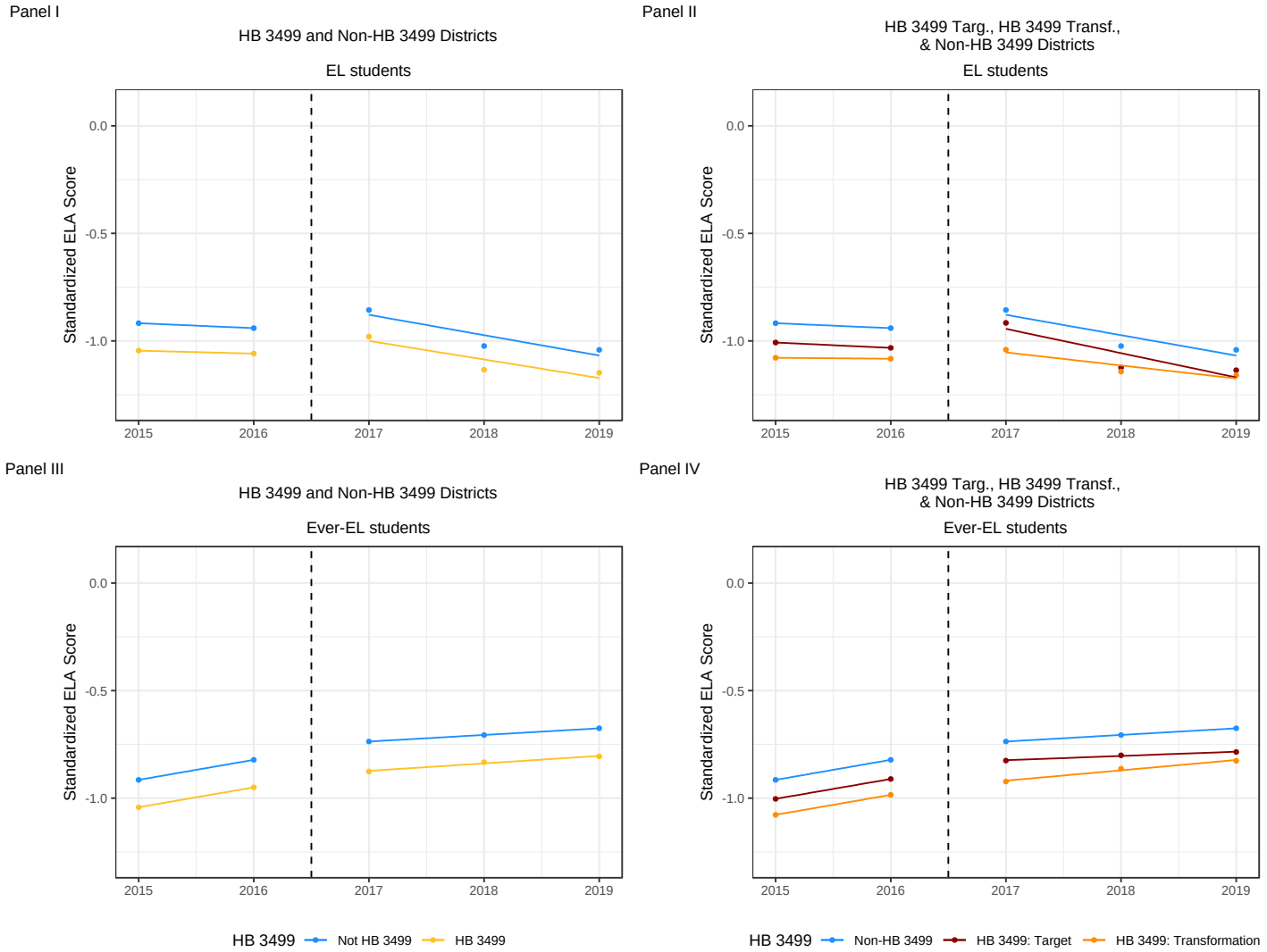
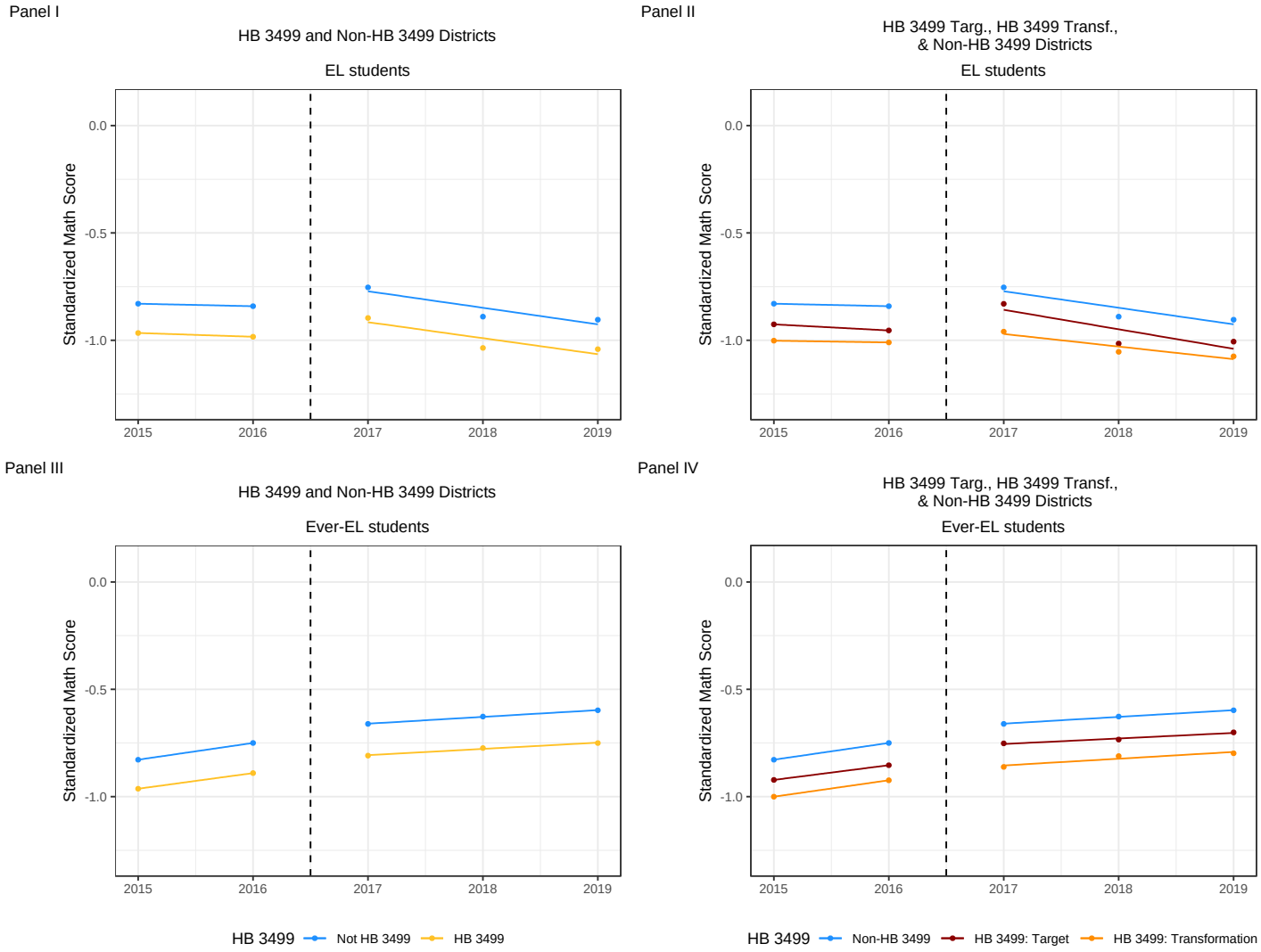


Figure A5. English Language Arts Outcomes Over Time, by Whether Districts were Identified for HB 3499



Note. Figures present a linear line of best fit using raw student level data, grouped by whether the outcome is from an HB 3499 district or not (Panels I and III) or a Target district, Transformation district, or non-HB 3499 district (Panels II and IV).

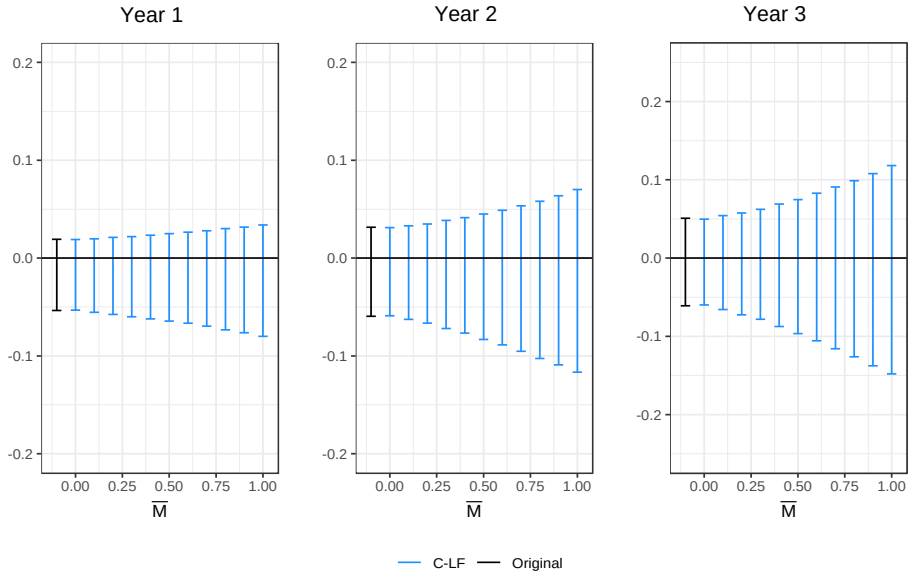
Figure A6. Math Outcomes Over Time, by Whether Districts were Identified for HB 3499



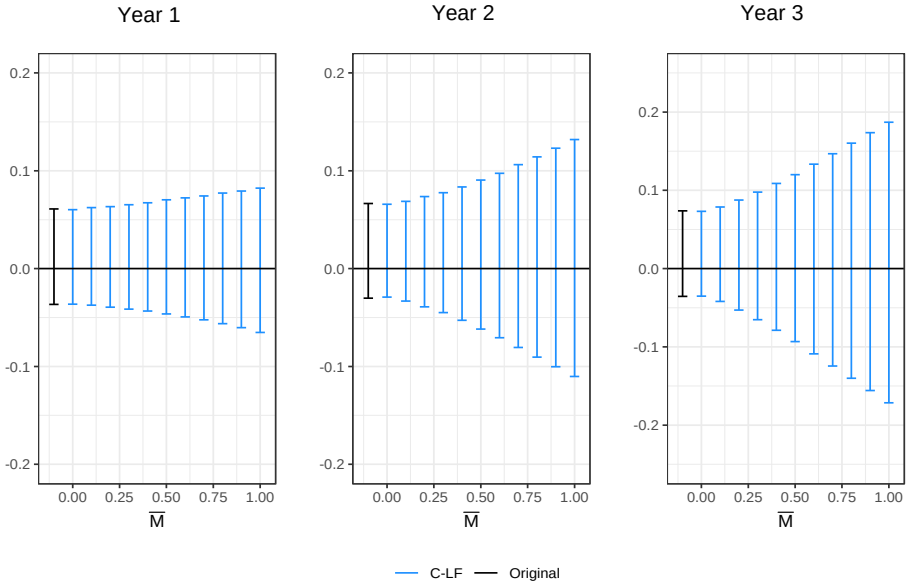
Note. Figures present a linear line of best fit using raw student level data, grouped by whether or not the outcome is from an HB 3499 district or not (Panels I and III) or a Target district, Transformation district, or non-HB 3499 district (Panels II and IV).

Figure A7. Changes in Teacher Endorsement and Bilingual Services Point Estimates Under Modeled Violations in Pre-Trends

Panel I. ESOL Endorsed Teachers/EL-Classified Students

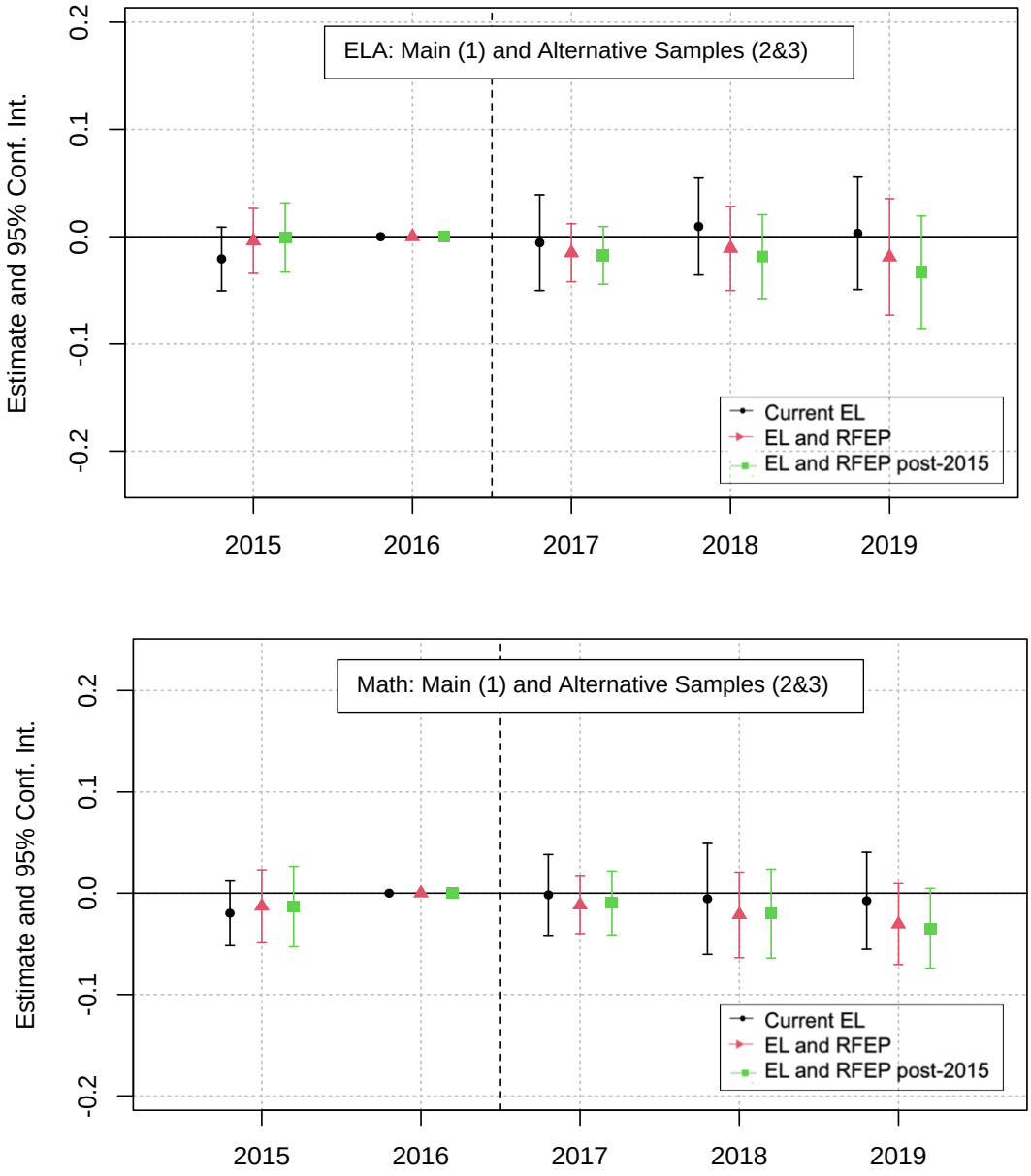


Panel II. Bilingual Services



Note. Estimates reflect modeled violations of the post-treatment trend as varying relative magnitudes in relation to pre-trend violations. Estimates were generated and plotted using the HonestDiD packages (Rambachan & Roth, 2022). Analytic code was supported with code examples from the DiD Reading Group GitHub (Wright, 2022).

Figure A8. Event Study Estimates for Alternative Samples: Academic Outcomes



Note. Each figure plots estimates from models that include grade, year, and district fixed effects, as well as student, school, and district covariates. The first estimate plotted is the main estimate, which includes outcomes for all current EL-classified students in a given year, while the next estimate comes from a sample of EL-classified students as well as students who were reclassified as fluent English proficient during the panel, regardless of whether they were reclassified before or after the intervention. The third estimate draws on a sample of current EL-classified students as well as students who were reclassified post-intervention.

Figure A9. Event Study Estimates of the Effect of HB 3499 Identification on Never-EL-classified Students' Outcomes

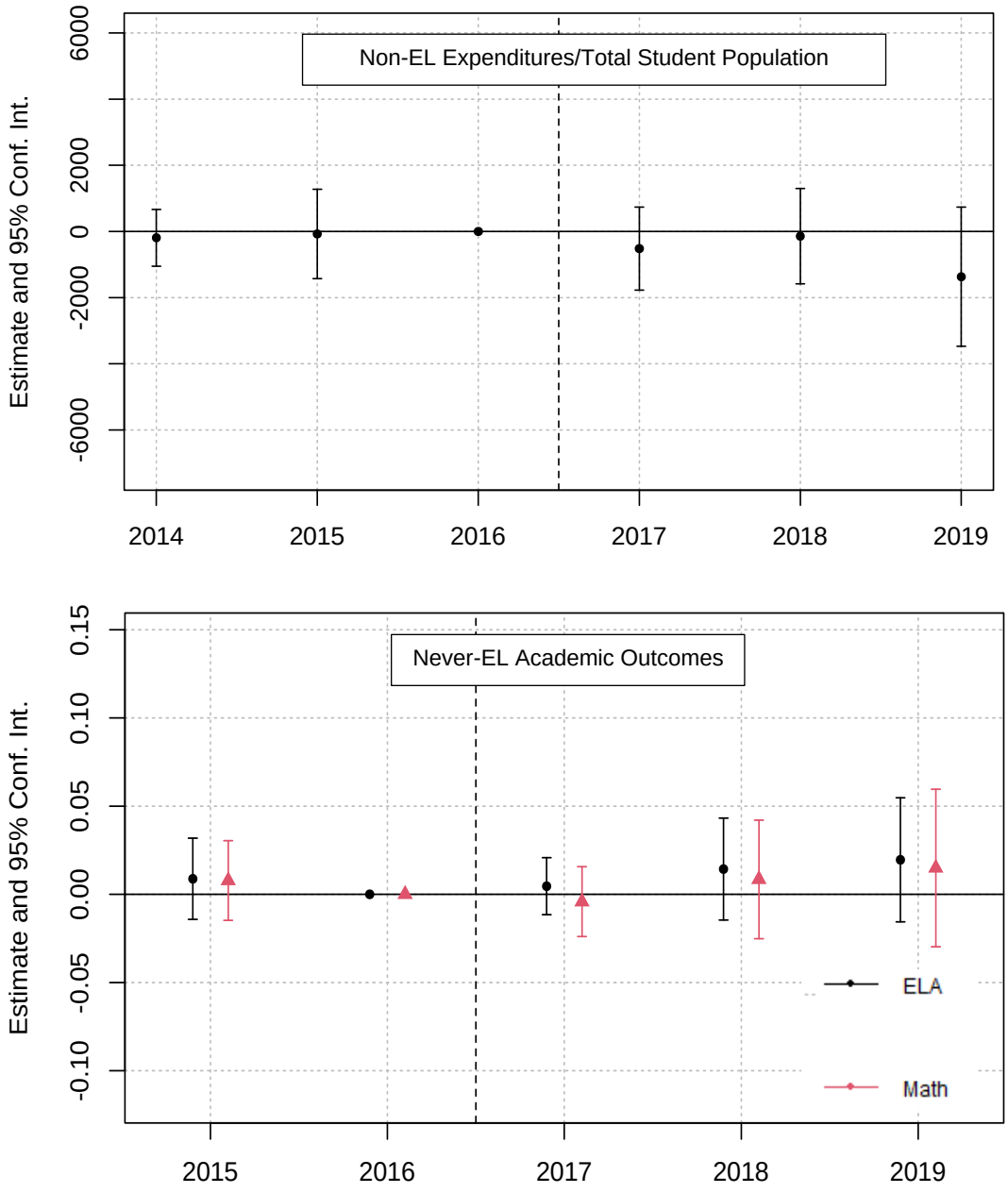
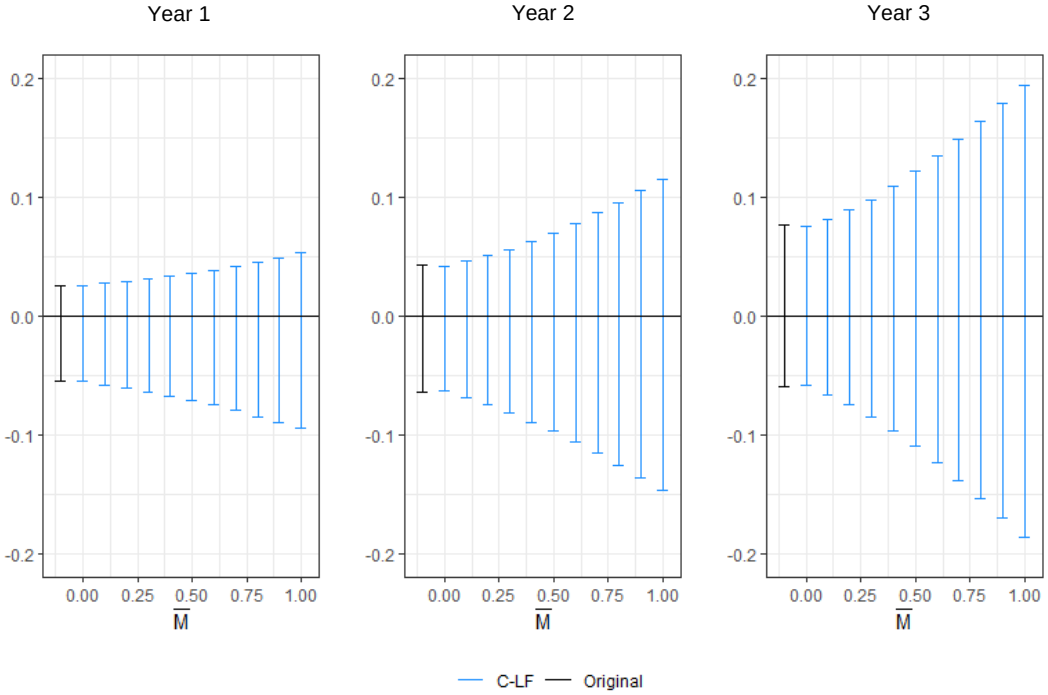
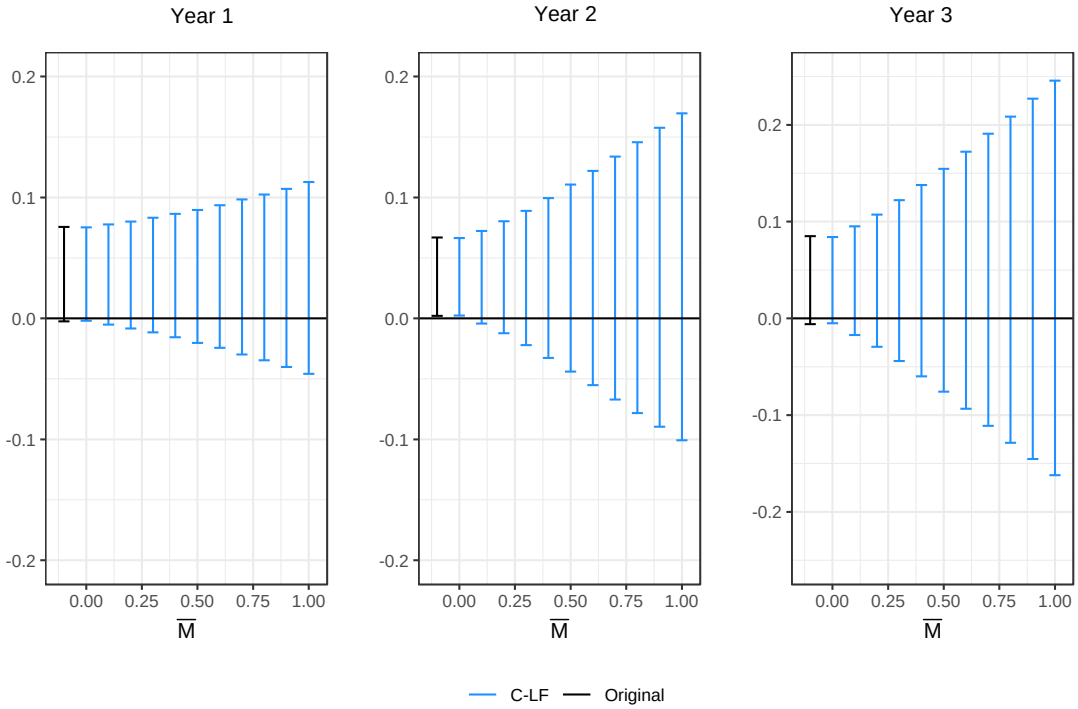


Figure A10. Changes in ESOL Endorsed Teachers and Bilingual Services Point Estimates Under Modeled Violations in Pre-Trends (Target Districts)

Panel I. ESOL Endorsed Teachers, Target Districts



Panel II. Bilingual Services, Target District



APPENDIX B

In this Appendix B, I provide a more detailed description of the regression discontinuity (RD) analyses and assumptions. I then provide more detailed results.

Analytic Approach

There are multiple ways to approach a multiple rating score RD (Reardon & Robinson, 2012). In cases where it is hypothesized that the effect may differ by threshold, fitting separate RDs at each threshold is appropriate (Dee, 2012; Reardon & Robinson, 2012). The two thresholds represent different district contexts—with one subset facing higher need, but higher in outcomes, and the other being relatively lower need, but also lower outcomes. As such, I argue it is tenable that an intervention that includes additional funds, technical assistance, and accountability may impact districts at each threshold differently. For example, districts with higher outcomes prior to treatment may have stronger instructional systems in place prior to the intervention than districts with lower outcomes. Having these systems in place may lead to a stronger, more immediate impact. As another example, districts that are higher in need may face compounding structural issues that mean an intervention may not be large enough in scope to see meaningful change, as compared to a relatively lower-need district context.

Working from the assumption that the effect may be different at each threshold, I create two subsamples of districts, one comprised of a bandwidth of districts above the needs threshold (high-need) and at the outcomes threshold, and one comprised of districts below the outcomes threshold (low-outcome) at the needs threshold. For each threshold I restrict the sample to be districts within ten percentile points from the threshold. As plotted in Figure B1, Panel 1, there is perfect compliance with the assignment mechanism among high-need districts at the outcome threshold, meaning that no district above the needs index threshold that had an index value above the outcomes-based threshold was identified for HB 3499, and that all districts above the needs index that had an index value below

the outcomes-based threshold were identified. In contrast, there is not perfect compliance with the assignment mechanism among low-outcome districts at the needs threshold. Specifically, 13 districts that fell below the outcomes threshold and were below the needs threshold (i.e., did not have high enough need) were identified, despite not aligning with the identification criteria (Figure B1, Panel II). However, within the bandwidth I specify of ten percentile points, treatment is perfectly predicted by whether a district is above or below the 47th percentile.

Description of Bandwidth Samples

For both bandwidth samples, there were differences between HB 3499 districts and non-HB 3499 districts. Table B1 presents student-level descriptive statistics in 2015/16 for the eleven high-need districts that were within ten percentile points of the outcomes threshold. Table B2 presents student-level descriptive statistics in 2015/16 for the thirteen low-outcome districts that were within ten percentile points of the needs threshold. Among high-need districts at the outcomes threshold, a lower proportion of EL-classified students in HB 3499 districts were Hispanic/Latinx, and a higher proportion were Black, White, and Asian/Pacific Islander. HB 3499 districts had, on average, student populations that had a lower proportion of EL or ever-EL classified, non-White, and eligible for free/reduced price lunch students. Additionally, in the year prior to identification, a far lower proportion of students received bilingual services in HB 3499 districts, and for the academic samples standardized scores were lower in HB 3499 districts. While not presented in tabular form, on average in the year prior to identification HB 3499 districts in the bandwidth of high-need districts around the outcome threshold had a slightly higher ratio of endorsed teachers to EL-classified students than non-HB 3499 districts and slightly higher EL expenditures per EL-classified students (see Figure B2).

Among low-outcome districts at the needs threshold, a higher proportion of EL-classified students were eligible for free/reduced price lunch in HB 3499 districts, while a lower proportion were ever identified for special education services. There were very small differences in the

racial/ethnic make-up of EL-classified students in the sample, as the vast majority of EL-classified students across the two groups were identified as Hispanic/Latinx. On average, HB 3499 districts had a lower proportion of students who were EL or ever-EL classified and non-White, and a slightly higher proportion of students eligible for free/reduced price lunch. Very few EL-classified students in either HB 3499 or non-HB 3499 districts received bilingual services, close to zero percent among non-HB 3499 districts, and 1 percent among HB 3499 districts. Standardized ELA and math scores were higher in HB 3499 districts. There were more endorsed teachers per EL-classified student, on average, in non-HB 3499 districts, while EL spending per EL-classified student was higher in HB 3499 districts. These differences prior to intervention are important indicators that the two groups of districts, at either threshold, do not have the same average characteristics prior to intervention, either with regard to their characteristics, or the outcomes of interest.

Additional Regression Discontinuity Assumptions

Additional assumption checks (Figure B2) reveal that there are also visible discontinuities in the characteristics of the EL-classified student population at the threshold. Among high-need districts at the outcome threshold, there is a visible jump in the proportion of EL-classified students who are eligible for free/reduced price lunch and non-white moving from the treated group to the untreated group. Among low-outcome districts at the needs threshold, there is a visible drop in the proportion of EL-classified students who are eligible for free/reduced price lunch and non-white moving from the untreated group to the treated group. It is concerning that there would be discontinuities at the threshold, potentially providing insight into how the threshold was placed in such a way to concentrate treatment among districts with certain characteristics.

Taken together, the compounding evidence that there are concerning differences in the districts within the bandwidth is an argument for integrating the RD approach with a difference-in-differences approach. The difference-in-differences approach allows for baseline differences on the

outcome of interest, as well as assignment to treatment that is not necessarily as good as random. The integrated approach, then, will reflect the use of an RD to hopefully create groups where treatment is more likely to not be related to district characteristics, as well as the use of a difference-in-differences to account for the baseline differences in treatment and control groups. However, the same assumptions as described in the main analysis should hold, including parallel pre-trends.

Parallel Trends Assumption

Figures B3 and B4 contain the plotted line of best of fit for the raw data among districts at the threshold for the different observed outcomes. For high-need districts at the outcome threshold, in the raw data the trends appear to be parallel for expenditure and academic outcomes. There is a slight convergence among the endorsed teacher trends, and for bilingual services trends are clearly not following parallel trajectories, with a negative trend among non-HB 3499 districts and a positive trend among HB 3499 districts. For low-outcome districts at the needs threshold, in the raw data the trends appear to be parallel for expenditure outcomes, as well as teacher endorsements. Math trends follow similar trajectories, while there is a slight convergence among average ELA scores. For bilingual services trends are also not following parallel trajectories, with a negative trend among non-HB 3499 districts and a positive trend among HB 3499 districts.

These raw data point to compounding concerns about the ability of the difference-in-differences approach, once integrated with an RD to identify theoretically more comparable groups, to estimate an unbiased effect of the policy in this context. Specifically, estimates among districts at the threshold for bilingual services are unlikely to represent the true effect of the policy, while there are also serious concerns for ELA estimates among low-outcome districts at the needs threshold. In the following section I present results for each outcome, but also note here, and throughout, that these estimates may be limited in the information we can take away from the supplemental analysis.

Findings

High-Need Districts at the Outcomes Threshold

There was no significant effect of HB 3499 identification on the five outcomes examined among the subset of high-need districts at the outcomes threshold. Table B3 presents the estimates from the preferred specification, with a null, imprecise estimate for expenditures, a precise null for ESOL endorsed teachers/EL-classified students, and null and somewhat imprecise estimates for bilingual services and academic outcomes. In all cases except teacher endorsements, findings are less precise than the main estimate, which is not surprising given the small number of districts in the bandwidth. Among the subsample, I am less confident in the assertation that there was no detectable effect, as the confidence intervals include moderate to large effects for all outcomes other than endorsements, meaning I am less able to rule out the presence of sizeable effects among the subset of districts. Another way that results are not aligned with the main findings are the expenditures, where no significant effect was detected. Again, the confidence interval is wide, making it difficult to make any claim about the policy impact.

Low-Outcome Districts at the Needs Threshold

In contrast with the other bandwidth, there was a significant effect of identification on both EL expenditures per EL-classified student and the ratio of ESOL endorsed teachers to EL-classified students. The expenditure estimate is larger in magnitude than the main estimates, but again the confidence interval spans moderate increases ($267.8 increase per EL-classified student) to large increases ($2,430.7 per EL-classified student), meaning that we cannot rule out changes that range widely in magnitude. For endorsed teachers, the effect is moderate, about one additional ESOL endorsed teacher per 13 EL-classified students, although we cannot rule out very small changes (0.012) or larger changes (0.142).

As with the other bandwidth, there was no significant effect on bilingual services or academic outcomes. The bilingual services estimate was small, with a confidence interval that spanned small

negative effects to moderate positive effects. For both ELA and math outcomes, the confidence interval spanned moderate to large negative and positive effects, with the imprecision inhibiting my ability to conclude whether the policy had no overall effect.

Appendix B Tables and Figures

Table B1. Descriptive Statistics for EL-Classified Students in the Bandwidth (10 Percentile Points) of High-Need Districts at the Outcomes Threshold (2015/16)

	Bilingual Program (K-12)		English Language Arts (3-8)		Math (3-8)	
	Non-HB 3499 (N =12,258)	HB 3499 (N =5,375)	Non-HB 3499 (N =5,755)	HB 3499 (N = 2,440)	Non-HB 3499 (N = 5,790)	HB 3499 (N = 2,469)
Female	0.47	0.46	0.46	0.44	0.46	0.44
Identified for FRPL	0.91	0.95	0.93	0.95	0.93	0.95
Identified for special education	0.19	0.21	0.22	0.24	0.22	0.24
Race/Ethnicity						
AIAN	0.00	0.00	0.00	0.00	0.00	0.00
Asian/Pacific Islander	0.09	0.16	0.09	0.15	0.09	0.15
Black	0.01	0.08	0.01	0.07	0.01	0.07
Hispanic	0.80	0.65	0.82	0.69	0.82	0.69
Multi-Ethnic	0.01	0.01	0.01	0.01	0.01	0.01
White	0.08	0.10	0.08	0.09	0.08	0.09
% In bilingual services	0.38	0.15	--	--	--	--
ELA score	--	--	-0.94	-1.06	--	--
Math score	--	--	--	--	-0.87	-0.99
District chars.						
% ELs	0.18	0.10	0.18	0.10	0.18	0.10
% Ever-ELs	0.32	0.20	0.32	0.20	0.32	0.20
% Non-White	0.51	0.43	0.51	0.43	0.51	0.43
% FRPL	0.66	0.54	0.67	0.54	0.67	0.54
% Special Education	0.16	0.14	0.16	0.14	0.16	0.14

Note. EL=English learner. HB=House bill. FRPL=Free/reduced price lunch. AIAN-American Indian or Alaska Native. ELA=English language arts. Statistics report the average value of student or district-level characteristics in the year prior to HB 3499 identification by whether the district was identified in the following year.

Table B2. Descriptive Statistics for EL-Classified Students in the Bandwidth (BW=10 Percentile Points) for Low-Outcome Districts at the Needs Threshold (2015/16)

	Bilingual Program (K-12)		English Language Arts (3-8)		Math (3-8)	
	Non-HB 3499 (N =666)	HB 3499 (N =2,406)	Non-HB 3499 (N = 302)	HB 3499 (N = 1,043)	Non-HB 3499 (N = 301)	HB 3499 (N = 1,069)
Female	0.46	0.46	0.42	0.46	0.42	0.46
Identified for FRPL	0.80	0.93	0.87	0.93	0.87	0.93
Identified for special education	0.24	0.19	0.34	0.24	0.35	0.23
Race/Ethnicity						
AIAN	0.01	0.00	0.00	0.00	0.00	0.00
Asian/Pacific Islander	0.01	0.02	0.01	0.02	0.01	0.02
Black	0.00	0.00	0.00	0.00	0.00	0.00
Hispanic	0.96	0.95	0.96	0.95	0.96	0.95
Multi-Ethnic	0.00	0.01	0.00	0.00	0.00	0.00
White	0.02	0.02	0.02	0.02	0.02	0.02
% In bilingual services	0.00	0.01	--	--	--	--
ELA score	--	--	-0.97	-0.95	--	--
Math score	--	--	--	--	-0.93	-0.84
District chars.						
% ELs	0.15	0.13	0.16	0.15	0.16	0.15
% Ever-ELs	0.32	0.25	0.31	0.27	0.31	0.27
% Non-White	0.47	0.40	0.49	0.47	0.49	0.47
% FRPL	0.55	0.57	0.62	0.65	0.62	0.65
% Special Education	0.13	0.13	0.14	0.14	0.14	0.14

Note. EL=English learner. HB=House bill. FRPL=Free/reduced price lunch. AIAN-American Indian or Alaska Native. ELA=English language arts. Statistics report the average value of student or district-level characteristics in the year prior to HB 3499 identification by whether the district was identified in the following year.

Table B3. Estimated Effect of HB 3499 Identification at Needs and Outcomes Threshold

	Expenditures				ESOL Endorsed Teachers				Bilingual Services				ELA				Math			
	Low-Outcome at Needs Threshold		High-Need at Outcomes Threshold		Low-Outcome at Needs Threshold		High-Need at Outcomes Threshold		Low-Outcome at Needs Threshold		High-Need at Outcomes Threshold		Low-Outcome at Needs Threshold		High-Need at Outcomes Threshold		Low-Outcome at Needs Threshold		High-Need at Outcomes Threshold	
	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI	Est.	95% CI
HB 3499	1,349.3*	[267.8; 2,430.7]	-491.5	[-1,302.3; 319.2]	0.077*	[0.012; 0.142]	-0.013	[-0.028; 0.003]	0.029	[-0.016; 0.074]	0.060	[-0.011; 0.132]	-0.026	[-0.128; 0.076]	0.047	[-0.005; 0.100]	0.009	[-0.137; 0.155]	0.053	[-0.003; 0.109]
Fixed Effects																				
Grd.	No		No		No		No		Yes		Yes		Yes		Yes		Yes		Yes	
Year	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Dist.	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Obs.	78		66		78		66		19,145		108,130		7,279		43,003		7,423		98,551	
R2	0.702		0.746		0.874		0.091		0.172		0.331		0.164		0.144		0.186		0.146	
Cond. R2	0.214		0.163		0.353		0.119		0.078		0.125		0.078		0.081		0.108		0.010	

Note. *p<.05, **p<.01, ***p<.001. Est=Estimate, CI=Confidence interval, Grd=Grade. All models include district and year fixed effects as well as district covariates, while student-level specifications include grade fixed effects, student covariates, and school covariates.

Figure B1. HB 3499 Identification Rate for Districts at the Needs (Top) and Outcomes (Bottom) Threshold

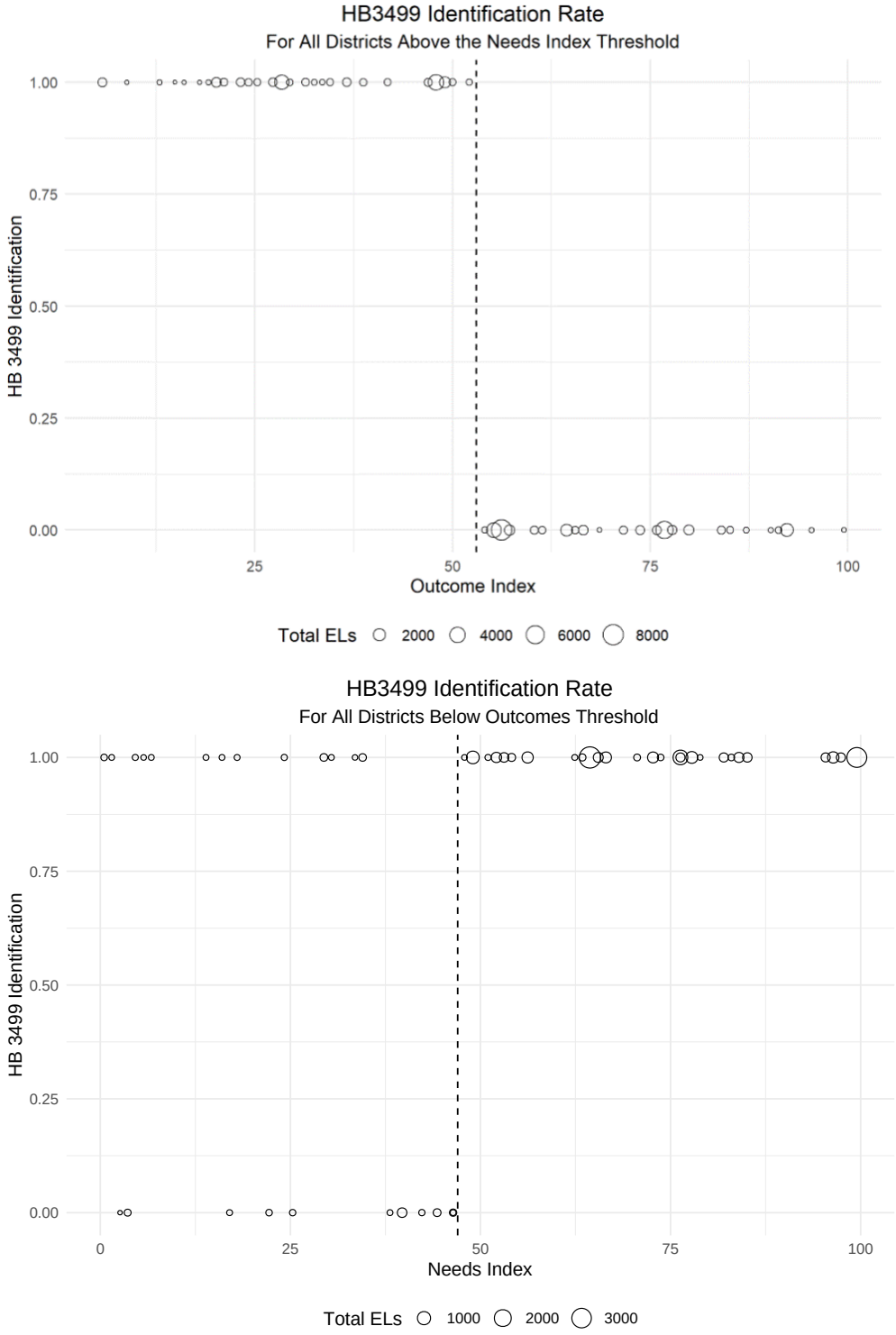


Figure B2. Plotted Covariates Among EL-Classified Students in at the Needs and Outcomes Threshold

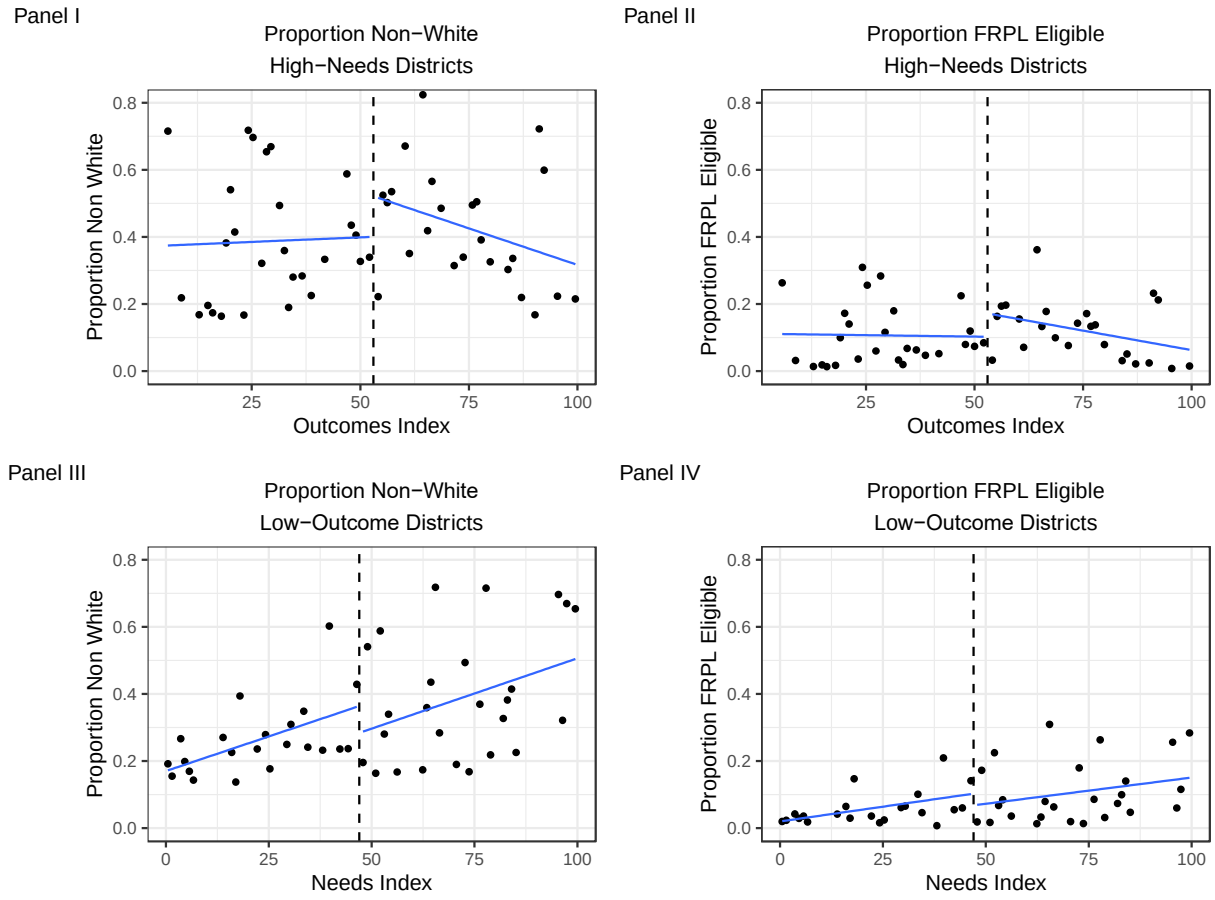


Figure B3. High Need Districts at Outcome Threshold

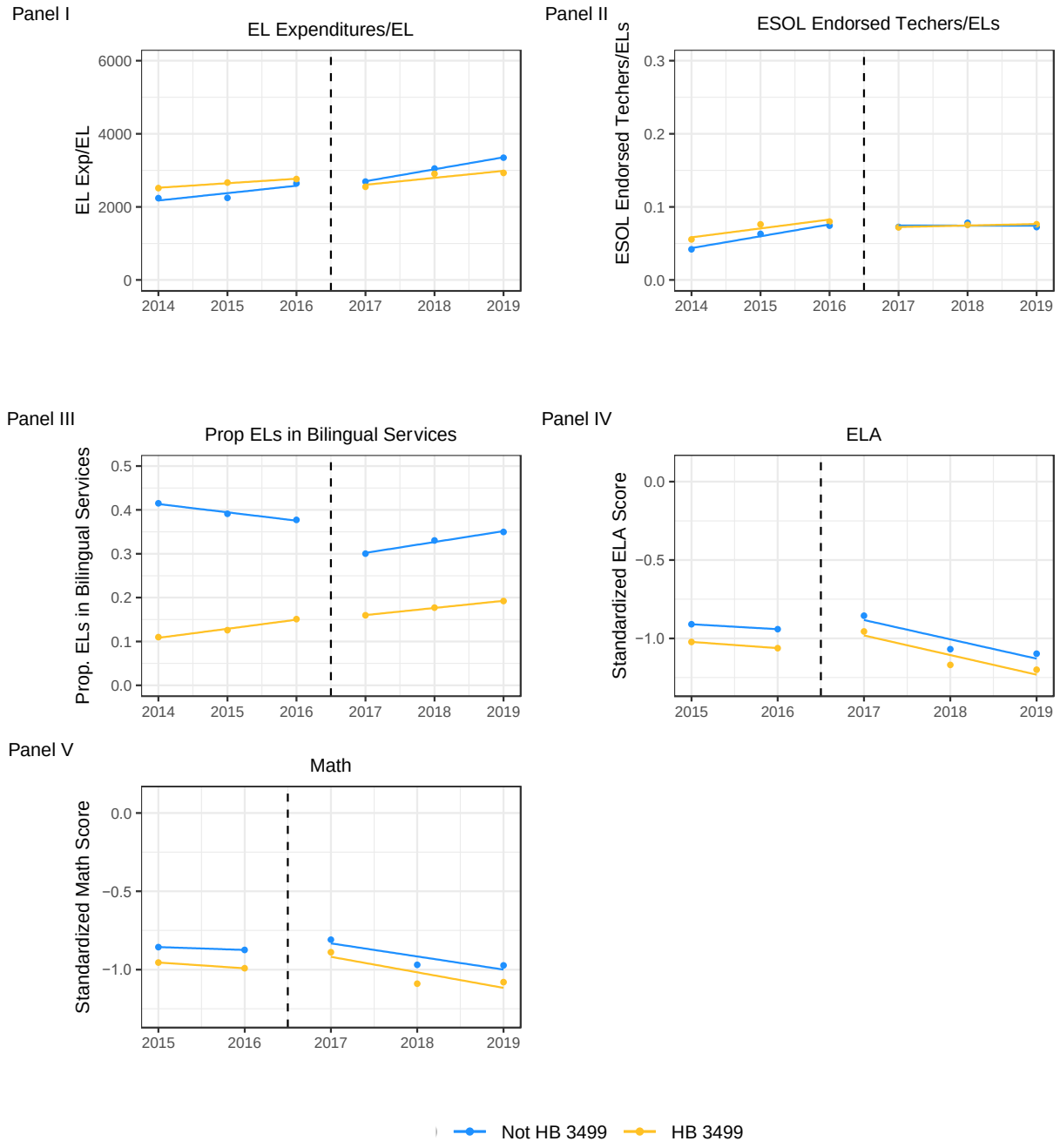
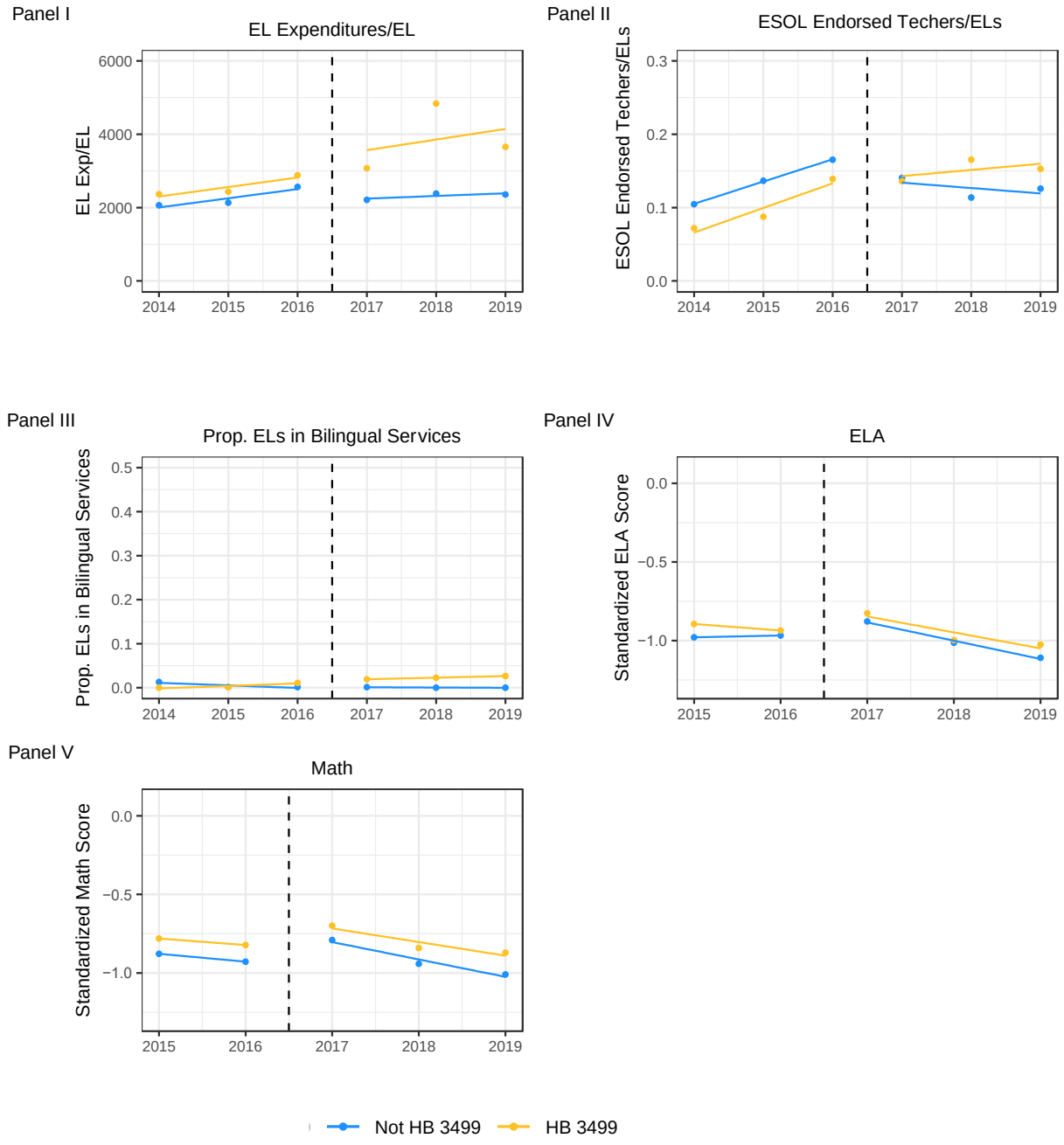


Figure B4. Low Outcome Districts at Needs Threshold



REFERENCES CITED

Angrist J. D., Pischke J. S. (2008). Mostly harmless econometrics: An empiricist's companion. Princeton University Press.

Armas, E., Lavadenz, M., & Olsen, L. (2015). Falling short on the promise to English Learners: A report on year one LCAPS. *Californians Together*. Retrieved from https://californianstogether. app. box. com/s/g4zg47vhgoqwzpm2dqpry2pdxzmbhman.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics, 119*(1), 249-275.

Bonilla, S., & Dee, T. S. (2020). The effects of school reform under NCLB waivers: Evidence from focus schools in Kentucky. *Education Finance and Policy, 15*(1), 75-103.

Callahan, R. M. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal, 42*(2), 305-328.

Callahan, R. M., & Shifrer, D. (2016, Aug). Equitable Access for Secondary English Learner Students: Course Taking as Evidence of EL Program Effectiveness. *Educational Administration Quarterly, 52*(3), 463-496.

Calonico, S., Cattaneo, M. D., & Farrell, M. H. (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal*, *23*(2), 192-210.

Carlson, D., & Lavertu, S. (2018). School improvement grants in Ohio: Effects on student achievement and school administration. *Educational Evaluation and Policy Analysis, 40*(3), 287-315.

Carnock, J. (2017). *Pioneering change: Leveraging data to reform English learner education in Oregon*. New America Foundation.

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

Chiang, H. (2009). How accountability pressure on failing schools affects student achievement. *Journal of Public Economics, 93*(9-10), 1045-1057.

Chitiyo, J. (2017). Challenges to the use of co-teaching by teachers. *International Journal of Whole Schooling*, *13*(3), 55-66.

Cunningham, S. (2021). *Causal inference: The mixtape*. Yale University Press.

Dabach, D. B. (2014). "I am not a shelter!": Stigma and social boundaries in teachers' accounts of students' experience in separate "sheltered" English learner classrooms. *Journal of Education for Students Placed at Risk 19*(2), 98-124.

Darling-Hammond, L. (2007). Race, inequality and educational accountability: The irony of 'No Child Left Behind'. *Race Ethnicity and Education*, *10*(3), 245-260.

Darling-Hammond, L., Bae, S., Cook-Harvey, C. M., Lam, L., Mercer, C., Podolsky, A., & Stosich, E. L. (2016). Pathways to new accountability through the Every Student Succeeds Act. *Palo Alto, CA: Learning Policy Institute*.

Dee, T. (2012). *School turnarounds: Evidence from the 2009 stimulus*. NBER Working Paper

Dee, T. S., & Dizon-Ross, E. (2019). School performance, accountability, and waiver reforms: Evidence from Louisiana. *Educational Evaluation and Policy Analysis, 41*(3), 316-349.

DeNicolo, C. P. (2016). "School within a school": Examining implementation barriers in a Spanish/English transitional bilingual education program. *Bilingual Research Journal*, *39*(2), 91-106.

Dickey-Griffith, D. (2013). Preliminary effects of the school improvement grant program on student achievement in Texas. *The Georgetown Public Policy Review, 2*, 21-39.

Dougherty, S. M., & Weiner, J. M. (2019). The Rhode to turnaround: The impact of waivers to No Child Left Behind on school performance. *Educational Policy, 33*(4), 555-586.

Education Commission of the States (2020). 50-State Comparison: K-12 and Special Education Funding. Retrieved from: https://www.ecs.org/50-state-comparison-k-12-and-special-education-funding/

Equal Educational Opportunities Act (1974) 20 USC Sec. 1701-1758

Figlio, D. N., & Ladd, H. F. (2008). School accountability and student achievement. *Handbook of Research in Education Finance and Policy*, 166-182.

Fullan, M. (1992). *Successful school improvement: The implementation perspective and beyond*. McGraw-Hill Education (UK).

Gándara, P., & Hopkins, M. (2010). English learners and restrictive language policies. *Teachers College Press.*

Gándara, P., Maxwell-Jolly, J., & Driscoll, A. (2005). Listening to teachers of English language learners: A survey of California teachers' challenges, experiences, and professional development needs.

Gándara, P., Maxwell-Jolly, J., Garcia, E., Asato, J., Gutierrez, K., Stritikus, T., & Curry, J. (2000). The initial impact of Proposition 227 on the instruction of English learners.

Gándara, P., & Orfield, G. (2012a). Segregating Arizona's English learners: A return to the "Mexican room"? *Teachers College Record, 114*(9), n9.

Gándara, P., & Orfield, G. (2012b). Why Arizona matters: The historical, legal, and political contexts of Arizona's instructional policies and US linguistic hegemony. *Language Policy, 11*(1), 7-19.

García, E., & Weiss, E. (2018). Student Absenteeism: Who Misses School and How Missing School Matters for Performance. *Economic Policy Institute*.

Goldschmidt, P., & Hakuta, K. (2017). Incorporating English Learner Progress into State Accountability Systems. *Council of Chief State School Officers*.

Guo, Q., & Koretz, D. (2013). Estimating the impact of the Massachusetts English immersion law on limited English proficient students' reading achievement. *Educational Policy*, *27*(1), 121-149.

Hamilton, L. S., Stecher, B. M., Russell, J. L., Marsh, J. A., & Miles, J. (2008). Accountability and teaching practices: School-level actions and teacher responses. In *Strong states, weak schools: The benefits and dilemmas of centralized accountability*. Emerald Group Publishing Limited.

Hanushek, E. A., & Rivkin, S. G. (2012). The distribution of teacher quality and implications for policy. *Annual Review Economics*, *4*(1), 131-157.

Heissel, J. A., & Ladd, H. F. (2016). School turnaround in North Carolina: A regression discontinuity analysis. Working Paper 156. *National Center for Analysis of Longitudinal Data in Education Research*.

Henry, G. T., Fortner, C. K., & Thompson, C. L. (2010). Targeted funding for educationally disadvantaged students: A regression discontinuity estimate of the impact on high school student achievement. *Educational Evaluation and Policy Analysis, 32*(2), 183-204.

Henry, G. T., Guthrie, J. E., & Townsend, L. (2015). Outcomes and impacts of North Carolina's initiative to turn around the lowest-achieving schools. *The Friday Institute for Educational Innovation, North Carolina State University*.

Hiatt, J. E., & Fairbairn, S. B. (2018). Improving the focus of English learner professional development for in-service teachers. *NASSP Bulletin*, *102*(3), 228-263.

Holbein, J. B., & Ladd, H. F. (2017). Accountability pressure: Regression discontinuity estimates of how No Child Left Behind influenced student behavior. *Economics of Education Review, 58*, 55-67.

Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies, 79*(3), 933-959.

Jackson, C. K. (2020). Does school spending matter? The new literature on an old question. In L. Tach, R. Dunifon, & D. L. Miller (Eds.), *Confronting inequality: How policies and practices shape children's opportunities* (pp. 165–186). American Psychological Association.

Jackson, C. K., Johnson, R. C., & Persico, C. (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *The Quarterly Journal of Economics, 131*(1), 157-218.

Jimenez-Castellanos, O., Gonzalez, S., & Long, D. (2022). *HB 3499 and Beyond: Report for the Oregon Department of Education*. Retrieved from: https://www.oregon.gov/ode/students-and-family/equity/EngLearners/Documents/ODE%20Directed%20Funding%20FULL%20Report_Oscar%20Jimenez-Castellanos_July%2031%202020.pdf

Jimenez-Castellanos, O., & Topper, A. M. (2012). The cost of providing an adequate education to English language learners: A review of the literature. *Review of Educational Research, 82*(2), 179-232.

Kanno, Y., & Cromley, J. G. (2013). English language learners' access to and attainment in postsecondary education. *TESOL Quarterly, 47*(1), 89-121

.

Kanno, Y., & Kangas, S. E. (2014). "I'm not going to be, like, for the AP" English language learners' limited access to advanced college-preparatory courses in high school. *American Educational Research Journal, 51*(5), 848-878.

Khong, T. D. H., & Saito, E. (2014). Challenges confronting teachers of English language learners. *Educational Review*, *66*(2), 210-225.

Levin, H. M. (1974). A conceptual framework for accountability in education. *The School Review, 82*(3), 363-391.

Linquanti, R. (2014). Fostering success for English Learners in turnaround schools: What state education agencies need to know and be able to do. *The State Role in School Turnaround: Emerging Best Practices*, 195.

Louis, K. S., Febey, K., & Schroeder, R. (2005). State-mandated accountability in high schools: Teachers' interpretations of a new era. *Educational Evaluation and Policy Analysis, 27*(2), 177-204.

Lucas, T. (2011). Language, schooling, and the preparation of teachers for linguistic diversity. *Teacher preparation for linguistically diverse classrooms: A resource for teacher educators*, 3-17.

Lyons, S., Dadey, N., & Garibay-Mulattieri, K. (2017). Considering English language proficiency within systems of accountability under the Every Student Succeeds Act. *Center for Assessment*.

Martínez, D. G., & Spikes, D. D. (2020). Se Acabaron Las Palabras: A Post-Mortem Flores v. Arizona Disproportional Funding Analysis of Targeted English Learner Expenditures. *Educational Policy*. *34*(1), 324-365.

Master, B., Loeb, S., Whitney, C., & Wyckoff, J. (2016). Different skills? Identifying differentially effective teachers of English language learners. *The Elementary School Journal*, *117*(2), 261-284.

Menken, K., & Solorza, C. (2014). No child left bilingual: Accountability and the elimination of bilingual education programs in New York City schools. *Educational Policy*, *28*(1), 96-125.

Menken, K., & Solorza, C. (2015). Principals as linchpins in bilingual education: The need for prepared school leaders. *International Journal of Bilingual Education and Bilingualism, 18*(6), 676-697.

Mitra, D. L. (2017). *Educational change and the political process*. Routledge.

Morando Rhim, L, & Redding, S. (Eds.). (2014). *The state role in school turnaround: Emerging best practices*. Information Age Publishing.

Morita-Mullaney, T., & Singh, M. (2021). Obscuring English learners from state accountability: The case of Indiana's language blind policies. *Educational Policy*, *35*(4), 621-645.

Office of Refugee Resettlement (2017). Refugee School Impact Grants. Retrieved from: https://www.acf.hhs.gov/orr/programs/refugees/school-impact

Okhremtchouk, I. S. (2017). The Politics of Schools and Money: Building Awareness about Channeling Practices for Supplemental Resource Allocations to Serve English Language Learners. *education policy analysis archives*, *25*(17), n17.

Oregon Department of Education (2019). English Learner Report 2018-19. Oregon Department of Education.

Oregon Department of Education (2020). *EL Advisory Group Meeting, April 2020* (Powerpoint). Retrieved from: https://www.oregon.gov/ode/students-and-family/equity/EngLearners/Documents/April%20EL%20Advisory%20Slidedeck.pdf

Oregon Department of Education (2021a). *EL Advisory Meeting Minutes*, *January 2021*. Retrieved from: https://www.oregon.gov/ode/students-and-family/equity/EngLearners/Documents/EL%20Advisory%20Meeting%20%20Minutes%20January%208,%202021.pdf

Oregon Department of Education (2021b). *HB 3499 Cohort 1 Districts: Supplemental Information Concerning the New Identification and Determination System*. Retrieved from: https://www.oregon.gov/ode/students-and-family/equity/EngLearners/Documents/HB%203499%20Cohort%201%20-%20New%20Identification%20and%20Determination%20System.pdf

Oregon Department of Education (2022). *HB3499 EL School & District Improvement Program Cohort 1* Retrieved from: https://www.oregon.gov/ode/students-and-family/equity/EngLearners/Pages/EL-Cohort-1.aspx

Pacific Research and Evaluation (2020). HB 3499 ELL District and School Improvement Evaluation Report: Eagle Point School District. Retrieved from: https://www.pacific-research.org/_files/ugd/2f81b8_247dc9b9e63e4445835435e957b0e5e9.pdf

Palmer, D., & Snodgrass Rangel, V. (2011). High stakes accountability and policy implementation: Teacher decision making in bilingual classrooms in Texas. *Educational Policy*, *25*(4), 614-647.

Pandya, J. Z. (2011). *Overtested: How high-stakes accountability fails English language learners*. Teachers College Press.

Parrish, T. B., Merickel, A., Pérez, M., Linquanti, R., Socias, M., Spain, A., Speroni, C., Esra, P., Brock, L., & Delancey, D. (2006). Effects of the Implementation of Proposition 227 on the Education

of English Learners, K-12: Findings from a Five-Year Evaluation. Final Report for AB 56 and AB 1116. *American Institutes for Research*.

Pompa D. (2016). New education law puts more pressure on states to serve English learners. Retrieved from https://edsource.org/2016/new-education-law-puts-more-pressure-on-states-to-serve-english-learners/92811.

Rambachan, A., & Roth, J. (2019). An honest approach to parallel trends. Unpublished manuscript, Harvard University.

Reardon, S. F., & Robinson, J. P. (2012). Regression discontinuity designs with multiple rating-score variables. *Journal of Research on Educational Effectiveness, 5*(1), 83-104.

Rosa, J., & Flores, N. (2017). Unsettling race and language: Toward a raciolinguistic perspective. *Language in Society, 46*(5), 621-647.

Roth, J. (2019). Pre-test with caution: Event-study estimates after testing for parallel trends. *Department of Economics, Harvard University, Unpublished manuscript*.

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2022). What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature. *arXiv preprint arXiv:2201.01194*.

Steele, J. L., Slater, R. O., Li, J., Zamarro, G., Miller, T., & Bacon, M. (2018). Dual-language immersion education at scale: An analysis of program costs, mechanisms, and moderators. *Educational Evaluation and Policy Analysis*, *40*(3), 420-445.

Strunk, K. O., & McEachin, A. (2014). More than sanctions: Closing achievement gaps through California's use of intensive technical assistance. *Educational Evaluation and Policy Analysis*, *36*(3), 281-306.

Strunk, K. O., McEachin, A., & Westover, T. N. (2014). The use and efficacy of capacity-building assistance for low-performing districts: The case of California's district assistance and intervention teams. *Journal of Policy Analysis and Management*, *33*(3), 719-751.

Sugarman, J. (2016). Funding an equitable education for English learners in the United States. *Washington DC: Migration Policy Institute*.

Sugarman, J., Lazarín, M., Capps, R., Cardoso, J. B., Brabeck, K., Fix, M., Soto, A. G. R., Lowenhaupt, R., Yammine, J., & Morales, M. (2020). Educating English Learners during the COVID-19 Pandemic: Policy Ideas for States and School Districts. *Migration Policy Institute*.

Sun, M., Kennedy, A. I., & Loeb, S. (2021). The longitudinal effects of School Improvement Grants. *Educational Evaluation and Policy Analysis*, *43*(4), 647-667.

Sun, M., Penner, E. K., & Loeb, S. (2017). Resource-and approach-driven multidimensional change: Three-year effects of school improvement grants. *American Educational Research Journal, 54*(4), 607-643.

Tanenbaum, C., Boyle, A., Soga, K., Le Floch, K. C., Golden, L., Petroccia, M., Toplitz, M., Taylor, J., & O'Day, J. (2012). National Evaluation of Title III Implementation: Report on State and Local Implementation. *Office of Planning, Evaluation and Policy Development, US Department of Education*.

Taylor, J., Stecher, B., O'Day, J., Naftel, S., & Le Floch, K. C. (2010). State and Local Implementation of the" No Child Left Behind Act". Volume IX--Accountability under NCLB. *Office of Planning, Evaluation and Policy Development, US Department of Education*.

Thompson, K. D., Umansky, I. M., & Rew, W. J. (2022). Improving understanding of English learner education Through an expanded analytic framework. *Educational Policy*, Online first.

Umansky, I., Hopkins, M., Dabach, D. B., Porter, L., Thompson, K., & Pompa, D. (2018). Understanding and Supporting the Educational Needs of Recently Arrived Immigrant English Learner Students: Lessons for State and Local Education Agencies. *Council of Chief State School Officers*.

Umansky, I. M. (2016). To be or not to be EL: An examination of the impact of classifying students as English learners. *Educational Evaluation and Policy Analysis, 38*(4), 714-737.

Umansky, I. M., & Porter, L. (2020). State English Learner Education Policy: A Conceptual Framework to Guide Comprehensive Policy Action. *education policy analysis archives, 28*(17), n17.

U.S. Census (2021). Public School Spending Increases by Largest Amount in 11 Years. Retrieved from: https://www.census.gov/newsroom/press-releases/2021/public-school-spending-per-pupil.html

U.S. Department of Education (2016) Title III Grant FAQs. Retrieved from: https://ncela.ed.gov/title-iii-grant-faqs

U.S. Department of Education (2019). Title I. Retrieved from: https://nces.ed.gov/fastfacts/display.asp?id=158

Valenzuela, A., & Rubio, B. (2018). Subtractive schooling. *The TESOL encyclopedia of English language teaching*, 1-7.

Van der Klaauw, W. (2008). Breaking the link between poverty and low student achievement: An evaluation of Title I. *Journal of Econometrics*, *142*(2), 731-756.

Versiegen, D. A. (2017). State finance policies for English language learners: New findings from a 50-state survey. *Journal of Education Finance*, *2*(1) 338-355.

Walker, A., Shafer, J., & Iiams, M. (2004). Not in my classroom": Teacher attitudes towards English language learners in the mainstream classroom. *NABE Journal of Research and Practice*, *2*(1), 130-160.

Williams, C. (2020). *The Case for Expanding Federal Funding for English Learners.* The Century Foundation. Retrieved from: https://tcf.org/content/commentary/case-expanding-federal-funding-english-learners/?agreed=1

Yatsko, S., Lake, R., Nelson, E. C., & Bowen, M. (2012). Tinkering toward Transformation: A Look at Federal School Improvement Grant Implementation. *Center on Reinventing Public Education.*

Zarate, M. E., & Gàndara, P. (2019). Can the LCFF improve teaching and learning for EL students? A review of the emerging research in California and directions for future implementation. *Peabody Journal of Education, 94*(2), 157-175.

CHAPTER IV: VARIATION IN TEACHER CONTRIBUTIONS TO EL-CLASSIFIED STUDENTS' PERFORMANCE ON STANDARDIZED ENGLISH LANGUAGE ARTS AND ENGLISH LANGUAGE PROFICIENCY ASSESSMENTS

Teacher quality is positioned as a critical element of education systems. Variation in teachers' instructional effectiveness has been found to have significant implications for students' academic performance (Aaronson et al., 2007; Nye et al., 2004; Loeb et al., 2014; Rivkin et al., 2005), behavior (Jackson, 2018), and life course outcomes such as access to postsecondary education and earnings (Chetty et al., 2014). For students classified as English learners (EL) who are developing their English language proficiency (ELP) while simultaneously building content area skills and knowledge, teachers are framed both as important sources of support as well as potential gatekeepers to opportunity (Bunch, 2013; Dabach, 2014; Faltis et al., 2010; Gándara & Santibañez, 2016; Shea et al., 2018). However, studies that attempt to quantify the variation in teacher effectiveness and implications of this variation for student performance typically overlook the nuances of EL education that likely impact conclusions about teachers' roles in shaping EL-classified students' educational outcomes (Jones et al., 2013). Specifically, the provision of additional instruction on the English language (a core feature of EL education) is not explicitly addressed in examinations of teachers' unique contributions to EL-classified students' academic performance (Jones et al., 2013), nor do examinations of teacher contributions to student outcomes include ELP as an outcome.

EL-classified students have the right to instruction on the English language and access to meaningful core content learning opportunities, with developing core content skills and knowledge and English proficiency as important learning outcomes (Every Student Succeed Act, 2016). The delivery of English language and core content instruction will look different across contexts, as federal requirements leave a great deal of discretion in determining program models to the state and local level (*Castaneda v. Pickard*, 1984). In secondary schools, however, the most common structure is for EL-classified students to be provided with a separate class period dedicated to English language

development (ELD), and access core content in general education classrooms aside EL- and non-EL-classified peers alike (Lewis & Gray, 2016). As the name suggests, instruction in an ELD classroom is focused on supporting students' ELP development, including instruction on the domains of English writing, reading, speaking, and listening (DiCerbo et al., 2014). For core content instruction, teachers are typically tasked with delivering core content material (i.e., English language arts, math, science) to EL- and non-EL-classified students in the same instructional setting while adapting their instruction in ways that are designed to be more accessible to EL-classified students (Verplaetse & Migliacci, 2017).

While ELD is distinct from other core subjects, just as there is observed overlap in standards and skills taught across core content subjects such as English language arts (ELA) and social studies (Koedel, 2009; Yuan, 2015), instruction on the English language in an ELD classroom may overlap with content being taught in other classrooms. This disciplinary overlap may be particularly true for ELD and ELA, given the intersection of concepts and domains such as literacy, writing skills, vocabulary, and grammar (Callahan, 2006). However, despite this conceptual overlap, observed changes in EL-classified students' ELA performance year to year are typically attributed to their ELA teacher in both research and policy, without explicit considerations to the role that ELD teachers may play (Loeb et al., 2014; Master et al., 2016).

Because of the interrelated nature of ELA and ELD content, ELD teachers may, through variability in their instructional approaches and effectiveness, contribute to differences in EL-classified students' performance on standardized ELA assessments. While ELA teachers are positioned as the key party responsible for students' ELA outcomes in both research and accountability systems (Grossman et al., 2013; Loeb et al., 2014; Master et al., 2016), EL-classified students' ability to fully participate in their ELA classroom instruction may be shaped, in part, by the strength of instruction they receive in an ELD classroom (Callahan, 2006). Equitable access to learning opportunities is a

core legal right of EL-classified students, yet there is a wide body of research that identifies ways in which, even when enrolled in core content classrooms, instruction fails to be structured such that EL-classified students can meaningfully participate in their learning (Hansen-Thomas & Cavagnetto, 2010; Harper & de Jong, 2009). In such settings, EL-classified students' ELD instruction may be a critical support for accessing core content, as students may need to rely on ELD instruction to develop English proficiencies that can then support them in core content classrooms. Additionally, variability in ELD instructor effectiveness may have implications for EL-classified students' performance on standardized assessments, as English-administered content assessments conflate English language proficiency with core content skills and knowledge (Abedi, 2002).

While ELA teachers are tasked with supporting students' development of ELA skills and knowledge, there may be wide variation in the degree to which core content teachers engage in instructional practices that impact EL-classified students' ELP performance. For example, some teachers may view adapting their instruction for EL-classified students as outside their role as content area instructors. This may mean they pay little attention to opportunities to modify instruction to support EL-classified students' core access, relying instead on modifications that do not focus on language development explicitly, or drawing on misconceptions of the language development process in their instructional approach in ways that fail to meaningfully support students (Reeves, 2006). In contrast, some core content teachers may have training and awareness around the opportunities that they have in the core content classroom to support EL-classified students' language and core content skill development simultaneously and integrate this pedagogy into their instruction (Bunch, 2013; Zwiep et al., 2011). This variation in ELA teachers' instructional approaches and effectiveness may have significant implications for EL-classified students' ELA and ELP outcomes.

There is a push for research and practice that moves beyond the pedagogical framing of EL education as a process by which students' English language development happens in one classroom

and content learning in another, towards a framing that recognizes all educators' roles in supporting language and content learning (Bunch, 2013; DiCerbo et al., 2014; López et al., 2013; Reeves, 2006; Zwiep et al., 2011). Yet, there is limited quantitative evidence on the role that different teachers play in contributing to differences in EL-classified students' academic and linguistic outcomes (Jones et al., 2013). Addressing these gaps can inform our conceptual understanding of how different sources of instruction throughout the education process contribute to differences in measured academic and linguistic performance among EL-classified students. This may also highlight different ways in which systems can or should invest resources and attention to strengthen EL instruction and outcomes.

In this study, using statewide data for Oregon EL-classified students who are in grades 6-8 and enrolled in both an ELD and ELA class, I explored two research questions focused on understanding the extent to which variation is measured across ELA and ELD teachers in their contributions to EL-classified students' performance on ELA and ELP assessments.[22] I ask:

1. To what extent does variation in ELA and ELD teacher effectiveness contribute to differences in EL-students' ELA performance?
2. To what extent does variation in ELA and ELD teacher effectiveness contribute to differences in EL-students' ELP performance?

EL-classified students bring rich assets to their schooling and are building valuable skills such as multilingualism and multiculturalism, but there are significant, documented challenges in accessing meaningful learning opportunities for EL-classified students (Callahan & Shifrer, 2016). There is a rich body of work that focuses on identifying specific skills that support effective instruction for EL-classified students (Faltis et al., 2010; Gersten & Baker, 2000; Goldenberg, 2013), but there are gaps

[22] While these relationships and questions are of import among students in both lower and higher grades, middle school is a setting where typically the majority of EL-classified students are provided with a separate class period dedicated to instruction on the English language in addition to their core content instruction. Additionally, data are available to link students to individual teachers and populations of EL-classified students shrink as they progress through grades, making secondary school estimates potentially less stable and reliable.

in our understanding as to how teachers' contributions to EL-classified students' academic and linguistic outcomes vary, and whether that variability is observed across subject areas (Jones & Buziak, 2013). In this study, I explore these gaps, with findings that can inform conversations around how to leverage the important role of teachers to strengthen EL education. Importantly, there are critical challenges and limitations to the practice of quantifying teachers' effectiveness. The estimates in this study do not fully capture the range of skills and impacts that teachers bring to the classroom. Instead, they represent the degree to which individual teachers may contribute to differences in student performance, year to year. This contribution likely encompasses many different elements, which in this study are reduced down to the notion of teacher "effectiveness"—a blunt measure, but one that also may help to illuminate the ways in which differences in what happens inside classrooms contributes to higher or lower ELP and ELA outcomes for EL-classified students.

Theoretical Framework

There are a set of theoretical assumptions that guide the study focus, embedded within research on teachers and EL-classified students. The first theoretical assumption is that variation in ELA and ELD teachers' training, ability, resources, and awareness will lead to variation in both ELA and ELD teachers' effectiveness at supporting students' ability to develop and demonstrate their ELA skills and knowledge. Similarly, the second theoretical assumption is that this variation in ELA and ELD teachers' training, ability, resources, and awareness also has implications for EL-classified students' assessed ELP. In this section, I outline the research that supports these assumptions.

Teacher Contributions to Student ELA Performance

There are many different factors that shape students' academic performance year to year (Hanushek, 2020). For EL-classified students, these factors range from the individual, including student-level characteristics and experiences, to characteristics of the school and community environments (Eamon, 2005). For example, student English proficiency and academic engagement

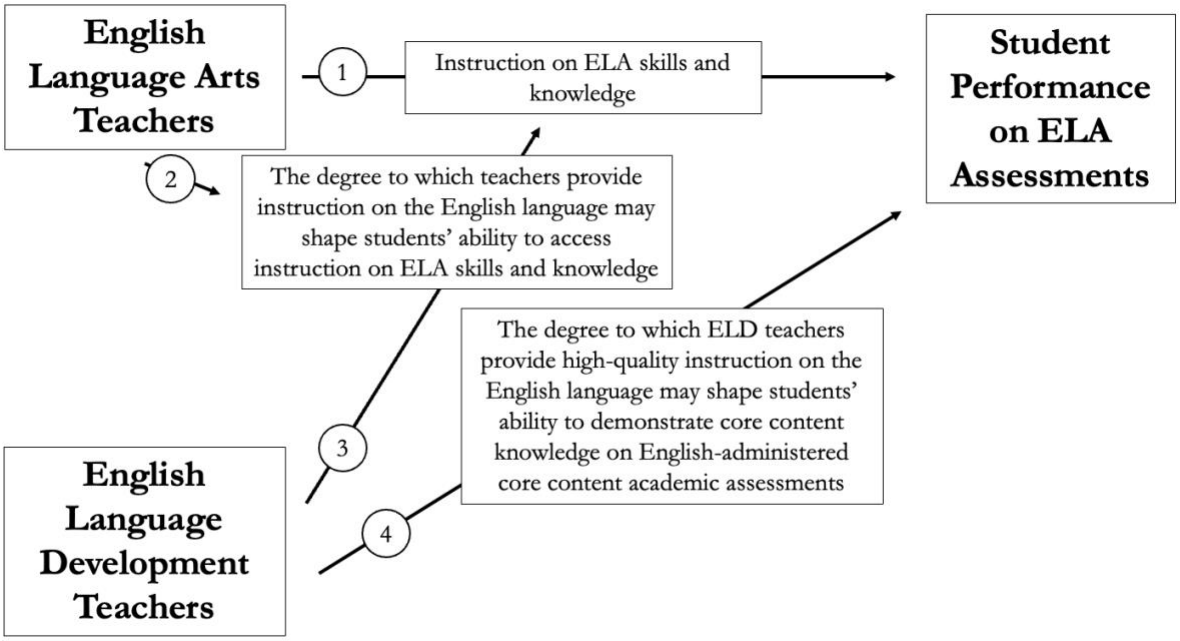
are associated with differences in academic growth, as are factors such as school segregation and poverty (Kieffer, 2008; Suárez-Orozco et al., 2010). EL-classified students' racial/ethnic and gender identity are also associated with different academic growth patterns (Polat et al., 2016). Beyond student, school, and community characteristics, teachers may also be important factors in EL-classified students' ELA performance.

Accounting for a host of factors (including prior assessed academic performance) significant variation has been documented in ELA teacher contributions to students' ELA performance among the broader population. For example, the standard deviation (*SD*) of Miami-Dade teacher contributions to non-EL students' reading performance ranged from 0.14 *SD* among elementary students to 0.09 *SD* among middle school students (Loeb et al., 2014). Similar estimates were found in Chicago Public High Schools (Aaronson et al., 2007), while slightly smaller estimates were found for high school students in San Diego Unified School District, 0.07 to 0.10 *SD*, as well as elementary and middle school students in North Carolina, 0.06 to 0.07 *SD* (Torre Gibney & Henry, 2020). These findings suggest that variation in ELA teacher effectiveness plays a significant role in students' academic outcomes. While these results represent outcomes for students overall, research on teacher contributions to academic outcomes for EL-classified students suggests that similar patterns hold. Specifically, Loeb and colleagues (2014) found that the variation in teacher effects on ELA performance for EL-classified students and non-EL students was similar across the two groups, differing no more than 0.01 *SD* across grades. Together, this work suggests that the ELA teacher a student is assigned to is associated with meaningful differences in student outcomes. However, these studies estimate ELA teacher contributions to student performance without explicitly considering how, for EL-classified students, instruction received through EL supports and services (such as ELD) may also be contributing to differences in student outcomes.

Figure 3.1 outlines the pathways through which theorized differences in teachers' instructional effectiveness may contribute to differences in EL-classified students' performance on standardized ELA assessments. First, as represented by Path 1, the quality of ELA instruction has the potential to impact EL-classified students' performance on standardized ELA assessments. This link is based in the research outlined above that identifies significant variation in ELA teacher contributions to students' ELA performance as measured by standardized ELA assessments, variation that has been found to be similar across EL-classified and non-EL-classified students (Loeb et al., 2014). Variation in effectiveness may stem from different levels of training and professional development that core content teachers receive focused on supporting EL-classified students in the classroom, as well as variability in awareness and ability to modify instruction appropriately (Gándara & Santibañez, 2016; Harper et al., 2011; Harper & de Jong, 2009). For example, some teachers may water down instruction in ways that constrict EL-classified students' access to rigorous content learning (Dabach, 2014; C. A. Harper & de Jong, 2009), while others may employ effective scaffolding techniques or other targeted instructional approaches that facilitate meaningful access to participation and learning in general core content classrooms (August et al., 2009; Llosa et al., 2016; Walqui & van Lier, 2010). There are specific competencies that may support secondary age EL-classified students in the classroom (Faltis et al., 2010); learning and employing these competencies may result in some ELA teachers' having larger, positive impacts on EL students' ELA performance or learning, while others who do not have access to sufficient training or support to do so may have little to no impact, or even be associated with lower performance outcomes (Santibañez & Gándara, 2018). Further, some ELA teachers may collaborate with EL specialists or teachers to adapt instruction in linguistically responsive ways or align instruction (Villavicencio et al., 2021), while others may see adaptations to EL instruction as "outside" their professional role (Pettit, 2011). Also seen in Figure 3.1, Path 2 connects ELA teachers indirectly to EL-classified students' ELA outcomes through instruction on the English language. This path

represents how some ELA teachers may integrate rich, embedded language pedagogy into their ELA instruction in ways that supports students' ELP development, which in turn can support core content learning (Bunch, 2013; Stoddart et al., 2002; Zwiep et al., 2011). These theorized relationships reflect the interrelated nature of language and literacy development (Callahan, 2006).

Figure 3.1 Theorized Pathways Through Which Variation in ELA and ELD Teacher Effectiveness may Shape EL-Classified Student Performance on ELA Assessments



Note. ELA=English language arts, ELD=English language development.

Figure 3.1 also includes ELD teachers as an input into EL-classified students' ELA performance in a given year, with two pathways identified. The first pathway (Path 3) through which ELD teachers may contribute to differences in EL-classified students' ELA performance connects variability in effectiveness of ELD instruction to EL-classified students' ability to participate in their ELA classrooms. In classrooms where English is the predominant or only language used and ELA teachers are not integrating effective instructional practices to support access for students who have varying levels of English proficiency, EL-classified students' core content learning and ELD likely become critically interrelated and interdependent (Gibbons, 2002; Lucas & Grinberg, 2008). For

example, if a student's ELD instruction is delivered in a way that supports linguistic growth and their sense of self-efficacy and linguistic identity (Snow & Katz, 2010) it is likely that the ELD instruction may support the student's ability to participate meaningfully in their ELA classroom. As noted above, this may become even more of an important source of instructional support if core content teachers are not structuring their classrooms such that learning opportunities are accessible or delivered with EL-classified students in mind. In such contexts, English proficiency may become a gatekeeper to accessing instruction (King & Scott, 2014). ELD instruction through a separate class period may support students' English language and literacy growth, suggesting that variability in the effectiveness in those classrooms may play a role in differential ELA outcomes for students (Saunders et al., 2006). There is also another pathway (Path 4), as variability in ELD instruction may also impact students' ability to demonstrate their ELA skills and knowledge on the standardized ELA assessments used. English-administered core content assessments are widely understood to conflate English proficiency with academic content (Abedi, 2002). As such, less effective ELD instruction may not support students' ability to demonstrate their ELA knowledge, while more effective ELD instruction may support students' ability to demonstrate their ELA knowledge.

Teacher Contributions to Student ELP Performance

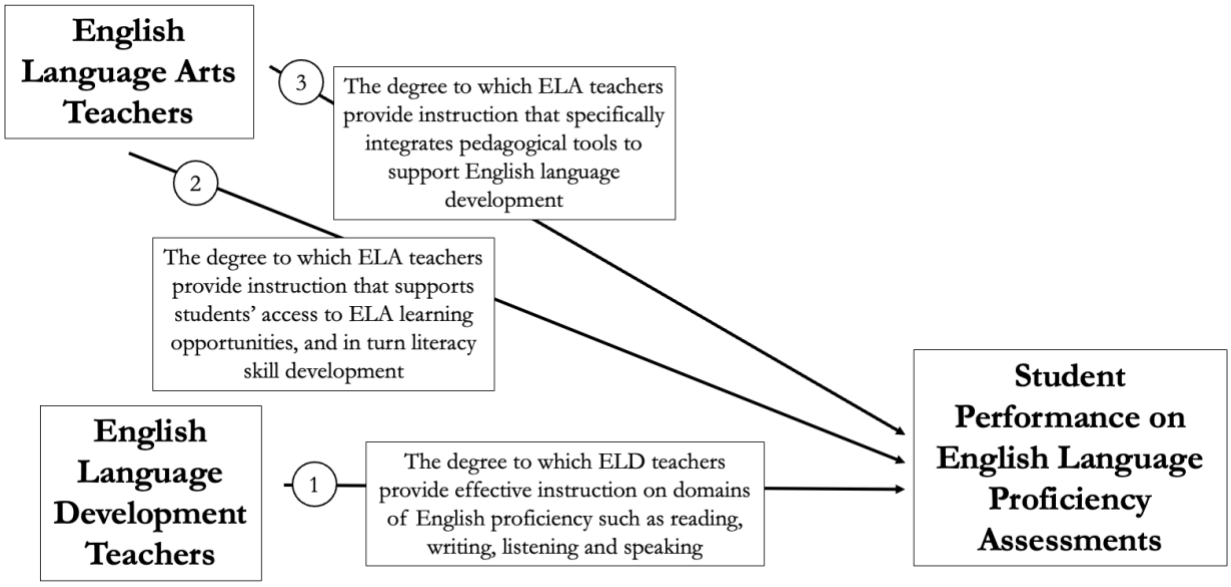
There are also many inputs that shape EL-classified students' assessed performance on standardized ELP assessments year to year. Many of these factors mirror those that shape academic growth. For example, students' socioeconomic status and race/ethnicity are predictive of students' ELP development, as are other factors such as initial home language literacy, gender, identification for special education services, and the school concentration of students in poverty and classified as EL (Kim et al., 2014; Thompson, 2017). While there is much less quantitative evidence on the role that teachers may play in shaping students' ELP performance than for ELA, I argue that prior research

supports the theory that variation in the effectiveness of both ELA and ELD teachers may shape EL-classified students' ELP performance.

Figure 3.2 outlines the different pathways through which differences in ELA and ELD teachers' instructional effectiveness may contribute to differences in EL-classified students' performance on standardized ELP assessments. The first pathway (Path 1) runs directly from ELD teachers to students' ELP performance, representing how differences in students' ELP performance may be attributable to their assigned ELD teacher through the degree to which that teacher engages in effective instruction on different domains of the English language. It is recommended that ELD instruction explicitly teach English language concepts such as vocabulary, syntax, morphology and conventions, as well as both conversational and more formal language, be interactive and incorporate opportunities for feedback, and be based on an understanding of language development and communication strategies (Saunders et al., 2013). The degree to which ELD teachers are provided with the supports and training to do so effectively for students may result in variability in their effectiveness, and in turn student performance. The second path (Path 2) runs from ELA teachers to students' ELP performance, representing the degree to which ELA teachers provide instruction that supports EL-classified students' opportunities to access and participate in ELA learning opportunities in the classroom. While ELA and ELD are separate subjects, the overlap across the skills and knowledge, as well as state standards (Oregon Department of Education [ODE], 2019), suggests that effective ELA instruction that has been appropriately modified or adapted for EL-classified students may also support students' reading, writing, listening and speaking skills, all of which are assessed on the ELP assessment. Relatedly, the degree to which ELA teachers are fostering meaningful dialogue and engagement in the classroom may support language development (Walqui, 2006). Path 3 also connects ELA teachers to students' ELP performance, based on the theory that differences in students' ELP performance may also depend on the degree to which ELA teachers integrate

pedagogical tools that support students' ELP growth into their instruction.

Figure 3.2. Theorized Pathways Through Which Variation in ELA and ELD Teacher Effectiveness may Shape EL-Classified Student Performance on ELP Assessments



Note. ELA=English language arts, ELD=English language development.

The theories that guide this paper can be summarized by the more simply put assertion of Bunch and colleagues (2012) that "Language development and cognitive development are interrelated and mutually dependent; ELs learn language as they learn content" (p. 2). While Bunch and colleagues are not explicitly discussing ELA and ELD, but rather core content and ELD broadly, the focus of this study is specifically on ELA and ELD given the interrelated nature of core content and language learning and the potential overlap in the instructional focus across ELA and ELD classrooms. This is exemplified in Oregon's ELA and ELP standards. The Oregon ELP standards for 6th-8th graders include standards such as, for example, "speak and write about grade- appropriate complex literary and informational texts and topics", "conduct research and evaluate and communicate findings to answer questions or solve problems", "determine the meaning of words and phrases in oral presentations and literary and informational text" and "make accurate use of standard English to communicate in grade-appropriate speech and writing" (ODE, 2019, pp. 23-24). There are clear

similarities and overlaps with some of the ELA and literacy standards, which include, for example, "Demonstrate command of the conventions of standard English grammar and usage when writing or speaking", "Use knowledge of language and its conventions when writing, speaking, reading, or listening", "Determine the meaning of words and phrases as they are used in a text, including figurative, connotative, and technical meanings; analyze the impact of a specific word choice on meaning and tone", and "Conduct short as well as more sustained research projects based on focused questions, demonstrating understanding of the subject under investigation" (ODE, 2019).

This is not to say that ELA and ELD classrooms will mirror one another. Rather, the instruction students receive in each classroom is likely not isolated to that classroom, but instead shapes how they engage across ELA and ELD classrooms with learning opportunities, as well as how they perform on the assessments given at the end of the year for both subjects.

Data

To answer the above research questions, I drew on statewide, teacher- and student-level administrative records from 2015/16 to 2018/19 for EL-classified students in grades 6-8. The sample that included students' ELA test score outcomes spans all four years, while the sample that included ELP test score outcomes spans three years (2015/16 to 2017/18).

To generate reliable, meaningful estimates of ELA and ELD contributions to students' performance that can be compared to one another, I took multiple steps to create a sample that included teachers who were linked to enough individual students who themselves were in both ELA and ELD classes. Using course enrollment data for all EL-classified students from 2015/16 to 2018/19, I undertook multiple steps to create individual records for each student, in each year, linked to an ELA and ELD teacher. First, I dropped all course-taking records where students' enrollment duration was less than 50% of the total enrollment period. I then dropped any class smaller than five students. I restricted the sample to EL-classified students who were enrolled in either an ELA or ELD

class. I also dropped students who were missing data across all years on eligibility for free/reduced price lunch, participation in migrant education, recent immigrant status, special education identification, gender, or race/ethnicity. For students missing fewer than all years on one or more of these variables, I imputed the modal value across their full series of data. For the ELA sample, I also dropped students missing an ELA score in the current year, while for the ELP sample I did the same for students missing an ELP score. I only retained ELA course-taking records for classes with NCES course codes that were non-elective or specialized ELA courses.[23] I retained all records with the NCES course code 1008 or 51008 for ELD coursework. I also dropped the small fraction of students whose ELA teacher was also their ELD teacher. If a student was connected to more than one ELA course/teacher combination after the above steps, I retained only their main, grade-specific ELA course while dropping the other (i.e., retain 6th grade ELA and drop Reading). After these steps, 13,145 student/year observations were still linked to either two ELA or two ELD teachers. I dropped these observations. For the main sample, I also excluded any student missing a prior year ELA, ELP Reading domain score, or math score. My analytic strategy relies on having the same student sample across both ELA and ELD teachers, meaning that students must be enrolled in both an ELA and ELD class.[24] Of the students remaining in the sample, almost 60% were enrolled in both an ELA and ELD class, while the remaining observations were dropped. Finally, to reduce concerns around instability of teacher effects (Chetty et al., 2014), I also drop any student-teacher record where the teacher is linked to fewer than ten EL-classified students. For ELA outcomes, this leaves a sample comprised of 9,513 student-year observations for 7,246 students, from 119 schools within 53 school districts. These students shared 388 ELA teachers and 211 ELD teachers. Student-level descriptive statistics

[23] ELA course codes that were included were: 1010, 1034, 1035, 1036, 1037, 1046, 1047, 1048, 1049, 1053, 1055, 1058, 1099, 1103, 1104, 1136, 1137, 1138, 1139, 1149, 1996, 51007, 51034, 51034, 51035, 51036, 51037, 51046, 51047, 51048, 51049, 51136, 51137, 51138, 51139.

[24] Having the same sample allows for comparisons in variability across ELA and ELD teachers (Koedel, 2009).

for this population are presented in Table 3.1. The student sample grows smaller as students progress through grades, reflecting the process by which students are reclassified from EL status to fluent English proficient. The majority of students were Hispanic/Latinx (around 80% in each grade). The proportion of students identified for special education ranges from 23.0% to 32.4%, increasing by grade. The majority of students were eligible for free/reduced price lunch (FRPL), over 90%. Students in the sample were, on average, in schools where about 19% of their peers were EL-classified and about 60% of their peers were students of color. Students in the ELA sample were in ELA classrooms where the average proportion of students who are EL-classified ranges from 23.5% in 8th grade to 30.4% in 6th grade, and the proportion of students who were nonwhite is higher than the school average. While it may be that some of these ELA classrooms are sheltered classrooms specifically designed to support EL-classified students, I am unable to identify which classes in the data. This is a limitation, as teachers of those classes may be more effective at increasing test scores among EL-classified students, both ELA and ELD, than those that are teaching non-sheltered classrooms.

Also presented in Table 3.1 are summary statistics for the student population that has non-missing scores on the ELP Reading domain, which the ELP domain score with the lowest missingness rate in the sample. This sample is comprised of 6,295 student-year observations for 5,020 students, from 106 schools within 47 school districts. These students shared 302 ELA teachers and 160 ELD teachers. The three other domains (Listening, Writing, and Speaking) have slightly fewer observations due to missingness, although summary statistics are virtually identical. Similar patterns observed among the students with non-missing ELA scores are evident in the smaller sample of students with non-missing ELP scores. However, ELD classroom-level proportions were different from their ELA classroom proportions. Specifically, on average students' ELD classroom peers were more likely to be eligible for special education services, immigrants, non-white, and FRPL eligible than their ELA classroom peers.

Table 3.1. Student Level Descriptive Statistics

	ELA			ELP		
	6	7	8	6	7	8
Male	0.556	0.571	0.583	0.555	0.561	0.597
Race/Ethnicity						
AIAN	0.017	0.045	0.060	0.023	0.048	0.064
Asian	0.060	0.072	0.085	0.063	0.069	0.099
Black	0.015	0.019	0.028	0.019	0.020	0.026
Hispanic/Latinx	0.849	0.802	0.767	0.836	0.804	0.742
White	0.056	0.058	0.057	0.057	0.056	0.065
Multi-Ethnic	0.003	0.004	0.002	0.001	0.004	0.004
SPED	0.230	0.270	0.324	0.232	0.293	0.351
FRPL Eligible	0.965	0.960	0.955	0.967	0.970	0.957
Home language						
Spanish	0.861	0.814	0.783	0.855	0.815	0.763
Migrant	0.133	0.138	0.142	0.139	0.136	0.146
ELPA21 Prof.						
1	0.012	0.044	0.057	0.012	0.046	0.057
2	0.783	0.887	0.896	0.787	0.883	0.896
3	0.205	0.069	0.047	0.201	0.071	0.047
(Missing)	1129	911	658	0	0	0
ELA	-1.095	-1.233	-1.254	-1.106	-1.286	-1.304
Math	-1.072	-1.185	-1.150	-1.106	-1.240	-1.206
ELP Reading	0.176	0.223	0.231	0.171	0.219	0.228
ELP Writing	0.223	0.262	0.276	0.218	0.255	0.271
ELP Listening	0.212	0.272	0.283	0.204	0.266	0.278
ELP Speaking	0.184	0.212	0.188	0.177	0.206	0.200
Prior Year ELA	-1.160	-1.313	-1.431	-1.152	-1.354	-1.491
Prior Year ELP Reading	-0.051	-0.050	0.000	-0.061	-0.043	-0.070
C. Prop. Sped	0.151	0.166	0.178	0.262	0.295	0.306
C. Prop. EL	0.304	0.264	0.235	0.996	0.994	0.995
C. Prop. Nonwhite	0.663	0.662	0.658	0.929	0.927	0.921
C. Prop. FRPL	0.822	0.819	0.814	0.953	0.951	0.949
S. Prop. Sped	0.160	0.159	0.157	0.160	0.159	0.156
S. Prop. EL	0.190	0.195	0.199	0.191	0.202	0.204
S. Prop. Nonwhite	0.606	0.610	0.611	0.606	0.618	0.611
S. Prop. FRPL	0.791	0.791	0.785	0.799	0.804	0.779
Total Students	4154	3271	2088	2831	2150	1314

Note. SPED=Identified for special education, AIAN=American Indian/Alaska Native, FPRL=Free/reduced price lunch, ELA=English language arts, C.=Classroom, S.=School, Prop.=Proportion, ELP=English language proficiency.

Summary statistics for the teacher population are presented in Table 3.2 by year. The majority of ELA and ELD teachers included in the study were female, with a slightly higher proportion of ELA teachers than ELD. The vast majority were White. For the ELA outcomes sample, about six to eight

percent of ELA teachers were identified as Hispanic/Latinx, while the proportion increases from around eight percent to about 14 percent by 2019 among ELD teachers. About one to three percent of ELA teachers reported Spanish as their primary language over the years observed, while about five to nine percent of ELD teachers reported Spanish as their primary language. Around 20 to 28 percent of ELA teachers held an active English for Speakers of Other Languages (ESOL) endorsement, while the proportion declined slightly over time among ELD teachers from about 91 percent to 76 percent. Similar patterns are observed among teachers included in the ELP outcomes sample. Of note, this is a selective sample of ELA and ELD teachers—not representative of the statewide population of ELA and ELD teachers who teach EL-classified students.

Table 3.2. Teacher Level Descriptive Statistics

	ELA Teachers				ELD Teachers			
	2016	**2017**	**2018**	**2019**	**2016**	**2017**	**2018**	**2019**
ELA Sample								
Female	0.811	0.817	0.763	0.779	0.732	0.693	0.711	0.709
Hispanic/Latinx	0.074	0.071	0.082	0.064	0.084	0.089	0.124	0.143
AIAN	0.033	0.037	0.027	0.020	0.053	0.041	0.066	0.079
Asian	0.024	0.022	0.016	0.003	0.034	0.036	0.023	0.056
Black	0.003	0.015	0.013	0.009	0.028	0.022	0.014	0.010
White	0.965	0.966	0.971	0.985	0.933	0.929	0.922	0.898
Spanish home language	0.030	0.017	0.027	0.012	0.056	0.076	0.096	0.097
ESOL Endorsed	0.284	0.195	0.255	0.206	0.911	0.862	0.775	0.760
Total	339	411	378	344	179	227	218	199

	2016	**2017**	**2018**	**2019**	**2016**	**2017**	**2018**	**2019**
ELP Sample								
Female	0.806	0.814	0.796		0.734	0.667	0.722	
Hispanic/Latinx	0.071	0.076	0.088		0.074	0.083	0.113	
AIAN	0.032	0.031	0.032		0.064	0.046	0.072	
Asian	0.025	0.028	0.021		0.032	0.037	0.031	
Black	0.004	0.019	0.021		0.021	0.019	0.021	
White	0.968	0.963	0.965		0.926	0.935	0.928	
Spanish home language	0.025	0.012	0.028		0.053	0.065	0.093	
ESOL Endorsed	0.286	0.226	0.277		0.904	0.852	0.825	
Total	281	318	280		94	109	97	

Note. ELA=English language arts, ELP=English language proficiency, ELD=English language development, AIAN=American Indian, Alaska Native, ESOL=English for speakers of other languages

ELA and ELP Outcomes

The two outcomes of interest in this study are students' scores on the standardized ELA assessment and students' scores from each standardized ELP assessment domain (reading, writing, speaking and listening). Both the ELA and ELP assessments are administered in the spring of each year, although the ELA assessment is administered to all students, while the ELP assessment is only administered to students who are currently EL-classified.

In Oregon, for the years examined, the standardized ELA assessment administered across the years of data was the Smarter Balanced Assessment Consortium (SBAC). SBAC scores range from 2000 to 3000. Scores were standardized within grade and year for the full EL-classified student population, with a mean of zero. As seen in Table 3.1, across grades six through eight students' ELA lagged and in-year scores were over a standard deviation lower than the full statewide population of students (EL and non-EL) who took the assessment in the same grade and year.

The ELP assessment used during the study window was the ELPA21 assessment. Oregon is a member of the ELPA21 consortium, a group of states that use the common ELPA21 assessment. The first year that Oregon transitioned to the ELPA21 was 2015/16,[25] the first year of data in this study. Scoring on the ELPA21 is based on a profile model, with students' English proficiency level being informed by their performance across domains, not a simple overall score (ODE, 2022). For this study, I analyzed teachers' unique contributions to students' standardized scores for each ELP domain—reading, writing, speaking, and listening—standardized within grade and year. As seen in Table 3.1, students' average ELP domain scores were above the average within-grade, within-year score for their EL-classified peers who took the assessment. This means that the sample is comprised of students who had slightly higher ELP scores than their grade-level EL-classified peers.

[25] Of note, lagged scores in 2015/16 are based off the prior ELP assessment (Oregon ELPA). I argue that these scores are still informative in the lagged-score models given that they are standardized within grade and year for the prior year, and therefore still capturing standardized information on students' ELP they bring into their current year.

Analytic Approach

The first research question is focused on estimating the variation in ELA and ELD teacher contributions to EL-classified students' ELA performance, accounting for a host of other factors, including assessed academic and English language performance from the end of the prior year. The second research question is similar but focuses on students' ELP performance as the outcome. The approach used in this study to estimate teacher contributions to students' outcomes is often called estimating a teacher value-added model, as it is framed as capturing the "value" an individual teacher contributes to students' performance, as well as capturing the overall variation in the "value-added" as a way to represent the degree to which it matters for students' outcomes whether a student is assigned to a more or less instructionally effective teacher. This approach is often framed as estimating how much the variation in teacher quality matters for student outcomes, as well as a way to identify more or less effective teachers (Aaronson et al., 2007; Hanushek, 2020; Koedel et al., 2015; Rivkin et al., 2005).

While I drew on methods from the value-added approach, I avoid using the term value-added because the term, by nominally tying teachers' value to student test scores, solely relies on standardized assessment scores as indicators of student learning, leaving out important other sources of information. The resulting values represent just one element of students' educational experiences and outcomes. I approach this study as an exploratory way of trying to understand the extent to which there is variation in teacher contributions to students' assessed content and ELP performance. The goal is not to identify "good" or "bad" teachers, nor hold up teachers of one subject over the other as more important. Rather, this study is designed, and hopefully executed, in a way that lends evidence to the discussion around the interrelated nature of language and content learning for EL-classified students and the role that teachers across classrooms play in shaping students' outcomes.

In this study, I fit a model that estimates a parameter for each individual ELA and ELD teacher a student is linked to in a given year in relation to their ELA score, as well as each ELP domain score, in addition to a host of covariates that include students' prior-year math, ELA, and ELP scores and a series of fixed effects: This model is conceptually similar to those that estimate teachers' contributions to students' academic performance that "spill over" from their assigned subject to students' performance on assessments measuring skills and knowledge from other subjects (Koedel, 2009; Yuan, 2015). Specifically, I fit the following mixed-effects model, estimated via restricted maximum likelihood:

$$SCORE_{it} = f(ELA_{i(t-1)}) + f(ELP\_READ_{i(t-1)}) + f(MATH_{i(t-1)}) + \boldsymbol{X}_{it} + \boldsymbol{Y}_{ct} + \boldsymbol{Z}_{st} + \sigma_g + \Gamma_t + \pi_s + (\varphi_{ela(i,t)} + \eta_{eld(i,t)} + \varepsilon_{it})$$

This model includes a random effect for each ELA teacher ($\varphi_{ela(i,t)}$) and each ELD teacher ($\eta_{eld(i,t)}$) to capture their unique contribution to students' test scores. The inclusion of teachers as random effects approximates a fixed-effect value-added modeling approach that applies Bayesian shrinkage to the resulting estimate to mitigate concerns around bias in the estimate stemming from yearly measurement error (Bitler et al., 2019; Kraft, 2019; Mulhern, 2020). This model can also be thought of as specifying a three-level error structure, in which there is an ELA teacher-level random effect, ELD teacher-level random effect, and an idiosyncratic student-level error term (Kraft, 2019).

In addition to ELA and ELD teacher random effects, I included grade (σ_g), school (π_s), and year (Γ_t) fixed effects. I also included cubic polynomials of students' prior year ELA, math, and ELP reading domain scores[26], as well as their current year demographic characteristics ($\boldsymbol{X}_{it}$; race/ethnicity, gender, FRPL eligibility, immigrant status, migrant status special education identification). I also included classroom- ($\boldsymbol{Y}_{ct}$) and school-level ($\boldsymbol{Z}_{st}$) averages of student demographics. In this approach,

[26] I lag students' ELP reading score because it has the lowest missingness level across ELP domain scores, and it also represents the domain I argue would most likely predict students' course placement, given the important role that reading plays across classrooms.

teachers' unique contributions to test score performance are estimated from differences in students' residualized test scores across different "pairings" of ELA and ELD teachers within schools. In an alternative main specification, I dropped school fixed effects and specify a four-level error structure, including school random effects rather than fixed.

The second research question mirrors the first, but instead of estimating the unique contribution of each ELA and ELD teacher to students' ELA assessment performance, I fit a series of models, each with students' ELP domain scores as the outcome of interest. The models are identical to that of the ELA outcome, with the only difference being a smaller subset of students.

The parameter of interest for both research questions is the standard deviation (*SD*) of the distribution of teacher random effects, a measure of variation in teacher contributions to student performance as captured through the remaining variation (residuals) in student test scores after the inclusion of the above covariates and fixed effects that is then attributed to a student's teacher. The significance and magnitude of this variation provides a way to frame whether the explained variance is approximately statistically distinguishable from zero, as well as how much variation in teachers' contributions is observed across the full distribution of ELA and ELD teachers' estimated contributions. For example, the *SD* of ELA teachers' contributions to students' performance on ELA assessments provides information on how much higher, on average, students' ELA test scores are when assigned to a teacher whose effectiveness is one *SD* above or below the average teacher effectiveness. This can be examined alongside the *SD* of ELD teachers' contributions to students' ELA performance, as well as the *SD* of ELA and ELD teacher contributions generated with ELP performance measures as the outcome of interest. As such, the resulting parameters provide evidence on the extent to which there is variation in ELA and ELD teachers' measured instructional effectiveness that has meaningful implications for student performance on ELA and ELP assessments.

Assumptions and Limitations

One assumption that underlies this approach is that teacher effects are additive across subjects, and therefore one can identify the unique contributions of each teacher, across subjects, to student performance through including both in the model (Koedel, 2009). This is similar to approaches used for co-teaching that rely on weights to identify the partial effect for teachers, except in this approach all weights are equal, approximating a situation in which each "co-teacher" plays an equal role in student learning (Hock & Isenberg, 2017).

Another set of assumptions is related to the variation from which the estimated contributions to student performance are estimated. Teachers' random effect estimates are generated from mean differences in students' residualized test scores across different teacher pairings. As such, for the estimates to be unbiased, the approach assumes that there is no "sorting" when it comes to ELA and ELD teacher pairing, conditional on the included fixed effects and covariates. A related assumption is that, conditional on included covariates, there is no sorting of students to teachers. Given that middle school EL-classified students' ELP levels and academic preparation inform and predict course access (Callahan, 2005; Estrada, 2014), I argue that the included covariates are well-situated to address sorting concerns. However, omitted variable bias and resulting sorting, including parental preference for certain teachers, or building leaders' preference in assigning students to specific teachers or combinations of ELA and ELD teachers cannot be fully ruled out and should be considered as a limitation to the interpretation of these findings.

There are also other limitations of note. First, given the structure of the course-taking data I have access to, my sample is limited to ELD teachers who provide ELD instruction through a separate class period approach. This varies from ELD teachers or specialists who provide English language support by pulling students out of class to receive supports or "pushing-in" to the classroom to support content learning through ELD instruction. This sample is also limited to ELA teachers who are linked to a sufficient number of EL-classified students who are concurrently enrolled in an ELA

class. While this limits the generalizability of this study to only ELD teachers who provide ELD through a separate class period, and to ELA teachers who provide content instruction in a school that has designated class period ELD, this likely represents a large proportion of schools. In 2016, 68% percent of surveyed U.S. high schools that had EL-classified students reported that they provided ELD through a designated class period (Lewis & Gray, 2016).

As another limitation, I do not have data on is whether there are other supports that ELA teachers are receiving in the classroom. For example, paraeducators or EL specialists may support EL-classified students in their ELA classroom. The additional supports may introduce bias into teachers' estimated contribution to students' performance. Additionally, some courses may be explicitly designed to support EL-classified students, and teachers in those classrooms may be particularly effective, with more training and supports. Finally, this is a very small sample of teachers and students, with a short panel of data. In addition to the other limitations noted above, this may introduce concerns of instability in the estimates.

In light of concerns regarding unobserved variable bias and sorting concerns, as well as a relatively small population with potentially limited variation in pairings across teachers, I frame this study as exploratory, with important results that should be interpreted within the context of these limitations. The results would benefit from being compared to similar analyses run with a longer panel and larger student and teacher population.

Sensitivity Analyses

In response to the limitations and concerns outlined above, I ran a series of sensitivity analyses for the above two research questions. Some of the sensitivity analyses involve different sample constructions, while others different model specifications.

For both research questions, I re-ran the models dropping all students identified as receiving special education services. One reason I ran this specification is, as with EL education, special

education services often come with additional instructional supports and services that may be contributing to variability in teacher estimated contributions. Excluding these students may help to get a cleaner estimate of teachers' contributions. Additionally, students who are eligible for special education, as with EL-classified students, may have particularly high rates of measurement error in their test scores outcomes, in comparison with the non-EL population without disabilities (Buzick & Laitusis, 2010). Measurement error, even with the random effects approach, may still be a concern for estimates' stability and validity.

I also re-ran the main specification with a larger sample, imputing prior year math, ELA, and ELP reading scores to the mean value for their given grade and year (Lipscomb et al., 2010). This sample is larger and includes more students and teachers, potentially strengthening the findings if they align with the main estimates. I also re-ran the analyses using a subset of the imputed dataset limited to students who are taught by ELA teachers who are linked to two or more ELD teachers within a given grade and year as well as ELD teachers who are also linked to two or more ELA teachers within a given grade and year. This is essentially reducing the variation to contexts where there is variation in teacher pairings within grade and year. While this will likely be a much smaller data set, similar results will provide further evidence for the main estimates, while divergent estimates may suggest that the findings are less reliable.

As alternative specifications, I first re-ran the model with the main sample but did not include school fixed effects, allowing for the estimate to draw on cross-school variation (Koedel et al., 2015). I also re-ran the model specification with the main sample, but instead of modeling ELA and ELD teacher parameters as random effects, I included them as fixed effects, and do not adjust them using a shrinkage procedure. While the random effects approach adjusts the estimates in a way that addresses key issues of measurement error and instability (e.g. Kraft, 2019), the resulting estimates may draw on multiple sources of variation without a clear way to disentangle the source of such variation. A fixed

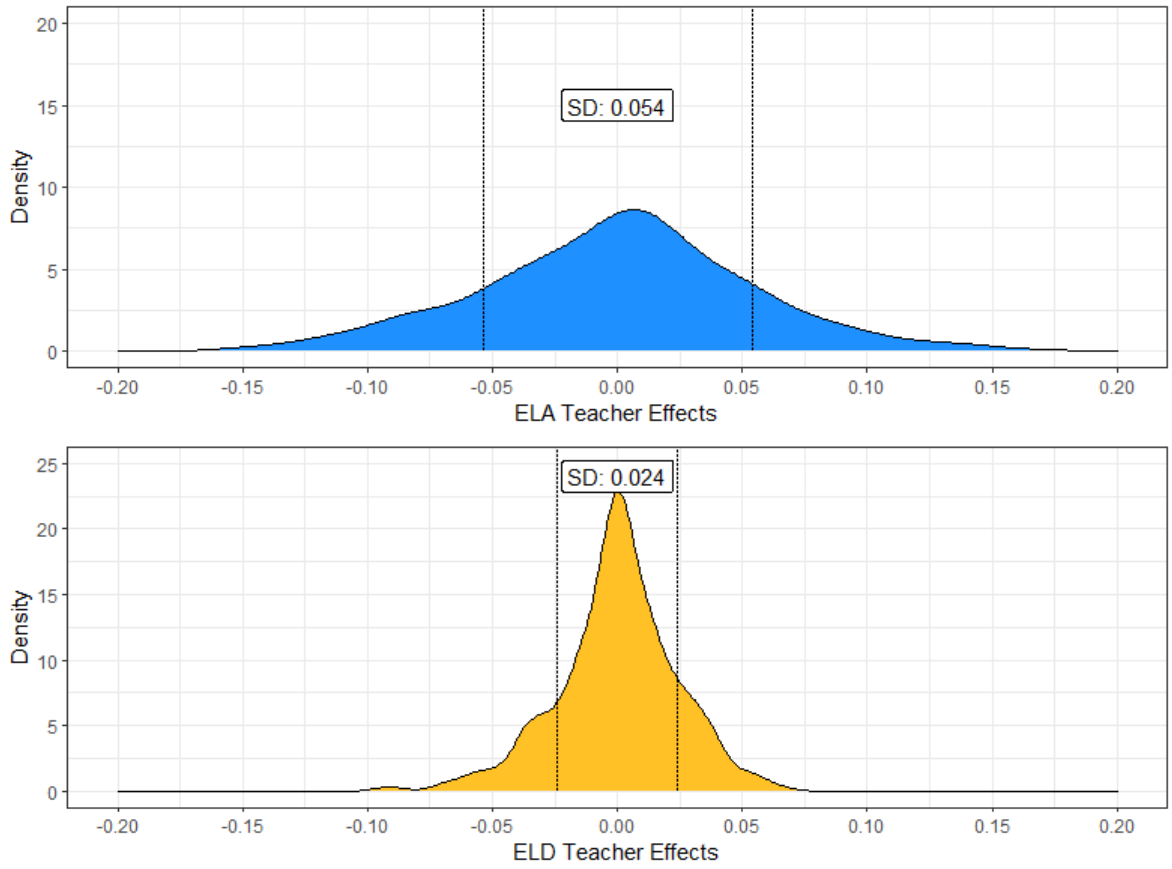
effects approach, while likely representing much noisier and often outsized estimates, may more cleanly capture overall variation in teacher contributions to student performance. In the fixed effects approach I included the same covariates as above and grade and year fixed effects (dropping school fixed effects because of collinearity concerns with the teacher parameters), reducing the variation to within grade and year. While likely much larger and noisier than the main estimates (Bitler et al., 2019), the distribution of these estimates can be compared to the distribution of main estimated contributions to see whether they appear to be similar. If the distribution is similar, that is evidence that the variation identified through the random effects approach in teacher contributions to student performance represents variation in teacher effectiveness overall, rather than drawing on year to year variation in individual teachers' effectiveness.

Results

ELA and ELD Teacher Contributions to Students' ELA Performance

In the main model specification, there was meaningful variation in both ELA and ELD teacher contributions to students' ELA performance. The full distribution of ELA teacher contributions to EL-classified students' ELA performance, as estimated in the main model, is plotted in the top panel of Figure 3.3, while the *SD* of teacher contributions from a series of model specifications is presented in Table 3.3. In the preferred main specification (Model III for ELA outcomes in Table 3.3), students who were assigned to an ELA teacher whose measured contribution to ELA performance was one *SD* above the mean ELA teacher contribution had, on average, ELA scores that were 0.054 *SD* higher. This estimate was similar in magnitude to estimates from a model where schools were included as random effects (Model IV), and in the model with no ELD teacher parameter included (Model I). Including the ELD teacher parameter does not seem to change conclusions about the extent to which variation in ELA teachers' instructional effectiveness shapes student ELA performance, suggesting that the ELD teacher estimate is being "picked up" through other parameters estimated in Model I.

Figure 3.3. Distribution of ELA and ELD Teacher Contributions to EL-Classified Students' ELA Outcomes



Variation in ELD teachers' effectiveness, as measured by contributions to EL-classified students' ELA performance, also appears to have implications for EL-classified students' ELA performance, although the variation is smaller in magnitude than for ELA teachers. Reported in Table 3.3, Model III for ELA outcomes (and plotted in the bottom panel of Figure 3.3), students assigned to an ELD teacher whose measured contribution to ELA performance was one *SD* above the mean had ELA assessment scores that were, on average, 0.024 *SD* higher. This is about half the magnitude of ELA teacher estimates. The estimated *SD* is substantively identical in the school random effects approach, and slightly smaller than the model that did not include an ELA teacher parameter.

Table 3.3. Estimated ELA and ELD Teacher Contributions to EL-Classified Students' ELA Outcomes

	ELA			
	Model I	**Model II**	**Model III**	**Model IV**
Teacher Intercepts (SD)				
ELA	0.058***	/	0.054***	0.058***
ELD	/	0.031***	0.024***	0.024***
Covariates				
Student Covariates	X	X	X	X
Classroom Covariates	X	X	X	X
School Covariates	X	X	X	X
Random Intercepts				
School				X
Fixed Effects				
Grade	X	X	X	X
Year	X	X	X	X
School	X	X	X	
N (student/year obs.)	9,513	9,513	9,513	9,513
N (ELA Teacher)	388	/	388	388
N (ELD Teacher)	/	211	211	211

Note. *p<.05, **p<.01, ***p<.001. All models include student covariates of prior year ELA, math, and standardized ELP scores, as well as FRPL, special education, migrant, male, ethnicity/race, as well as classroom- and school-level averages of student covariates (other than lagged scores). Models I, II and III include school, grade, and year fixed effects. Model I includes an ELA teacher random effect, while Model II includes just an ELD teacher random effect. Model III includes both. Model IV includes both an ELA and ELD teacher random effect, but instead of a school fixed effect, the school is modeled as a random effect.

ELD and ELA Teacher Contributions to Student ELP Performance

When estimating ELD and ELA teacher contributions to students' listening, reading, writing, and speaking ELP domain scores as the outcomes of interest, there was significant variation in ELD teachers' instructional effectiveness across all four domains, while there was less variation in ELA teachers' instructional effectiveness across all four domains, with one instance where the variation was not statistically distinguishable from zero. For ELD teachers, there were differences in the variation across domains. Results are reported in Table 3.4, while the full distribution of ELD teacher contributions are plotted in Figure 3.4, and ELA teacher contributions are plotted in Figure 3.5.

The largest *SD* of ELD teacher contributions to EL-classified students' ELP domain-specific scores was estimated for speaking scores. In the main specification, students who were assigned to an ELD teacher whose contribution to students' speaking scores was one *SD* above the mean had, on average, speaking scores that were 0.095 *SD* higher. The *SD* of ELD teacher contributions to student

performance was similar when the outcome of interest was reading scores (*SD*=0.067) and listening scores (*SD*=0.055). In contrast, there was less variation in ELD teacher contributions to students' ELP writing performance, with a *SD* of 0.039. Together, this evidence suggests that variability in ELD teacher effectiveness had significant implications for students' ELP performance. These implications were particularly outsized for students' speaking performance, while ELD teachers' instructional variability seemed to have less of an impact on students' writing performance.

The estimated *SD* of ELA teacher contributions to students' performance on the different ELP assessment domains tended to be similar to one another and were much smaller than for students' ELA performance. For listening, speaking, and reading performance, the estimated *SDs* of ELA teacher contributions to student performance were very similar and small, 0.022, 0.027, and 0.028, respectively. There was the least amount of variation documented in ELA teachers' instructional effectiveness for writing performance, as the *SD* of 0.012 was not statistically different from zero. These estimates suggest that variability in ELA teacher effectiveness may have implications for EL-classified students' ELP performance, although changes in student performance are relatively small across all domains, and not significant for writing.

In sum, for the different ELP domain outcomes there were interesting patterns that suggest variation in ELD teachers' instructional effectiveness may matter more for students' performance on certain domains, with the variation for speaking performance being much larger than for writing, and moderate variation for reading and listening performance. Variation in ELA teachers' instructional effectiveness appears to have less of an implication for students' ELP performance across domains, with estimates that are statistically distinguishable from zero when looking at students' listening, reading, and speaking ELP domain scores, but small in magnitude.

Figure 3.4. Distribution of ELD Teacher Contributions to EL-Classified Students' ELP Domain Outcomes

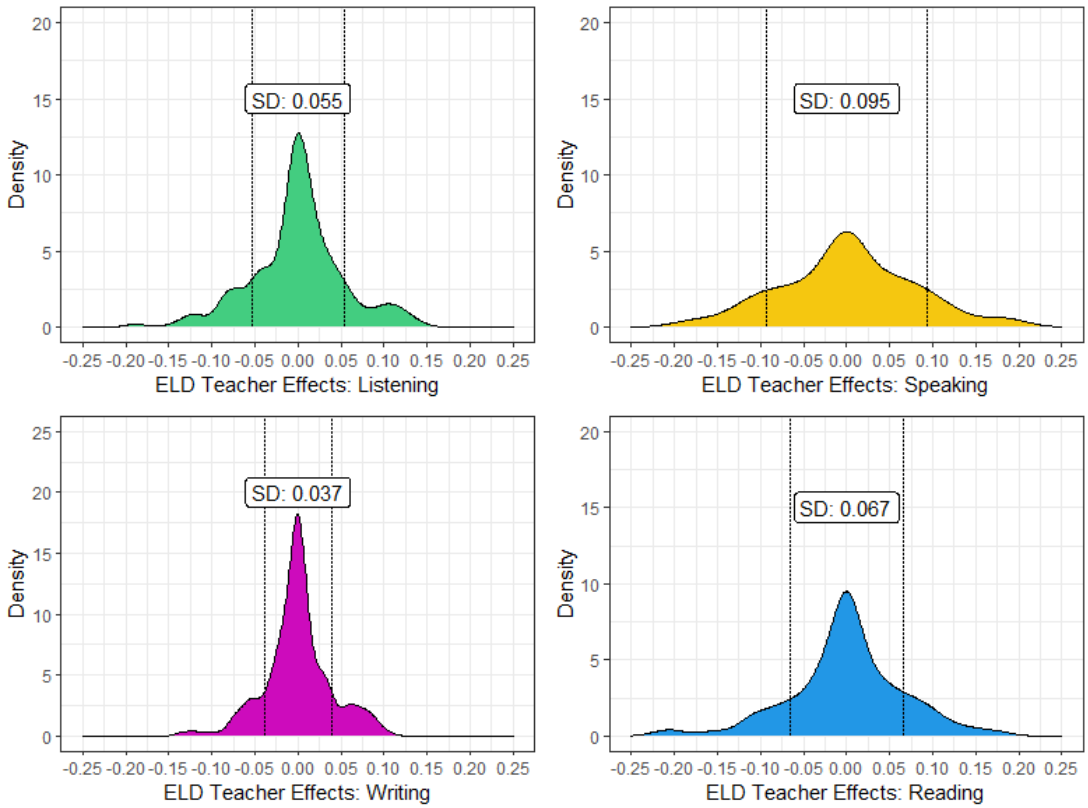


Figure 3.5. Distribution of ELA Teacher Contributions to EL-Classified Students' ELP Domain Outcomes

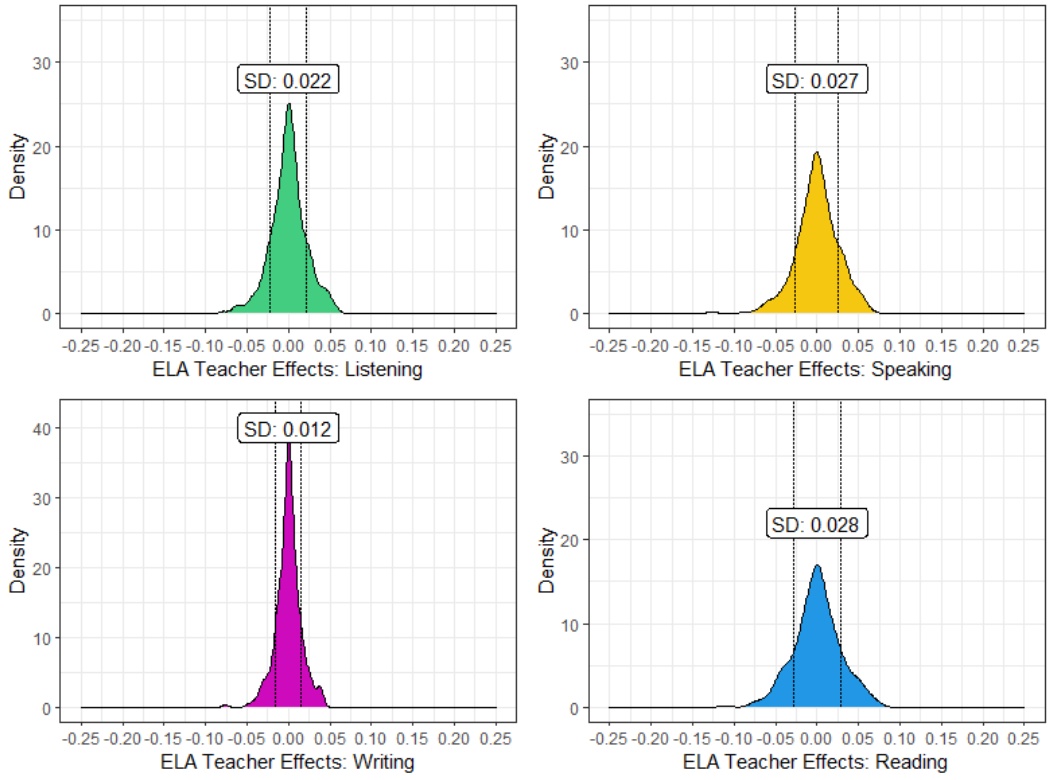


Table 3.4. Estimated ELA and ELD Teacher Contributions to EL-Classified Students' ELP Domain Outcomes

	ELP: Listening				ELP: Speaking				ELP: Writing				ELP: Reading			
	I	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**
Tch Int. (SD)																
ELA	0.032***	/	0.022**	0.025**	0.039**	/	0.027*	0.029*	0.021*	/	0.012	0.012~	0.034***	/	0.028**	0.032**
ELD	/	0.057***	0.055***	0.067***	/	0.096***	0.095***	0.110***	/	0.038***	0.037***	0.045***	/	0.066***	0.067***	0.074***
Cov.																
Stu. Cov.	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Class. Cov.	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Sch. Cov.	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
RE																
Sch.				X				X				X				X
FE																
Grade	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Year	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Sch.	X	X	X		X	X	X		X	X	X		X	X	X	
N (obs.)	6,293	6,293	6,293	6,293	6,271	6,271	6,271	6,271	6,277	6,277	6,277	6,277	6,295	6,295	6,295	6,295
N (ELA Teacher)	302	/	302	302	302	/	302	302	302	/	302	302	302	/	302	302
N (ELD Teacher)	/	160	160	160	/	160	160	160	/	160	160	160	/	160	160	160

Note. *p<.05, **p<.01, ***p<.001. All models include student covariates of prior year ELA, math, and standardized ELP scores, as well as FRPL, special education, migrant, male, ethnicity/race, as well as classroom- and school-level averages of student covariates (other than lagged scores). ELA=English language arts, ELP=English language proficiency, ELD=English language development Cov.=Covariates. Models I, II and III include school, grade, and year fixed effects. Model I includes an ELA teacher random effect, while Model II includes just an ELD teacher random effect. Model III includes both. Model IV includes both an ELA and ELD teacher random effect, but instead of a school fixed effect, the school is modeled as a random effect.

Sensitivity Analyses: ELA Outcomes

I conducted a series of sensitivity analyses to see whether the estimated variation in teacher contributions to students' ELA performance were similar across varying samples and model specifications. Results from five alternative approaches are reported in Table A1.

The first three sensitivity analyses were alternative samples that are used in the main model specification. The estimated *SD* was similar for ELA teachers when excluding all students who are identified for special education (Alt. 1), although the ELD teacher *SD* is smaller and no longer statistically distinguishable from zero at the 0.05 level. Estimates from the imputed sample (Alt. 2) were slightly different in magnitude than the main estimates, as the ELA teacher *SD* was slightly smaller, and the ELD teacher *SD* slightly larger. The third alternative sample (Alt. 3) is very small, with just 118 ELA teachers and 70 ELD teachers. The ELA teacher *SD* shrunk to 0.040, and the ELD teacher *SD* (0.011) is no longer statistically significant. Together, these alternative samples suggest that ELA teacher estimates held across specifications, although there was less variation with both the smaller and larger samples. Estimates of variation in ELD teachers' contributions to ELA performance shrank quite a bit with smaller samples.

In two additional specifications I used the main sample but fit two different model specifications. The first (Alt. 4) did not adjust for school-level nesting. The estimated *SDs* of ELA and ELD teacher contributions to student ELA performance were slightly larger than the main estimates. The second was a fixed effects approach (Alt. 5), where I model teacher parameters as fixed, rather than random, effects. The estimates were much larger for both ELA teachers and ELD teachers. As seen in the top panel of Figure A1, the distribution of ELA teacher contributions to students' ELA performance is relatively similar in shape to the overall estimate, despite the estimates themselves ranging to much larger negative and positive values than the main random effect estimate. As seen in the bottom panel of Figure A1, the distribution of ELD teacher contributions to students' ELA

performance was also similar to the main random effect estimates. However, these estimates are also much larger in magnitude and there are considerably more noisy estimates that pull out the tails of the distribution. These fixed effects results highlight how the approach may result in much larger coefficients on teacher contributions, potentially driven by large changes in student performance year to year for a smaller number of students. The similar distribution provides evidence that the overall findings are capturing variation in teacher effectiveness broadly, not just identifying off individual teachers' contributions varying within themselves year to year.

Sensitivity Analyses: ELP Outcomes

As with the ELA outcomes, I conducted a set of five different sensitivity analyses for each ELP domain outcome, three with alternative samples and two with the main sample, but in an alternative model specification. Results are reported in Tables A2 and A3. As with ELA outcomes, the analysis that excluded students identified for special education (Alt. 1) uniformly resulted in a smaller *SD* for ELD teachers, and for ELA teachers the *SD* in contributions shrunk to estimates that were not statistically significant. The imputed sample (Alt. 2) estimates were generally similar to the main estimates for both ELA and ELD teachers. Reducing the imputed sample to just students who were taught by ELA teachers who are linked to two or more ELD teachers within a given grade and year as well as ELD teachers who are also linked to two or more ELA teachers within a given grade and year (Alt. 3) resulted in similar estimates of the variation in ELD teacher contributions to students' domain performance, although estimates tended to shrink towards the mean. For ELA teachers the *SD* was very small in magnitude and the variation was not statistically distinguishable from zero.

Results from the model specification that did not adjust for school-level nesting (Alt. 4) were similar to the main estimates, although there tended to be more observed variation in ELD and ELA teacher contributions when not adjusting for school-level clustering. As with ELA outcomes, the fixed effect specification (Alt. 5) results in an estimated *SD* across outcomes that was much larger than the

random effects approach. Additionally, the *SD* of ELD teacher contributions was no longer much larger than ELA teachers, with estimates that were similar or with a larger *SD* of ELA teacher contributions. This is a departure from the main model, although given concerns about the already small sample size in the study and measurement error concerns that come with unadjusted fixed effect specifications (Bitler et al., 2019), these estimates are likely influenced by noisy outcomes and, while a useful check, do not overly influence my interpretation of the main results.

Across sensitivity checks other than the approach using unadjusted fixed effects, the variation in contributions of ELA teachers to ELA performance and ELD teachers to ELP performance (within-subject contributions), stayed relatively constant. Cross-subject contributions (i.e., ELA teachers and ELP Reading outcomes, ELD teachers and ELA outcomes) in the smaller samples did not consistently hold, with the *SD* shrinking and, in many specifications, becoming statistically indistinguishable from zero. These sensitivity analyses give room for pause. This was an already relatively small sample. Reducing the sample further may result in challenges in identifying teacher contributions. Additionally, smaller samples may introduce more noise and less stability, which are already concerns with the study design. As such, this is further evidence that some findings may be more robust than others, with the cross-subject contributions being more tentative than within-subject. Further work could strengthen these conclusions through a longer panel and larger sample.

Discussion

This study is a preliminary examination of the variation in ELA and ELD teachers' contributions to EL-classified students' performance on ELA and ELP assessments. The first research question focuses on quantifying the extent to which variation in ELA and ELD teacher effectiveness contributes to differences in EL-classified students' ELA performance. The second question is similar, but instead of ELA performance, focuses on ELP performance. I found that variation in both ELA and ELD teachers' instructional effectiveness contributed to differences in EL-classified students'

ELA and ELP performance, although variability in teacher effectiveness had larger implications for student performance within, rather than across, the two subjects.

Models that include both ELA and ELD teacher parameters suggested that ELA teachers' instructional variability has more of an implication for EL-classified students' ELA performance than ELD teachers' instructional variability, although both had measurable impacts. This finding lends validity to the modeling approach and problematizes current theory and modeling practices that conceptualize ELA teachers as solely responsible for students' ELA performance. It also provides more evidence to support the claim that core content teachers play an important role in EL-classified students' content outcomes (Loeb et al., 2012). These results suggest that there are ELA teachers who bring strong instructional skills to the classroom that are supporting EL-classified students' performance on ELA assessments, while other students may be assigned to teachers who either are less effective or less attentive to adapting ELA instruction for EL-classified students or providing instruction that supports both ELA and ELP skill development in ways that translate to improved ELA assessment performance.

A novel, preliminary finding in this study is that variation in ELD teachers' effectiveness also had implications for students' ELA performance. While the magnitude is smaller than ELA teachers and in smaller samples not always distinguishable from zero, I found that ELD teachers' contributions to EL-classified students' ELA performance were non-trivial in the main specification. Teacher spillover effects have been documented among non-EL classified student populations, with findings that variation in math teacher effectiveness has implications for students' ELA performance (Aaronson et al., 2007; Koedel, 2009; Yuan, 2015), as well as findings that variation in ELA teacher effectiveness has implications for students' math and social studies performance (Yuan, 2015). However, cross-subject contributions have not been explored for ELA and ELD.

Thinking through this cross-classroom contribution, this may be because some ELD teachers are providing instruction that supports students' ability to demonstrate their ELA knowledge on the standardized assessment while others are not (Figure 3.1, Path 4), or some are providing instruction and supports that help students access and excel in ELA core instruction, while and others are not (Figure 3.1, Path 3). This adds further weight to calls for instruction across ELA and ELD classrooms to move from siloed educational experiences to consider the ways in which students' instruction across the classrooms may overlap to support students' growth and performance (Bunch et al., 2012).

Just as there was more estimated variation in ELA teacher contributions to students' ELA performance, variation in ELD teacher effectiveness appears to play a larger role in shaping EL-classified students' ELP performance than ELA teachers. Importantly, however, the specification where ELD teachers are included as fixed, rather than random effects, introduces evidence that the main specification may be picking up on variation both within teachers (e.g. year to year) and across teachers (e.g. some ELD teachers are more effective at improving students' ELP performance in the data than others). Thinking through both sources of variation may point even more towards putting in place strong systems of consistent support such that individual teachers' effectiveness is not changing year to year with changing classrooms and contexts, and that teachers are supported in developing their instructional skills and abilities.

ELA teachers' contributions to students' ELP performance tended to be significant in the main specification, although not for all domains. As perhaps expected, the variation was smaller in magnitude. However, these findings suggest that ELA teachers, through their instruction, may play a role in supporting EL-classified students' ELP growth. As with the observed contributions of ELD teachers to students' ELA performance, these cross-subject contributions open up conversations around the interrelated nature of content and language development, especially across ELA and ELD classrooms, which focus on overlapping literacy and language concepts (Callahan, 2006).

Implications

A key takeaway from this study is that, as many studies have found, teachers matter for student performance. One contribution of this study is that modeling teacher contributions across ELA and ELD classrooms to students' ELA and ELP outcomes shows how, for EL-classified students concurrently enrolled in an ELA and ELD classroom, the quality of instruction in one classroom may have implications for their performance in the other. While this may not be surprising, especially given the overlap in standards for ELA and ELD (ODE, 2022), oftentimes policy investments and supports tend to focus acutely on single classrooms, subjects, or teachers, and fail to step back and think critically about how the full set of instructional inputs an EL-classified student is receiving can be strengthened to ensure coherent instruction. This study can inform conversations around the importance of investing in supports focused on the interrelated nature of language and core content learning for EL-classified students. For schools and districts interested in improving EL-classified students' ELA performance, providing supports to strengthen instruction in both ELA and ELD classrooms, while also continuing to support teachers who attend to language and literacy development holistically, may reduce gaps in teacher effectiveness that contribute to some students having higher test score outcomes than others.

EL-classified students' ELP growth is a key measure of school outcomes in state accountability systems (ESSA, 2016). Policy debates and framings tend to focus on the decision of which program model ELD should be delivered through, with less attention towards understanding the training, supports, and effectiveness of ELD teachers. For EL-classified students in middle school, many of whom have been classified as EL since kindergarten enrollment, acquiring the level of ELP necessary to be reclassified to fluent English proficient is an important goal. Reclassification has been found to have positive effects on student outcomes when it occurs earlier in their academic trajectory (Pope, 2016). To the extent that EL classification forecloses learning opportunities in middle school

(Umansky, 2016), investing in supports that can improve students' ELP outcomes may be an important focal area for states, schools, and districts. This study, while preliminary, provides evidence to inform the investment in stronger supports for ELD and ELA teachers to strengthen the instruction of those who may be less effective.

Differences in teacher variation across ELP domains highlight how variation in both ELA and ELD teacher effectiveness seem to be particularly important for students' English language speaking performance. It may be of interest to further explore and understand the instructional techniques employed in ELD and ELA classrooms that support the development of English oral skills and provide supports for teachers who may struggle to support students' English oral skill development to address what appears to be a wide-ranging effectiveness across both ELA and ELD teachers, but ELD teachers in particular.

Beyond investing in individual teachers, cross-subject contributions highlight the potential for collaboration and alignment across content and language development classrooms—which has been documented as a strong practice (Villavicencio et al., 2021). If both ELA and ELD teachers are contributing to the development of skills and knowledge that are assessed on ELA and ELP assessments, there may be clear pathways through which to collaborate and create instructional environments for students that allow for authentic, embedded literacy development. Importantly, test scores are a single, imperfect measure of students' educational experiences and quality of educational offerings. Conversations around strengthening instruction both within and across classrooms should not just focus on test score improvement but look at test score performance alongside other important indicators of instructional quality for EL-classified students.

Finally, while some uses of value-added modeling have been to identify weaker and stronger teachers with high-stakes implications, this study does not provide evidence to support this type of application or action. Such practices can be inequitable, rely on standardized assessments as singular

measures of effectiveness, and reduce instruction to a dimension that cannot capture what it means to be a "good" teacher (Paige & Amrein-Beardsley, 2020). Additionally, this study provides no evidence on whether the provision of ELD itself is associated with higher or lower test scores, only documenting the variability across ELD teachers' classrooms for students who are concurrently enrolled in both ELD and ELA. This study provides broad-strokes evidence on how variability in teacher effectiveness may shape students' linguistic and academic outcomes across ELA and ELD classrooms. Such evidence may provide insights into the important role of teachers and the degree to which variation in instructional effectiveness may contribute to differences in student outcomes but should not be used as standalone evidence of which teachers are more effective than others, nor whether ELD or ELA teachers are "more important" for student performance.

Future research can explore this cross-classroom overlap in more depth to understand the types of instructional approaches in ELD and ELA classrooms that support ELP and ELA skill development. It may be of interest to model teacher contributions across ELD classrooms and other subjects, such as math, science, and social studies, to see if cross-subject variation can be observed in other subject areas. Further research may also probe into the pairings of teachers to understand the implications of students being shared across teachers who have varying levels of instructional effectiveness, and how building leaders can think about creating the conditions for EL-classified students to receive aligned, coherent language and content instruction across classrooms.

Conclusion

This study is a preliminary exploration of the extent to which variability in ELA and ELD teachers' instructional effectiveness may have implications for EL-classified students' ELA and ELP performance. Results suggest that variation in effectiveness does contribute to students' assessed performance. While within-subject variation appears to have more of an impact on student performance, there was evidence that both ELA and ELD teachers' instructional effectiveness levels

have important implications across the subject assessments. This information provides more evidence to inform policy investments in strengthening teachers' supports for working with EL-classified students, while also providing important evidence into the important role that ELD teachers play, a teaching role that may be overlooked in conversations around strengthening instruction and outcomes for EL-classified students. This work can inform hiring, framing investment in effective teachers as a critical lever for supporting EL-classified students. Additionally, this work can inform attention towards supports for current teachers, supports that focus on ensuring that all teachers are provided with the training and tools to support students' interrelated language and content development.

APPENDIX

Table A1. Estimated ELA and ELD Teacher Contributions to EL-Classified Students' ELA Outcomes, Alternative Approaches

	ELA				
	Alt 1	**Alt 2**	**Alt 3**	**Alt 4**	**Alt 5**
	Excluding SPED	Imputed	Imputed & 2+ ELA/ELD by G&Y	Main Sample, no school FE or RE	Teacher parameters as fixed effects
Teacher Intercepts (SD)					
ELA	0.057***	0.048***	0.040***	0.068***	0.322***
ELD	0.021~	0.036***	0.011	0.043***	0.284***
Covariates					
Student Cov.	X	X	X	X	X
Classroom Cov.	X	X	X	X	X
School Cov.	X	X	X	X	X
Fixed Effects					
Grade	X	X	X	X	X
Year	X	X	X	X	X
School	X	X	X		
N (obs)	6,142	11,776	3,303	9,513	9,513
N (ELA Teacher)	283	472	118	388	388
N (ELD Teacher)	156	236	70	211	211

Note. *p<.05, **p<.01, ***p<.001. SPED=special education. All models include student covariates of cubic polynomials of prior year ELA, math, and standardized ELP scores, as well as FRPL, special education, migrant, male, ethnicity/race, as well as classroom- and school-level averages of student covariates (other than lagged scores). Alternative model 1 is an alternative sample, restricted to students who are not identified for special education services. Alternative model 2 is also an alternative sample--the main sample, but augmented with students who were previously missing prior year scores with their scores imputed to the mean for their given grade and year within the sample. Alternative model 3 is also an alternative sample, alternative sample 2, but only restricted to students taught by ELA and ELD teachers who are each linked to two or more ELA or ELD teachers. Alternative model 4 is the main sample, but the model does not have school or random effects included. Alternative model 5 is the main sample, but with ELA and ELD teacher parameters fit as fixed, rather than random effects.

Table A2. Estimated ELA and ELD Teacher Contributions to EL-Classified Students' ELP Listening and Speaking Outcomes, Alternative Approaches

	Listening					Speaking				
	Alt 1	**Alt 2**	**Alt 3**	**Alt 4**	**Alt 5**	**Alt 1**	**Alt 2**	**Alt 3**	**Alt 4**	**Alt 5**
	Excluding SPED	Imputed	Imputed & 2+ ELA/ELD by G&Y	Main Sample, no school FE or RE	Teacher parameters as fixed effects	Excluding SPED	Imputed	Imputed & 2+ ELA/ELD by G&Y	Main Sample, no school FE or RE	Teacher parameters as fixed effects
Tch. Int. (SD)										
ELA	0.021	0.029**	0.022	0.033**	0.313***	0.017	0.037**	0.001	0.036**	0.403***
ELD	0.048***	0.053***	0.028*	0.099***	0.276***	0.097***	0.088***	0.072***	0.156***	0.351***
Covariates										
Student Cov.	X	X	X	X	X	X	X	X	X	X
Classroom Cov.	X	X	X	X	X	X	X	X	X	X
School Cov.	X	X	X	X	X	X	X	X	X	X
Fixed Effects										
Grade	X	X	X	X	X	X	X	X	X	X
Year	X	X	X	X	X	X	X	X	X	X
School	X	X	X			X		X		
N (obs)	3,878	6,981	1,972	6,293	6,293	3,878	6,944	1,957	6,271	6,271
N (ELA Teacher)	202	369	98	302	302	202	369	98	302	302
N (ELD Teacher)	116	177	51	160	160	116	177	51	160	160

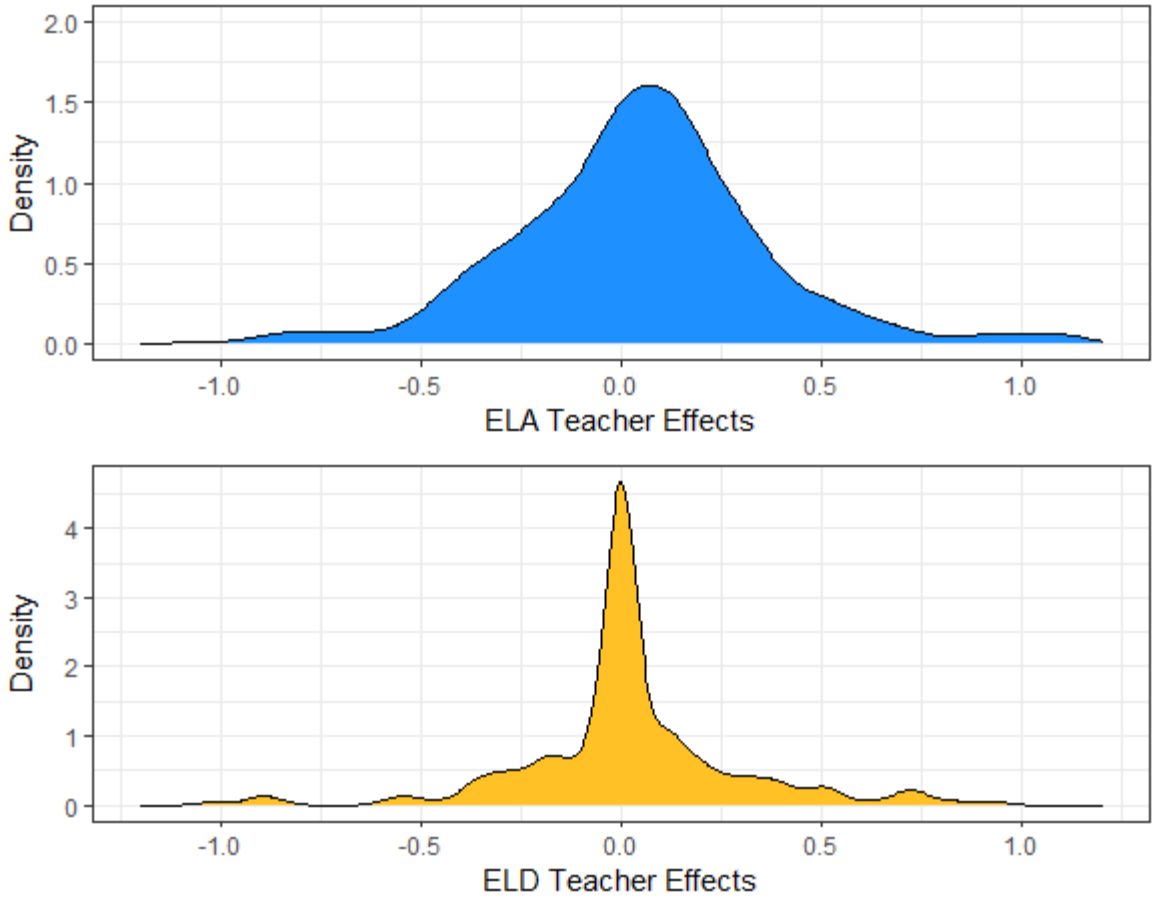
Note. *p<.05, **p<.01, ***p<.001. All models include student covariates of cubic polynomials of prior year ELA, math, and standardized ELP scores, as well as FRPL, special education, migrant, male, ethnicity/race, as well as classroom- and school-level averages of student covariates (other than lagged scores). Alternative model 1 is an alternative sample, restricted to students who are not identified for special education services. Alternative model 2 is also an alternative sample--the main sample, but augmented with students who were previously missing prior year scores with their scores imputed to the mean for their given grade and year within the sample. Alternative model 3 is also an alternative sample, alternative sample 2, but only restricted to students taught by ELA and ELD teachers who are each linked to two or more ELA or ELD teachers. Alternative model 4 is the main sample, but the model does not have school or random effects included.

Table A3. Estimated ELA and ELD Teacher Contributions to EL-Classified Students' ELP Writing and Reading Outcomes, Alternative Approaches

	Writing					Reading				
	Alt 1	**Alt 2**	**Alt 3**	**Alt 4**	**Alt 5**	**Alt 1**	**Alt 2**	**Alt 3**	**Alt 4**	**Alt 5**
	Excluding SPED	Imputed	Imputed & 2+ ELA/ELD by G&Y	Main Sample, no school FE or RE	Teacher parameters as fixed effects	Excluding SPED	Imputed	Imputed & 2+ ELA/ELD by G&Y	Main Sample, no school FE or RE	Teacher parameters as fixed effects
Tch. Int. (SD)										
ELA	0.016	0.016*	0.022	0.026**	0.300***	0.020	0.036***	0.004	0.041**	0.264***
ELD	0.039***	0.036**	0.023	0.068***	0.246***	0.064***	0.063***	0.051**	0.103***	0.248***
Covariates										
Student Cov.	X	X	X	X	X	X	X	X	X	X
Classroom Cov.	X	X	X	X	X	X	X	X	X	X
School Cov.	X	X	X	X	X	X	X	X	X	X
Fixed Effects										
Grade	X	X	X	X	X	X	X	X	X	X
Year	X	X	X	X	X	X	X	X	X	X
School	X	X	X			X		X		
N (obs)	3,878	6,861	1,920	6,277	6,277	3,878	6,827	1,901	6,295	6,295
N (ELA Teacher)	202	369	98	302	302	202	369	98	302	302
N (ELD Teacher)	116	177	51	160	160	116	177	51	160	160

Note. *p<.05, **p<.01, ***p<.001. All models include student covariates of cubic polynomials of prior year ELA, math, and standardized ELP scores, as well as FRPL, special education, migrant, male, ethnicity/race, as well as classroom- and school-level averages of student covariates (other than lagged scores). Alternative model 1 is an alternative sample, restricted to students who are not identified for special education services. Alternative model 2 is also an alternative sample--the main sample, but augmented with students who were previously missing prior year scores with their scores imputed to the mean for their given grade and year within the sample. Alternative model 3 is also an alternative sample, alternative sample 2, but only restricted to students taught by ELA and ELD teachers who are each linked to two or more ELA or ELD teachers. Alternative model 4 is the main sample, but the model does not have school or random effects included.

Figure A1. Distribution of ELA and ELD Teacher Contributions to EL-Classified Students' ELA Outcomes, Fixed Effect Specification



REFERENCES CITED

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics 25*(1) 95–135.

Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometrics issues. *International Journal of Phytoremediation*, *21*(1), 231–257.

August, D., Branum-Martin, L., Cardenas-Hagan, E., & Francis, D. J. (2009). The impact of an instructional intervention on the science and language learning of middle grade English language learners. *Journal of Research on Educational Effectiveness*, *2*(4), 345–376.

Bitler, M., Corcoran, S. P., Domina, T., & Penner, E. K. (2021). Teacher effects on student achievement and height: A cautionary tale. *Journal of Research on Educational Effectiveness*, *14*(4), 900-924.

Bunch, G. C. (2013a). Pedagogical language knowledge. *Review of Research in Education*, *37*(1), 298–341.

Buzick, H. M., & Laitusis, C. C. (2010). Using growth for accountability: Measurement challenges for students with disabilities and recommendations for research. *Educational Researcher*, *39*(7), 537-544.

Callahan, R. M. (2006). The intersection of accountability and language: Can reading intervention replace English language development? *Bilingual Research Journal*, *30*(1), 1–21.

Callahan, R. M., & Shifrer, D. (2016). Equitable access for secondary English learner students: Course taking as evidence of EL program effectiveness. *Educational Administration Quarterly*, *52*(3), 463–496.

Castaneda v. Pickard. United States Court of Appeals, Fifth Circuit. 648 F.2d 989 (5th Cir. *1981*)

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, *104*(9), 2633–2679.

Dabach, D. B. (2014). "I am not a shelter!": Stigma and social Boundaries in teachers' accounts of students' experience in separate "Sheltered" English learner classrooms. *Journal of Education for Students Placed at Risk*, *19*(2), 98–124.

DiCerbo, P. A., Anstrom, K. A., Baker, L. L., & Rivera, C. (2014). A review of the literature on teaching academic English to English language learners. *Review of Educational Research 84*(3), 446–482.

Eamon, M. K. (2005). Social-demographic, school, neighborhood, and parenting influences on the academic achievement of Latino young adolescents. *Journal of Youth and Adolescence*, *34*(2), 163–174.

Estrada, P. (2014). English learner curricular streams in four middle schools: Triage in the trenches. *Urban Review*, *46*(4), 535–573.

Every Student Succeeds Act, 20 U.S.C. § 6301 (2015).

Faltis, C., Arias, M. B., & Ramírez-Marín, F. (2010). Identifying Relevant Competencies for Secondary Teachers of English Learners. *Bilingual Research Journal*, *33*(3), 307–328.

Gándara, P., & Santibañez, L. (2016). The teachers our English language learners need. *Educational Leadership 64(*2) 1-22.

Gersten, R., & Baker, S. (2000). What we know about effective instructional practices for English-Language learners. *Exceptional Children*, *66*(4), 454–470.

Goldenberg, C. (2013). Unlocking the research on English learners: What we know--and don't yet know--about effective instruction. *American Educator*. 1-14.

Grossman, P., Loeb, S., Cohen, J., & Wyckoff, J. (2013). Measure for measure: The relationship between measures of instructional practice in middle school English language arts and teachers' value-added scores. *American Journal of Education*, *119*(3), 445–470.

Hansen-Thomas, H., & Cavagnetto, A. (2010). What do mainstream middle school teachers think about their English language learners? A tri-state case study. *Bilingual Research Journal*, *33*(2), 249–266.

Hanushek, E. A. (2020). Education production functions. In *The Economics of Education: A Comprehensive Overview* (pp. 161–170). Elsevier.

Harper, C. A., & de Jong, E. J. (2009). English language teacher expertise: The elephant in the room. *Language and Education*, *23*(2), 137–151.

Harper, C., de Jong, E., & Coady, M. (2011). From preservice to practice: Mainstream elementary teacher beliefs of preparation and efficacy with English Language Learners in the state of Florida. *Bilingual Research Journal*, *34*(2), 223–239.

Hock, H., & Isenberg, E. (2017). Methods for accounting for co-teaching in value-added models. *Statistics and Public Policy*, *4*(1), 1–11.

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

Jones, N. D., Buzick, H. M., & Turkan, S. (2013). Including students with disabilities and English learners in measures of educator effectiveness. *Educational Researcher*, *42*(4), 234–241.

Kieffer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology*, *100*(4), 851–868.

Kim, Y. K. yong, Curby, T. W., & Winsler, A. (2014). Child, family, and school characteristics related to English proficiency development among low-income, dual language learners. *Developmental Psychology*, *50*(12), 2600–2613.

King, E. T., & Scott, L. M. (2014). English as gatekeeper: Linguistic capital and American schools. *Journal for Multicultural Education*, *8*(4), 226–236.

Koedel, C. (2009). An empirical analysis of teacher spillover effects in secondary school. *Economics of Education Review*, *28*(6), 682–692.

Koedel, C., Mihaly, K., & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review 47*(2), 180–195.

Kraft, M. A. (2019). Teacher effects on complex cognitive skills and social-emotional competencies. *Journal of Human Resources*, *54*(1), 1–36.

Lipscomb, S., Teh, B. R., Gill, B., Chiang, H., & Owens, A. (2010). *Teacher and Principal Value-Added: Research Findings and Implementation Practices*. Mathematica Policy Research, Inc.

Llosa, L., Lee, O., Jiang, F., Haas, A., O'Connor, C., van Booven, C. D., & Kieffer, M. J. (2016). Impact of a large-scale science intervention focused on English language learners. *American Educational Research Journal*, *53*(2), 395–424.

Loeb, S., Soland, J., & Fox, L. (2014). Is a good teacher a good teacher for all? Comparing value-added of teachers with their English learners and non-English learners. *Educational Evaluation and Policy Analysis*, *36*(4), 457–475.

López, F., Scanlan, M., & Gundrum, B. (2013). Preparing teachers of English language learners: Empirical evidence and policy implications. *education policy analysis archives*, *21*(0), 20-42.

Master, B., Loeb, S., Whitney, C., & Wyckoff, J. (2016). Different skills?: Identifying differentially effective teachers of English language learners. *Elementary School Journal*, *117*(2), 261–284.

Mulhern, C. (2020). Beyond teachers: Estimating individual guidance counselors' effects on educational attainment. *Unpublished Manuscript, RAND Corporation*.

Nye, B., Konstantopoulos, S., & Hedges, L. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, *26*(3), 237–257.

Oregon Department of Education (2019). Oregon English Language Arts and Literacy Standards. Retrieved from: https://www.oregon.gov/ode/educator-resources/standards/ELA/Documents/Grade%206.pdf

Paige, M. A., & Amrein-Beardsley, A. (2020). "Houston, we have a lawsuit": A cautionary tale for the implementation of value-added models for high-stakes employment decisions. *Educational Researcher*, *49*(5), 350-359.

Pettit, S. K. (2011). Teachers' beliefs about English language learners in the mainstream classroom: A review of the literature. *International Multilingual Research Journal* 5(2) 123–147.

Polat, N., Zarecky-Hodge, A., & Schreiber, J. B. (2016). Academic growth trajectories of ELLs in NAEP data: The case of fourth- and eighth-grade ELLs and non-ELLs on mathematics and reading tests. *Journal of Educational Research*, *109*(5), 541–553.

Reeves, J. R. (2006). Secondary teacher attitudes toward including English Language Learners in mainstream classrooms. *Journal of Educational Research 99*(3) 131–143

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417–458.

Saunders, W. M., Foorman, B. R., & Carlson, C. D. (2006). Is a separate block of time for oral English language development in programs for English learners needed? *Elementary School Journal*, *107*(2), 181–198.

Shea, L. M., Sandholtz, J. H., & Shanahan, T. B. (2018). We are all talking: a whole-school approach to professional development for teachers of English learners. *Professional Development in Education*, *44*(2), 190–208.

Snow, M. A., & Katz, A. (2010). *English language development: Foundations and implementation in kindergarten through grade five*. California Department of Education

Stoddart, T., Pinal, A., Latzke, M., & Canaday, D. (2002). Integrating inquiry science and language development for English language learners. *Journal of Research in Science Teaching*, *39*(8), 664–687.

Suárez-Orozco, C., Hee Jin Bang, & Onaga, M. (2010). Contributions to variations in academic trajectories amongst recent immigrant youth. *International Journal of Behavioral Development*, *34*(6), 500–510.

Thompson, K. D. (2017). English learners' time to reclassification. *Educational Policy*, *31*(3), 330–363.

Torre Gibney, D., & Henry, G. (2020). Who teaches English learners? A study of the quality, experience, and credentials of teachers of English learners in a new immigrant destination. *Teaching and Teacher Education*, *90(1)*, 102-143.

Verplaetse, L. S., & Migliacci, N. (2017). Making mainstream content comprehensible through sheltered instruction. In *Inclusive pedagogy for English language learners: A handbook of research-informed practices* (pp. 127-166). Routledge.

Villavicencio, A., Jaffe-Walter, R., & Klevan, S. (2021). "You can't close your door here:" Leveraging teacher collaboration to improve outcomes for immigrant English Learners. *Teaching and Teacher Education*, *97*, 103-227.

Walqui, A., & van Lier, L. (2010). *Scaffolding the Academic Success of Adolescent English Language Learners: A Pedagogy of Promise*. WestEd

Yuan, K. (2015). A value-added study of teacher spillover effects across four core subjects in middle schools. *education policy analysis archives*, *23*(0), 1–24.

Zwiep, S. G., Straits, W. J., Stone, K. R., Beltran, D. D., & Furtado, L. (2011). The integration of English language development and science instruction in elementary classrooms. *Journal of Science Teacher Education*, *22*(8), 769–785.

CHAPTER V. CONCLUSION

This set of dissertation studies is unified by a focus on different elements of EL education and policy. The goal of this dissertation was to produce evidence to inform policymaking that strengthens services and system conditions for EL-classified students. Drawing on data from one U.S. state, the studies ask a range of research questions that span different groups of EL-classified students and examine multiple different ways in which system conditions may be supporting or constricting EL-classified students' opportunities and outcomes. From this body of work, two key themes emerged. One theme is identifying key areas for policy intervention. Another is the challenges in developing, implementing, and evaluating policy levers designed to support EL-classified students.

Key Areas for Policy Intervention

All three studies examined different ways in which EL-classified students' opportunities and outcomes are shaped. Research questions for chapters II and IV, specifically, focused on identifying potential areas for policy intervention to support EL-classified students.

In Chapter II, I find evidence that, among immigrant EL-classified students who arrive in secondary grades, access to core content is inequitable, both in relation to their peers and across a host of student characteristics among immigrant EL-classified students. Raw differences show how Black and Hispanic/Latinx immigrant EL-classified students who arrive in secondary grades are enrolled in core content at lower rates than their non-Black and Latinx immigrant EL-classified peers, while FPRL-eligible students and students with interrupted formal education are also enrolled in core content courses at lower rates. Modeling different predictors of student enrollment suggest that enrollment is largely predicted by students' initial ELP and SIFE status. When paired with the raw differences, this provides evidence that schools and districts are not structuring student schedules in ways that ensure all students have equitable access, with under-enrollment concentrated among students who face compounding intersections of historical marginalization. Chapter II also provides

examples of how data can be used toward this aim, encouraging state, district, and local education agencies to use descriptive statistics and analyses to explore how course access (and other indicators of opportunities to learn) may differ by student characteristics in ways that not only constrict opportunity, but do so for students who are members of historically marginalized communities. Chapter II also includes perspectives from the field, which provide further insights for policymakers into how local policy implementation and agency are shaping critical enrollment decisions—with opportunities again to see these decision-making points as policy entry points and develop guidance in response to quantitative patterns and reported practices and assumptions.

Chapter IV also identifies a key area for policy intervention. By providing preliminary evidence on the implications of variability in teacher effectiveness for EL-classified students' ELA and ELP performance, the findings from this study highlight the important role that teachers play. These findings encourage conversations around how strengthening instruction in both ELA and ELD classrooms could translate into significant improvements in instruction and outcomes for EL-classified students. Cross-subject contributions to learning across ELA and ELD teachers, while relatively small in magnitude, can encourage policy to explore how ELA and ELD teachers are supported through training and professional development to support EL-classified students' language and literacy growth. Supports can recognize the overlap in core content taught and the opportunities to align instruction across classrooms in ways that can support students' English language and literacy skills and knowledge development. Together, this work provides evidence for policymakers that investing in teachers is a critical policy lever through which EL education can be strengthened.

Challenges in Developing, Implementing, and Evaluating EL Policies

Two studies include an explicit focus on evaluating whether policy levers result in significant changes in EL-classified students' instructional environments and opportunities—with Chapter II examining whether newcomer programs seem to support course access among immigrant EL-

classified students, and Chapter III examining whether a state policy providing technical support and funding while introducing accountability pressures resulted in changes to the instructional environments and academic outcomes of EL-classified students. Both studies find little to no evidence that these policy levers resulted in meaningful changes for the populations of interest. I found evidence that newcomer programs did not support fuller course access and may have even constricted course access, while receiving supports through HB 3499 identification did not result in significant increases in the ratio of ESOL endorsed teachers to EL-classified students or the likelihood of receiving bilingual services, nor improved academic test score outcomes over the years observed. These findings speak to some of the challenges in developing and evaluating EL policies.

The first challenge is that the structural barriers that EL-classified students face in U.S. districts and schools are rooted in systems that overwhelmingly have been created to provide teaching and learning for white, monolingual English-speaking students (Flores et al., 2015; Flores & Rosa, 2015). While there are policies and practice developed to support EL-classified students, including those evaluated in Chapters II and III, these policies are designed and implemented to support students who often face critical, intensive barriers such as structural racism, language hierarchies, and under-resourced schools, taught by teachers who themselves often report that they are underprepared to provide appropriate instruction (Flores et al., 2015; Flores & Rosa, 2015; Gándara et al., 2003; Harper & de Jong, 2009; Valenzuela, 1999).

These different intersecting challenges, for example, may help to contextualize the newcomer program findings. For arriving immigrant EL-classified students, many of whom are students of color navigating racialized systems and the transition to a new community (Brown & Chu, 2012; Rodriguez et al., 2022), newcomer programs may provide an important source of support in many ways, such as socioemotional support, cultural orientation, and targeted academic supports (Hos, 2016). However, newcomer programs may not be developed and implemented in ways that address key issues of

exclusion from core content, instead focusing on addressing other critical challenges that students may be experiencing. The nexus of challenges that both schools and students report may not be fully able to be addressed through newcomer programs, as implemented.

For HB 3499 districts, the intervention was designed in recognition that districts may need additional resources and support to strengthen their EL services (Carnock, 2017). As implemented, the intersection of funding, support, and accountability resulted in a significant increase in EL expenditures per EL-classified student but did not translate into significant changes or improvements in measured services or outcomes in relation to non-identified districts and trends prior to identification. Contextualized within the challenges and need for supports and resources that districts report (Tannenbaum et al., 2012), HB 3499 may have represented an important step towards strengthening local services for EL-classified students but been too limited in scope to see meaningful improvements across identified districts. While there may be pockets of improvement, perhaps individual districts, schools, or even classrooms saw substantial changes in response to the intervention, the evidence suggests that the scope of the intervention did not fully address the challenges districts and schools face.

There are also challenges in evaluating whether EL policy levers have an impact. Policy evaluation is limited in the conclusions that can be made depending on what has been measured and within what time frame. For both Chapters II and III, there are salient challenges in evaluating whether the policies resulted in meaningful change. For example, Chapter II examined course access, an important indicator of opportunity to learn (Callahan & Shrifer, 2016). However, given the set of linguistic, academic, and socioemotional challenges that arriving immigrant students may face, there are many more important outcomes that may be of importance to schools and students not captured in this policy evaluation, such as quality of peer relationships, sense of school belonging, and home language development. Similarly, in Chapter III, I look at both proximal and distal outcomes, but there

are many other ways to think about how HB 3499 could have impacted the educational environment and experiences of EL-classified students. Additionally, three years may not have been a sufficient timeline along which to evaluate HB 3499's policy impact, given evidence that school improvement timelines can be extended (Sun et al., 2021).

These challenges are not new in EL education nor policy evaluation but can be challenging to navigate when thinking about how to evaluate policy impacts and use this evidence to inform future policymaking. For example, the conclusion I draw from both Chapters II and III is not that the policies have "failed". Rather, for newcomer programs I argue that the evidence on newcomer programs suggests that policy design and implementation should be examined closely for ways in which the mechanics and goals of programs may be leading to exclusion from core content. This examination can inform guidance for implementing these programs, if local education agencies feel that the program designs are providing critical supports. For HB 3499, I suggest that the policy's expected outcomes should be contextualized within the scope of the intervention, and that policymakers should consider thinking about whether the scope is large enough to see meaningful change at the district level.

All three chapters in this dissertation add to the body of work on EL education and EL policy, with the goal of supporting evidence-based policymaking. Central conclusions from these studies align with other research or understandings about how to support EL-classified students. I find that opportunity to learn is a critical issue for policy intervention (Callahan, 2005; Callahan & Shrifer, 2016; Johnson, 2019; Umansky, 2016), that teachers' instructional effectiveness plays an important role in EL-classified students' education (Dabach, 2014; Harper & de Jong, 2009), and that EL-classified students' language and content development are intertwined (Bunch, 2013; Bunch et al., 2012; Callahan, 2006). Together, in concert with prior research and current policy efforts, the evidence provided through this set of studies can inform reflection, priorities, and investment in policymaking,

with the hope of translating into meaningful supports and guidance for districts and schools that serve EL-classified students, as well as stronger supports and services for EL-classified students.

REFERENCES CITED

Brown, C. S., & Chu, H. (2012). Discrimination, ethnic identity, and academic outcomes of Mexican immigrant children: The importance of school context. *Child Development, 83*(5), 1477-1485.

Bunch, G. C. (2013). Pedagogical language knowledge. *Review of Research in Education*, *37*(1), 298–341.

Callahan, R. M. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal, 42*(2), 305-328.

Callahan, R., & Shifrer, D. (2016). Equitable access for secondary English learner students: Course taking as evidence of EL program effectiveness. *Educational Administration Quarterly, 52*(3), 463-496.

Carnock, J. (2017) *Pioneering Change: Leveraging Data to Reform English Learner Education in Oregon.* New America Foundation

Dabach, D. B. (2014). "I am not a shelter!": Stigma and social boundaries in teachers' accounts of students' experience in separate "sheltered" English learner classrooms. *Journal of Education for Students Placed at Risk 19*(2), 98-124.

Flores, N., Kleyn, T., & Menken, K. (2015). Looking holistically in a climate of partiality: Identities of students labeled long-term English language learners. *Journal of Language, Identity & Education*, *14*(2), 113-132.

Gándara, P., Rumberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English Learners in California Schools: Unequal resources,'Unequal outcomes. *education policy analysis archives, 11*(36), 1-24.

Harper, C., & de Jong, E. J. (2009). English language teacher expertise: The elephant in the room. *Language and Education*, *23*(2), 137–151.

Harper, C., de Jong, E., & Coady, M. (2011). From preservice to practice: Mainstream elementary teacher beliefs of preparation and efficacy with English Language Learners in the state of Florida. *Bilingual Research Journal*, *34*(2), 223–239.

Hos, R. (2016). Caring is not enough: Teachers' enactment of ethical care for adolescent students with limited or interrupted formal education (SLIFE) in a newcomer classroom. *Education and Urban Society, 48*(5), 479-503.

Johnson, A. (2019). A matter of time: Variations in high school course-taking by years-as-EL subgroup. *Educational Evaluation and Policy Analysis, 41*(4), 461-482.

Sun, M., Kennedy, A. I., & Loeb, S. (2021). The longitudinal effects of School Improvement Grants. *Educational Evaluation and Policy Analysis*, *43*(4), 647-667.

Tanenbaum, C., Boyle, A., Soga, K., Le Floch, K. C., Golden, L., Petroccia, M., Toplitz, M., Taylor, J., & O'Day, J. (2012). National Evaluation of Title III Implementation: Report on State and Local Implementation. *Office of Planning, Evaluation and Policy Development, US Department of Education*.

Umansky, I. M. (2016a). Leveled and exclusionary tracking: English learners' access to academic content in middle school. *American Educational Research Journal, 53*(6), 1792-1833.

Valenzuela, A. (1999). Subtractive Schooling: U.S.-Mexican youth and the politics of caring. SUNY Series. *The Social Context of Education*.