Title
SEParating the Cream: Selective Enrollment Public Schools and their Districtwide Effects

Author
Irwin, Veronique

Publication Date
2020

SEParating the Cream:
Selective Enrollment Public Schools and their Districtwide Effects

By

Véronique Tara Irwin

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Sociology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Samuel R. Lucas, Chair
Professor Daniel Schneider
Professor Prudence Carter

Summer 2020

ABSTRACT

SEParating the Cream: Selective Enrollment Public Schools and their Districtwide Effects

by

Véronique Irwin

Doctor of Philosophy in Sociology

University of California, Berkeley

Professor Samuel R. Lucas, Chair

Fierce local debates throughout the United States surround the equity of admitting students to public schools using academic criteria. Although research has evaluated the central assumption of these debates—that Selective Enrollment Public (SEP) schools enhance the welfare of students who attend them—none has addressed the district-level outcomes associated with these schools. This is important because the selectivity and scope of SEP schools produce tiered school systems (SEP *districts*). This district-level process, in turn, calls for an analysis of district-level educational outcomes. To address this gap, I compile an original list of SEP schools using an innovative web scraping procedure. I combine these data with secondary sources from the Common Core of Data, the Education Demographic and Geographic Estimates, the Stanford Education Data Archive, and the Civil Rights Data Collection to evaluate the districtwide effects of SEP school on three sets of key educational outcomes: segregation, achievement, and access to college preparation.

Before presenting these three quantitative analyses, I offer a detailed definition and background of SEP schools, which have been given little theoretical consideration by sociologists. The substantive analyses then begin by considering the first order question of whether sorting students on the basis of performance has implications for students sorting along other dimensions, namely race. I find evidence of some differences in patterns of segregation among white and under-represented minority (URM) students in SEP versus non-SEP districts, particularly in high school, which may suggest less stability for URM students attending predominantly white schools in SEP districts. With a better understanding of how SEP schools shape student sorting, I then evaluate districtwide academic outcomes. Using a grade-level Difference-in-Differences design, I find evidence of slower math achievement growth and widening white-Latinx gaps in SEP districts from third to eighth grade. Finally, evidence from Seemingly Unrelated Regression models suggests that, despite a narrower provision of Advanced Placement (AP) courses, AP enrollment rates are just as high (or higher) and exam passing rates are higher for white and Latinx AP participants in SEP versus non-SEP districts. However, white-URM gaps in college preparation are widest in SEP districts.

Overall, evidence of the effects of SEP districts on the average *quality* of education is mixed, but evidence suggesting that SEP districts exert significant influence on racial *inequality* in educational opportunity is relatively consistent. Even increases in efficiency appear to come at the cost of equity. Most importantly, this work highlights the importance of considering SEP schools as part of a differentiated school *system*, rather than isolated elite institutions.

CONTENTS

ACKNOWLEDGEMENTS

Graduate school is meant to be challenging. Even so, 2020 has seemed determined to set up unexpected hurdles in the last leg of the race, both globally and personally. These added challenges make it all the more precious to have this space to thank those without whom I would never have found the finish line.

First, I want to thank my dissertation committee. To Sam Lucas, thank you for always encouraging me to celebrate my accomplishments. It has been a privilege to work with and learn from such an intelligent, dedicated, and accomplished scholar from my very first day in graduate school. To Danny Schneider, thank you for always having an open door. In the relatively short time we worked together, you have contributed hugely to my growth as a researcher. I am grateful for your guidance, your feedback, and your belief in my work, especially when I did not seem to believe in it much myself.

I am also grateful to Prudence Carter, whose research largely inspired both my undergraduate and Masters theses. It has been a gratifying full-circle experience to have you on my dissertation committee. I would also like to extend a special thank you to Tony Jack, who has offered ongoing support and friendship since I was an undergraduate, and without whom I would not be in graduate school in the first place.

But of course, feedback, advice, and especially support came not only from advisors, but from my fellow graduate students. Graduate school would have been impossible without them. I was lucky to have this support from the very beginning, thanks to the amazing group of people in the Berkeley Sociology Cohort of 2013. Aya, David, Fatinha, Ghaleb, Jonathan, Liz, Mark, Martin, Ogi, Phung, Robert, Santiago, Seth, and Steve—thank you for giving me an instant community and for making me feel at home in a new place and a new phase of life. Over time, my Berkeley community grew even larger, especially when I moved into 362 Barrows and quickly made myself known as the denizen of the third floor. I am especially thankful for all of my amazing office mates over the last five years—David, Dani, Gowri, Mikaela, Allison, Alex, Krista, Margaret, Kim, and Robert, who indulged me in keeping our door open—and for everyone who took the open door as an opportunity to stop by and say hello. For having a pleasant place to go when I left my apartment each day, I also owe a special acknowledgement to Tamar Young, who went above and beyond to make Barrows a more congenial place to work. I was also lucky to have several small research communities throughout my time at Berkeley, from REIW, to MESS, to Shift. Workshopping my research in these spaces has both broadened and sharpened my thinking. But, just as importantly, these have also been spaces to recognize (and commiserate about) the challenges of doing research in graduate school, and to support each other through the process. Finally, to those who helped me blow off steam, and especially those who helped me *smash* it—Phung, Kit, Zawadi, Matt, and Jason—thank you for providing the best possible outlet. In these socially isolating times, it is these interactions I miss the most.

No acknowledgement of my Berkeley Sociology support system could be complete without specifically thanking two people: David Showalter and Phung Su. David, you were my first friend in graduate school. It took just a few months—of brilliant conversation and nineties movies—to know I'd made a lifelong friend. Thank you for the innumerable baked goods and chats over drinks. Phung, you have taken me on adventures, broadened my mind, championed my accomplishments, and been a fierce friend. Thank you for all the excuses to eat good food and for putting our spare room to use. I am beyond grateful to you both, for the intellectual stimulation, the heart-to-hearts, and the genuine support over the years.

To my family. À ma famille Canadienne Française, qui envoie de loin des ondes positives et de la fierté qui me rend vraiment humble. To my Buffalo family, who have made holidays precious during these years with too few visits home. And to my Bachhuber-Miller family, who have treated me like one of their own since I was seventeen. Thank you all.

But especially to my parents. Thank you for always supporting me, for being such shining examples of hard work, and for letting me know you are proud of me. Particularly at this moment in time, as graduates everywhere are mourning the loss of pomp and circumstance, I feel so lucky to have parents who have *always* let me know they are proud of me, not just at major milestones, but every single day. Although I am sad not to have this opportunity to celebrate with you, my disappointment is far outweighed by my daily gratitude for having two such incredible parents who cheer me on, from near or far, in every challenge I face.

Finally, I am exceptionally lucky to have two "other halves." In completely different ways, they each make me whole. I could never thank them enough.

Kayla, you inspire me constantly. You have taught me so many lessons about perseverance and about taking the time to notice and find joy in seemingly small, everyday beauties. As I struggled through this final stretch of graduate school, writing my dissertation in quarantine, these lessons were more relevant than ever. Being far from you has been one of the biggest challenges of the last seven years. Thank you for being my reminder to stop and smell the jasmine, even from three thousand miles away.

Eric, for the last five years, you have commuted four hours a day, five days a week, so that we could face the ups and downs of graduate school (and life) together. Although we studied and worked in different cities, your sacrifice gave us the comfort of being true partners through it all. It feels fitting that—although at different times and for different reasons—we both wrote the final chapters of our dissertations sitting side by side. Even though we had separate finish lines, we crossed them together. Thank you for being a constant source of encouragement and laughter along the way. You make every burden lighter and every day brighter. I love you.

PREFACE

In my graduate school application, I opened my personal statement by describing the school I attended from fifth to twelfth grade, an exam-school school that prided itself on educating "the best and the brightest." I described it as "a protective bubble in many ways. However, it's location in the middle of the Fruit Belt, one of Buffalo's most distressed neighborhoods, provided a daily reminder of exactly what that bubble was protecting us from." After a graduate degree in Sociology, I cringe at the language I used in that statement, although my intended sentiment holds. I was given an opportunity that I knew very few students in Buffalo were given—access to a rigorous college preparatory education, with an established track record of placing students in four-year, often elite, universities. It was a gateway to upward mobility that wasn't easy to come by in a city experiencing one of the longest economic declines and highest rates of poverty in the country, challenges we could see acutely in the very neighborhood where my alma mater is located.

The dissonance of such a privilege amidst extreme disadvantage is the experience to which I generally credit my attraction to sociology. And the more tools I gained in this discipline, the more I came back to it. Although my dissertation does not focus on Buffalo, I consider it my tribute to the budding sociological imagination of my teenage self. It feels fitting to conclude my formal study of Sociology where it informally began.

In 2018, about year into my dissertation research, Gary Orfield and Jennifer Ayscue published a book titled *Discrimination in Elite Public Schools*. It was a case study of Buffalo, NY. This study posed exactly the questions that have sat with me for the better part of two decades and vividly described the stakes of maintaining such a system in a historically segregated city. Yet, despite having experienced this system first hand, and sharing their very concerns, I must admit, reading about Buffalo and City Honors like this… stung. It's never easy to read criticism of a place you loved.

While I was a student, the biggest assembly of the year was the "Keeping the Dream Alive" assembly, which was held every January in honor of Dr. Martin Luther King, Jr. It was a tradition begun by a long-time aide at the school who had herself played an integral role in *Brown v Board* as a teenager in Virginia. That history was woven into the legacy of our school, even if most of us didn't know or appreciate it at the time. By the time I enrolled, new court rulings had rolled back legal protections for considering race in admissions, catalyzing the resurgent inequalities that eventually prompted Orfield and Ayscue's report. During my time at City Honors, there was also a real sense among the students that the school was slipping from what it had once been—students had less autonomy, and a growing emphasis was placed on test-taking over free thinking. This could be felt in new assemblies that celebrated our school's ranking on lists from *Newsweek* or *US News and World Report* that were based on the *number* of tests we took, not how well we did on them. In fact, students generally struggled to earn their IB diplomas because of inadequate test scores. It was not the coursework, but the people who helped me most to grow. I made better friends and had better conversations at City Honors than I did in my next four years at Harvard. It is these people, not my college classmates, who fill my Facebook newsfeed with strong and informed viewpoints about the need for equality in housing, healthcare, pay, and *schooling*.

City Honors does *not* provide adequate opportunity for poor and minority students in Buffalo. It *does* disproportionately admit students from white and middle class families whose parents would likely have found decent educational opportunities in its absence, students like me. But for those who attended—at least while I was there—it provided a diversity of experiences and viewpoints that were not replicated elsewhere in the city or surrounding suburbs. My classmates and I were likely to do fine without City Honors; maybe not quite as well, but fine. This is precisely what the entrance exam was selecting on, after all. Rather, what City Honors gave to us were these personal connections and the tools to grow into informed, engaged, and empathetic citizens. But in a system where other schools struggle to produce graduates who can read at a high school level, this two-tiered system is consequential.

In sharing my connection to the schools that are the focus of this dissertation, I hope to convey the sense of gravity with which I approached my research. On the one hand, I loved City Honors and I gained a lot from it. On the other, my journey as a sociologist is rooted in a deep uneasiness with the inequalities I witnessed in the Buffalo Public Schools, inequalities to which City Honors appears to contribute. Because of this, there is no finding that this research could produce that would not weigh heavily on me. I have done my best to let the data speak for itself, and it is my sincere hope that this work will contribute to the expansion of opportunities like the ones that I had, in whatever form they may take.

I. INTRODUCTION:
A TALE OF TWO(-TIERED) DISTRICTS

In the United States, the landscape of public school choice has grown extremely complex—evolving from a system of residence-based neighborhood schools to one that includes charter schools, magnets, private vouchers and inter-district transfers. In the last 20 years alone, participation in school choice has increased by 20%[1] (National Center for Education Statistics 2018). This means that, particularly in large cities, a large proportion of students must now submit an application (or several) in order to secure their preferred seat in school. In some cities, this process involves academically competitive selection into the district's top schools, or Selective Enrollment Public (SEP) schools. Some of these schools—like Stuyvesant and Bronx Science in New York City and Boston Latin School in Boston—are among the most famous high schools in the country. In recent years, these schools have also been the subject of widely circulated op-eds highlighting the strained relation between educational excellence and equity (Ali and Chin 2018; Finn Jr 2012; Gay 2019).

The central tension in these debates is that these elite and high-performing schools are an important resource for the students and communities they serve, but that their academically selective admissions procedures produce significant disparities in access (e.g. Ebbert 2016; Irizarry 2017; Rey 2017). Demands for reform have typically centered around this issue of access, under the implicit assumption that SEP schools enhance the wellbeing of students who gain admission. Lost in this argument, however, is any consideration of the broader distributional consequences of having SEP schools in a district in the first place. Yet, insofar as competitive application procedures shape the composition of SEP schools, this process of cream skimming—or separating the cream from the milk—necessarily shapes the composition of neighborhood schools throughout the district as well. Rather than isolated institutions, these schools are a constituent part of distinct two-tiered systems, which may have important implications for the distribution of educational opportunities and outcomes districtwide. I refer to these two-tiered systems as SEP districts.

In January 2019, *The Boston Globe*'s published an article about this two-tiered structuring of the Boston public school system. In this article, a former valedictorian of one of Boston's traditional public schools, named Shanika, recounts being told by a teacher that her class rank and GPA "didn't mean much compared to kids at [the SEP school,] Latin," (Gay 2019). This experience highlights the interconnectedness between SEP and non-SEP schools within the same district. Although Shanika did not attend a SEP school herself, she was acutely aware of their existence. More importantly, the existence of these schools clearly shaped her school experience. On the one hand, the departure of high achieving students from her neighborhood school may have paved the way for Shanika to win her valedictorian status in the first place. On the other, she may have encountered lowered expectations from teachers and peers based on the truncated definition of "high achievement" established by the district's SEP school admissions exam. In either case, Shanika's experience highlights the potential for spillover effects of SEP schools beyond choice participants themselves. Yet we lack an understanding of what effect, if any, SEP schools have on outcomes throughout the district.

SEP districts might shape educational outcomes in a myriad of ways. The theory of

[1] The percentage of students *not* attending an assigned public school increased from 25.9% in 1999 to 31.2% in 2016 (including both public and private choice options).

(school) peer effects, for instance, argues that learning is shaped by the students with whom one shares a classroom or school. These effects might derive directly from peers in the form of motivation or competition or indirectly from teachers whose expectations or (sense of) efficacy is shaped by the composition of their classroom. By producing a two-tiered system, in which students identified as the "cream" are separated from the rest, SEP districts may thus have important consequences for peer effects.

Beyond sorting *students*, the two-tiered structure of SEP districts might also have consequences for the distribution of *resources* between schools. Sorting students between schools on the basis of their demonstrated preparation or motivation to participate in particular programs or activities might result in a complementary distribution of resources, from the most experienced teachers to teachers of English as a second language, and from laboratory equipment and college preparatory curricula to athletic fields and sports programs. The distribution of such resources, whether more likely to be used by students in SEP or non-SEP schools, shape the opportunities available to students—the passions they can develop, the resumes they can build, the postsecondary educational opportunities they are likely to pursue. On the one hand, cash-strapped districts may be able to more *efficiently* deploy limited capital if students are sorted into schools based on their (apparent) appetites for different resources. On the other, this kind of sorting early in the educational career may circumscribe opportunities and prematurely limit potential.

To the extent that SEP schools shape peer effects or resource distribution strategies within a district, they do so within the context of broader inequalities. Although innumerable inequalities shape the American school system, racial inequalities occupy a unique position in the landscape of U.S. education policy, in general, and in the criticism of SEP schools, in particular. At the district level, racial inequalities in access to SEP schools may exacerbate intra-district segregation by affording more advantaged families greater opportunity to exit neighborhood schools. Importantly, these same families may also be those who, in the absence of this opportunity, would exit the district altogether in favor of more exclusive options, such as wealthy suburban districts or private schools. In other words, SEP schools may impact inter-district segregation as well. Such a district-level influence on school choice would have important implications for whether SEP districts expand opportunities for "traditional" public school students or instead serve as an educational subsidy for the more advantaged.

Accordingly, this dissertation moves beyond existing research, which has focused on the effects of SEP schools on enrolled students, in order to examine whether important educational outcomes differ in SEP *districts* compared to other traditional public school districts. Further, in addition to achievement outcomes, I also consider the effect of SEP districts on outcomes like school segregation and resource distribution. Specifically, I consider three sets of district-level outcomes (1) school segregation among white and under-represented minority (URM) students, (2) math achievement, and (3) the distribution of college preparatory resources. Together, these outcomes provide indicators not only of the *average quality* of education in SEP districts compared to non-SEP districts, but of the *(in)equality* of opportunity as well.

To date, data limitations have prevented any such analyses. No existing dataset on the universe of U.S. public schools gathers data on selective assignment procedures. Thus, I use an innovative web scraping technique to produce an original dataset of SEP schools and the SEP districts that house them throughout the United States. I identify 90 SEP districts, including 369 SEP schools, operating throughout the United States and at all grade levels. These data are then combined with secondary data from the National Center for Education Statistics (NCES), the

Education Demographic and Geographic Estimates (EDGE) from the American Community Survey (ACS), and the Civil Rights Data Collection (CRDC), as well as newly available national achievement data from the Stanford Education Data Archive (SEDA, Reardon et al. 2018)[2]. Together, these data provide the unique ability to assess the role of SEP schools in shaping districts' overall ecology of educational opportunity. I deploy these data in a series of quantitative analyses, which are described briefly below.

Chapter Summaries

Chapters 2 introduce SEP schools in greater detail. I begin by offering a theoretical definition of these schools, rooted in the concept of student differentiation—the common practice of sorting students into subgroups for instructional purposes (Sørensen 1970). Following this definition, I briefly situate these schools within the history of U.S. education policy. Next, I delve into the extant literature to interrogate SEP schools' elite status. In so doing, I offer further justification for the focus on SEP *districts*, which are the object of interest throughout the empirical analysis.

Chapter 3 describes the semi-automated process of original data collection, using web-scraping tools in Python. Having defined *what* SEP schools are in Chapter 2, I introduce these data by describing *where* (geographically) and *when* (at what grade levels) these schools exist, as well as *how* they admit their students. Finally, I present data on *who* SEP districts and SEP schools serve in order to suggest that (1) SEP districts serve a significant share of U.S. public school students and (2) may exert disproportionate influence on racial disparities in educational opportunity, making them an important—and, so far, understudied—model of student differentiation.

Continuing the thread of *who* SEP districts serve, I begin my formal empirical analysis with the question of school segregation in Chapter 4. First, a descriptive analysis of inter-district segregation provides a new angle as to who SEP districts serve by examining how they influence student sorting through residential and school district choice. I show that SEP districts are (non-causally) associated with greater demographic dissimilarity among districts within a Core Based Statistical Area (CBSA). I then provide a more detailed analysis of intra-district segregation by grade level. Overall, I find that structural school transitions bring significantly different changes in the level of segregation in SEP versus non-SEP districts, specifically in terms of exposure to outgroup peers. Further, I identify trends in segregation over the course of high school which may suggest greater exit rates (transfers or dropouts) among URM students enrolled in predominantly white schools in SEP districts.

Next, Chapter 5 asks: do SEP districts produce different (1) overall levels of achievement growth or (2) inequality of achievement than non-SEP districts? And, given the documented concerns around equity in access, (3) do these effects differ for students of different racial/ethnic backgrounds? These analyses draw on newly available data from the Stanford Education Data Archive to examine the effect of SEP schools on district-level math achievement outcomes from

[2] At the time of this writing, data from these secondary sources provide relatively detailed breakdowns by student race, whereas breakdowns by family socioeconomic status are either less detailed (e.g., from the CCD and CRDC) or unavailable altogether (SEDA). Given this ongoing data constraint, analyses focus on segregation and group differences in outcomes by race. This is also in keeping with the primary focus of stakeholders to date on racial inequities in access to SEP schools.

third to eighth grade. Analyses follow a difference-in-differences design, using grade level as the longitudinal dimension. This approach facilitates a falsification test, using *future* treated districts (i.e., those which begin offering SEP schools in high school), to reject spurious causation. I find evidence of slower math achievement growth for the average student in SEP districts, overall, and separately for the average white, Black, and Latinx student. SEP districts also see an increase in the white–Latinx math achievement gap.

In the final empirical chapter, I ask whether the efficiency of and exposure to college preparatory resources differs for students in SEP versus non-SEP districts. In other words, this chapter asks whether college prep resources—specifically Advanced Placement (AP) courses—are allocated differently in districts that selectively differentiate students between schools. I find evidence that AP courses are more narrowly distributed in SEP districts than non-SEP districts and that the white-URM enrollment gaps are largest in SEP districts. However, white and Latinx students are more likely to pass at least one AP exam in SEP districts than non-SEP districts (specifically, those that also do not have magnet programs), whereas Black students are no more likely to pass exams. Given that AP courses are allocated more narrowly, this exam performance may indicate more "efficient" resources distribution. However, given that white-URM gaps also increase, it is important to consider that this efficiency may come at the cost of equality.

Taken together, this work highlights the importance of considering SEP schools as part of a differentiated school *system*, rather than isolated elite institutions. In particular, although evidence of effects on the average *quality* of education is mixed—from negative effects on math achievement in middle school to positive effects (for certain students) on AP participation and performance—the evidence is relatively consistent in suggesting that SEP schools exert significant influence on racial *inequality* in educational opportunity throughout the district. I expand on these implications in greater detail in the concluding chapter.

II. SELECTIVE ENROLLMENT PUBLIC SCHOOLS: ELITE, SEPARATE, AND UNEQUAL

Most Americans think of public school choice as the prerogative of students and their parents. Although it is a high-stakes process in many places, it is typically driven by *interest* on the part of families and *capacity* in the schools. In Selective Enrollment Public (SEP) schools, however, the choice process is reciprocal. It is not just school choice, but *schools'* choice. SEP schools are full schools in traditional public school districts that admit their entire student body using academically selective admissions criteria. This makes SEP schools an unusual form of public school choice in the United States. Yet, perhaps because they are unusual—and relatively uncommon—relatively little is known about these schools, at least on a national level.

In 2014, an Op-Ed was published in the *New York Times* with the title "Elite, Separate, and Unequal" (Kahlenberg 2014). This article describes the tension between excellence and equity in New York City's SEP schools[3]. This tension has been endemic to the provision of public schooling in the U.S. since the Supreme Court decisions that offer inspiration to the title of the piece. In this chapter, I will offer a brief overview of this history, with a particular focus on desegregation and gifted education, in order to suggest how it may have shaped SEP schools as they exist today. I then interrogate each of the three characteristics proffered by Kahlenberg—that SEP schools are (1) elite, (2) separate, and (3) unequal—to establish what is known about these schools from extant research and to expose the gaps in our understanding that form the motivation for the current work.

I will show that existing work has focused on the relatively *elite* students who attend SEP schools. This work suggests that SEP schools may be "winner-take-all" institutions (Attewell 2001) which have little to no positive effect on the achievement of marginally admitted students (Abdulkadiroğlu et al. 2014; Dobbie and Fryer Jr. 2011). While rigorous, this work fails to critically engage the fact that these schools are *separate*, i.e., that they selectively differentiate students between schools. This is consequential because students may experience very different educational environments, characterized by different educational expectations and resources. Finally, I turn to concerns around *equality*, which raise questions about whether SEP school may exacerbate opportunity and achievement gaps between students from different racial/ethnic backgrounds.

First, however, I begin by providing a theoretical definition of SEP schools based on two dimensions of student differentiation.

Academic Differentiation and the Place of Selective Enrollment Public Schools

Student differentiation refers to the practice of sorting students into subgroups for instructional purposes. The effect of student differentiation on achievement is the subject of long standing and unsettled debate in the sociology of education. Sørensen (1970) argued that one reason for this is the inadequacy of theory to identify specific mechanisms through which differentiation might have its effect. He therefore theorized several dimensions of differentiation. In defining SEP schools, I am interested specifically in two dimensions of student differentiation:

[3] Specifically, the piece focuses on New York's "specialized" high schools—eight of the most selective SEP schools in the city, which admit students solely based on their ranked performance on the Specialized High School Admissions Test (SHSAT).

(1) selectivity of assignment procedures, and (2) scope[4].

Figure 2.1. Defining SEP Schools

<table>
<tr><th></th><th></th><th colspan="3">Scope</th></tr>
<tr><th></th><th></th><th>Low</th><th>Medium</th><th>High</th></tr>
<tr><td rowspan="3">Selectivity of Assignment Procedure</td><td rowspan="2">Low</td><td rowspan="2">General Electives</td><td rowspan="2">Traditional Magnet Programs</td><td>Charters</td></tr>
<tr><td>Traditional Magnet Schools</td></tr>
<tr><td>High</td><td>Traditional Tracking</td><td>Selective Programs</td><td>SEP Schools</td></tr>
</table>

Selective Enrollment Public (SEP) schools are unique in the public school sector in the United States in that they are both (1) selective in their assignment procedures and (2) high in scope. Selectivity refers to the degree of homogeneity the assignment procedure is intended to produce in terms of students' demonstrated performance or preparation. Scope refers to the extent to which differentiation confines students to educational spaces shared only by students of the same educational subgroup. In other words, a school is more selective when the assignment procedure is enforced by gatekeepers on the basis of academic criteria as opposed to student interest[5], and scope is higher when differentiation occurs at the level of the school rather than the classroom. Figure 2.1 illustrates the position of several forms of student differentiation in the context of public school choice, including SEP schools, along these two dimensions. Specifically, SEP schools are defined here as full schools, not within-school tracks or programs, that use *academic* criteria[6] to admit their entire student body. SEP schools range in their level of

[4] Other important dimensions, according to Sørensen, are inclusiveness and whether differentiation is horizontal or vertical. Inclusiveness refers to the number of opportunities available at a given level of schooling (i.e., K-8 education) and is of little analytic value here, because schooling is compulsory in the U.S. until students are sixteen. Vertical differentiation reduces variation in students' presumed capacity to learn course content, while horizontal differentiation reduces the variation in the content delivered in a particular course setting. These latter concepts overlap significantly with the selectivity of assignment procedures.

[5] Sørensen's framework defines electivity—the extent to which subgroup placement is a choice for students—as a distinct characteristic of assignment procedure. In theory, an elective process could produce a homogenous (i.e., select) group. In practice, however, when assignment is not purely the prerogative of students, this either results in or is the result of selective criteria.

[6] As distinct from schools for the performing arts that use auditions or portfolios.

selectivity, from requiring basic proficiency to gifted identification, and utilize admissions criteria that might include prior grades, writing samples, standardized testing and/or specialized admissions exams. While many SEP schools use a pure ranked admissions procedure, others enter all academically eligible students into a lottery.

Put simply, the "sector" of schools covered here is not a homogenous one, and encompasses many kinds of learning environments[7]. For analytic purposes, however, I restrict the definition of SEP schools to schools that are funded and managed by a central, "traditional" public school district (i.e., districts that include other traditional public schools with neighborhood-based catchments areas). This is essential to allow for the analysis of system-wide (here, district-level) achievement outcomes. Students from throughout the district compete for admissions and the 'brightest' students gain admission to the best schools, while the rest are relegated to neighborhood schools. This results in a clearly tiered system, bolstered by the guise of meritocracy, which might reasonably shape not only peer groups, but also student exposure to expectations and resources throughout the district, key mechanisms theorized by Sørensen in the relationship between differentiation and student achievement. It is these whole districts, which I refer to as SEP districts, that constitute the educational systems of interest throughout the present study.

The Dual Legacy of Separate Schooling in the US

Before turning to the research on SEP schools more specifically, it is important to consider the broader education policy landscape of which they are a part. SEP schools are situated at an intersection between policies aimed at fostering advanced learning and future leadership on the one hand, and equalizing opportunities in education on other. While they are designed, by definition, to deliver specialized curricula to high-performing students, they may also expand overall opportunities by augmenting school choices offerings. Yet SEP schools do not occupy an established position in largescale policy "movements" for either of these objectives, such as, for example, the broad expansion of the Advanced Placement program or the proliferation of charter schools. In fact, because little is known about when these schools were founded (or became selective), they are somewhat difficult to place in history. In this section, I will highlight key moments in the history of U.S. public education, which—although direct influence is difficult to establish—necessarily shape the provision of SEP schools.

The Pursuit of Gifted Education

As World War II gave way to the Cold War, the United States faced a new existential threat from Soviet Union, which extended to education policy. The launch of Sputnik in 1957 raised fears that the United States was not keeping up in math, science, and technology. Despite a successful moon landing in 1969 and the establishment of the Office of Gifted Education in 1974, these fears persisted, coming to a head with the publication of *A Nation at Risk* in 1983.

[7] These schools are referred to by many different names by their operating districts, including selective enrollment schools, specialized schools, exam schools, criteria schools, etc. Notably, some of these schools are designated as magnets, while others are not. Magnets are public schools that offer specialized programs intended to attract students from throughout a district. Canonically, these programs are designed to foster integration and do not use admissions criteria (Fuller and Elmore 1996; Wells 1993). Variation in magnet status among SEP schools may therefore reflect different levels of emphasis on student integration associated with these schools.

The report, commissioned by the Reagan administration, warned of a "rising tide of mediocrity" (National Commission on Excellence in Education 1983). If the United States were to maintain its status as a world leader, it would have to recommit to Gifted and Talented education.

The most direct legislative response was the passage of the Jacob K. Javits Gifted and Talented Students Education Act of 1988. This act initially earmarked about $10 million a year, the majority of which went to the newly formed National Research Center on the Gifted and Talented (NRC/GT). Despite being defunded in 2011, the Javits Act has been rescued by Congress and began issuing grants again in 2014 (U.S. Department of Education 2019). The goal of this funding is to support research on the identification of and (separate) programming for gifted students, particularly those from groups that are typically underserved by gifted programs, like students from lower socioeconomic backgrounds or those with other special needs (Renzulli, Callahan, and Gubbins 2014). To this effect, the NRC/GT has developed guidelines for what they view to be equitable identification, including multiple criteria and specialized instruments for different areas of giftedness, in order to better identify students from different cultural backgrounds. However, uptake of these instruments and guidelines has been inconsistent, making the overall impact of the Javits Act difficult to assess (Gubbins, Callahan, and Renzulli 2014). Moreover, although the Javits Act remains intact, it has faced regular challenges since the passage of the No Child Left Behind Act (NCLB) in 2001, which turned attention away from gifted students by imposing school accountability measures based on basic proficiency standards.

Notably, the first step established by NCLB to incentivize improvement among schools failing to make Adequate Yearly Progress (AYP) was not to help provide supplemental educational services (this came in step two), but rather to give students the often trivial "choice" of attending another (AYP-compliant) school in the district. Thus, in addition to ushering in an era of unprecedented emphasis on basic testing standards, NCLB also exemplifies the longstanding American tradition of using "choice" to correct inequalities or inefficiencies in the education system without dismantling the underlying issue. I turn next to desegregation, another defining policy of the latter half of the 20th century, which exhibits important similarities.

Choice as a Solution to Segregation

At the same time that the Cold War was beginning to raise concerns about American excellence and the ability to compete internationally, the Civil Rights Movement was raising concerns about equity at home. In 1954, the Supreme Court decided *Brown v. Board*, holding that separate schools were inherently unequal. This decision meant that school systems, which had previously been legally segregated under the premise that they were "separate but equal," were now mandated to integrate. "School choice" made its first real splash as a strategic conservative response to this decision, based on the logic that if a Black student could *choose* to attend a white school, then the system was not technically separate and the district's constitutional responsibility was fulfilled. However, lack of transportation services and discomfort attending school in a racially hostile environment remained powerful deterrents to integration (Orfield 2013). Thus, school choice functioned as a way to circumvent desegregation, and practical progress towards the implementation of integration stagnated until the Civil Rights Act of 1964.

As progress got underway, battles over what desegregation could and should entail produced a wave of legal battles that crested in the 1970s. In 1973, nearly two decades after *Brown*, the Supreme Court decision of *Keyes v School District No. 1, Denver* ruled that de facto segregation was also unconstitutional. In other words, even districts without codified intent to operate racially separate school systems were subject to court-ordered desegregation if it could be plausibly demonstrated that any law or official practice exacerbated racial segregation in schools. However, this expansion of the purview of desegregation was quickly checked. The next year, in the decision of *Milliken v Bradley* (1974), the Supreme Court ruled that the responsibility to desegregate resided *within*, not between, district boundaries. The practical implication of this ruling was that there could be no legal recourse for labor market or housing policies that spurred white flight and made access to suburban districts prohibitive for Black families.

Nevertheless, a series of desegregation orders were issued by the courts in subsequent years. Two key policies emerged in response. On the one hand, busing emerged as a way to transport students to schools outside their neighborhoods, where students would be assigned in a racially equitable manner. This policy produced a severe backlash and resentment among white parents, including famous protests in cities like Boston (Formisano 2004). A second approach was magnet schools (Wells 1993, Fuller & Elmore 1996). Rather than circumvent the privilege white parents felt they had earned through their residential "choices," as busing did, magnet schools were designed to actively *attract* families through specialized programming. Moreover, although municipalities could not be legally compelled to desegregate across district lines, these programs could be used as a carrot to prevent white flight by providing desired resources.

From its origins as a bad faith offering in the Jim Crow South to a strategy for school innovation, school choice has gained traction as a lever of educational equality in the United States. To have real potential as a carrot for families who may otherwise exercise choice through exit, however, these options must also foster excellence. I turn next to the question of whether and how SEP schools have been shaped by the intersection of this dual legacy of separate schooling in the United States.

Situating SEP Schools

To date, information on the motivation for and establishment of SEP schools still primarily resides with the local districts and communities they serve. The only available national data[8] suggests that some of these schools have been selective since at least the early 20th century, but that the plurality are likely to have been founded in the wake of the Civil Rights and gifted education legislation just discussed (Finn Jr and Hockett 2012). While it is certainly the case that all of these schools aim to attract high-ability students, and likely the case that many districts saw these schools as strategic incentives to retain white and wealthy families, it is not necessarily the case that these schools were established with the explicit aim of responding to court ordered desegregation.

In terms of their place in the history of gifted education, I am aware of no explicit recommendations by the NRC/GT for or against SEP schools. However, under the umbrella of improving identification, at least one study funded by the Javits Act evaluated admissions procedures to selective high schools. Specifically, this 1993 study evaluated admissions at nine state-run residential STEM high schools, which had opened in the preceding decade (Jarwan and Feldhusen 1993). This timeline suggests that some selective schools may well have been founded in response to the publication of *A Nation at Risk*. Importantly, although the study—in keeping with the stated priorities of the Javits Act—included among its recommendations the need to

[8] These data are limited by low response rates and lack of clarity between when schools were founded and when they became selective.

"continue efforts to find qualified minority students and to develop counseling and instructional methods to help them succeed once enrolled," it also argued that selective school admissions is simply a "measurement process" and that the only defensible approach is through standardized scores and/or regression analysis, not professional "judgment" (Jarwan and Feldhusen 1993, p 35). Whether or not any precedent was set by this study is unclear. However, if one were, it would seem to suggest that, while gatekeepers at selective schools should perhaps be responsible to seek more (diverse) qualified *candidates*, a fair admissions procedure should treat these candidates the same once they enter the applicant pool. This is emblematic of the eventual retrenchment of race-conscious admissions in desegregation policies.

If desegregation was ever an explicit goal in SEP schools, it quickly came under attack. For instance, before the *Keyes* decision was even passed, New York City preemptively passed a state law[9] in 1972 that established the Specialized High School Admissions Test (SHSAT) as the sole criteria for admissions to the city's "specialized" high schools—some of the oldest and best-known SEP schools in the country, including Stuyvesant and Bronx Science (*NYT* Editorial Board 2018). Given persistent inequalities that shape students' preparation for the test, this produces predictable inequalities in access to these schools[10]. In other cities with SEP schools, where integration policies were successfully implemented, white parents eventually filed a spate of so-called "reverse discrimination" lawsuits in the 1990s, indignant that their children were apparently being held to higher admissions standards than their Black peers. These cases, which were brought in cities like Boston, San Francisco, and Buffalo, New York, resulted in significant rollbacks to the race-conscious policies that had stimulated integration in the preceding decades (Orfield and Ayscue 2018). Of course, this retrenchment of desegregation policy was not restricted to SEP schools. Rather, these court cases occurred in the larger context of the steady removal of court oversight of desegregation in these districts[11] and throughout the country. Today, these schools operate in a national context where race has been all but eliminated as a valid consideration in individual admissions decisions. At the K-12 level, this was formalized by the Supreme Court's split decision in *Parents Involved in Community Schools v Seattle School District No. 1* (2007), which held that, while districts may pursue racial integration, they cannot do so by individually assigning students to schools on the basis of race.

To summarize, SEP schools have both shaped and been shaped by the history of gifted education and school integration in the United States, resulting in their present condition as "elite, separate, and unequal" institutions. Yet while these are still important concerns among parents, national education policy towards these ends has become much less direct. Rather, the current policy landscape revolves around school choice—as a mechanism for accountability, an incentive for innovation, and a coveted parental prerogative. As a result, the process of sorting students between programs and schools has become increasingly complex. Even in this context, however, SEP schools are characterized by a relatively unique approach to student differentiation among U.S. public schools.

---

[9] Passing this law at the state level makes it harder to overturn, because it requires intervention by the governor (*NYT* Editorial Board 2018).

[10] Black and Latinx students make up about three quarters of New York City's public school population, but only about 10% of students in the city's eight specialized high schools (Corcoran and Baker-Smith 2018; Shapiro 2019)

[11] Boston was released from court oversight in 1987, Buffalo in 1995, and San Francisco in 2005 (Reardon et al. 2012).

Elite: The Experience of SEP Enrollment

This section will review the literature that exists about SEP schools. While rigorous, this research has focused on relatively narrow questions around (1) characteristics of SEP schools themselves and (2) their effects on enrolled students. Given the district-level process of differentiation that shapes the student body in these schools and the traditional public schools throughout their districts, I will argue that SEP *districts* merit further attention because of the potential for both intended and unintended spillover effects.

High Standards for All?

To the extent that a national image of SEP schools exists, it is likely to reflect the character of SEP schools in New York City and Boston, simply by virtue of the broader circulation of these cities' "local" papers—the *New York Times* and the *Boston Globe*—which occasionally feature coverage of these schools and their admissions practices (e.g., Ali and Chin 2018; Ebbert 2016; Finn Jr 2012; Gay 2019; Irizarry 2017; Rey 2017). Through these editorials, we learn about the oldest and most elite public schools in the United States. But, even among SEP schools, Stuyvesant and Boston Latin are probably unusual.

One study, which attempted to compile a list of SEP schools at the high school level, finds that these schools are in fact not uniformly high performing (Finn and Hockett 2012). They identify 165 schools[12], of which 143 received ratings from Great Schools[13]. Of these schools, 85 earned the top score of 10 out of 10, while 12 earned a rating of 5 or less. Based on brief observations in eleven of these schools[14], Finn and Hockett argue that teachers and students in these schools demonstrate a general commitment to rigorous college preparation[15], but that they enjoy different amounts of resources with which to do so. However, although these schools may not uniformly rank among the most elite educational spaces in the country, they are almost necessarily relatively high-achieving *within their district*. Rather than comparing these schools to the average U.S. public school (as Finn and Hockett do), this suggests the importance of evaluating SEP schools relative to other schools in their district.

Beyond the question of whether these schools are objectively "elite" is the question of whether attending an elite school is good for students. Although Finn and Hockett themselves do not examine *effects* of the schools they study, research has long complicated the notion that elite schools produce positive outcomes for their students in terms of academic self-concept (Marsh 1988; Marsh and Hau 2003), attainment (Attewell 2001; Davis 1966; Espenshade, Hale, and

[12] Finn and Hockett's definition of exam schools (which differs slightly from the definition of SEP schools used here) includes charters, university laboratory schools, and state-run schools, which are not rated by Great Schools.

[13] Great Schools is a nonprofit online resource designed to provide information to parents about school quality. Their website suggests that they have recently improved their rating scheme from an initial system based primarily on test scores (likely the system in place at the time this resource was used by Finn and Hockett) to one that includes a broader range of criteria (https://www.greatschools.org/gk/ratings/).

[14] Finn and Hockett do not make claims to have randomly sampled from their list of schools and only observe each school for a maximum of two days.

[15] A college preparatory curriculum was a condition of selection for Finn and Hockett's list of exam schools.

Chung 2005), or achievement (Abdulkadiroğlu et al. 2014; Dobbie and Fryer Jr. 2011). By definition, only a small proportion of students at any school can be ‘top of the class’. Research shows that occupying lower positions in elite institutions—or being a ‘small frog in a large pond’ (Davis 1966) can have negative consequences for high achieving students. For example, (relatively) lower-performing students in extremely high-performing settings are less likely to take AP classes in high school (Attewell 2001) and tend to apply to lower-prestige jobs after college than students of similar demonstrated ability in average institutions (Davis 1966). This, of course, is in addition to the fact that class rank itself can be of structural importance, as it is often considered during the admissions process at selective universities (e.g. Stevens 2007). As a result, the benefits of attending “star” schools[16] reflect a winner-take-all phenomenon whereby benefits may only accrue to a very small number of students (Attewell 2001; Frank and Cook 1995).

Beyond the Margins: Regression Discontinuities and Remaining Questions

More recently, scholars have conducted rigorous studies of the achievement effects of public “exam schools”[17] in Boston (Abdulkadiroğlu et al. 2014), New York City (Abdulkadiroğlu et al. 2014; Dobbie and Fryer Jr. 2011), and Chicago (Allensworth et al. 2016), and grammar schools in the UK (Clark 2007). Using regression discontinuity designs, these studies find no important effect of SEP schools on high school achievement among marginally admitted students. In keeping with the winner-take all phenomenon, these findings suggest that SEP schools may not be doing a great job supporting their lower performing students (or at least not a better job than other schools are doing (Angrist, Pathak, and Zárate 2019)).

It is important to consider, however, that marginally admitted SEP students may be the least likely to benefit from SEP schools, not only among admitted students, but throughout the district. First, the fact that marginal admissions winners do not benefit from SEP schools by no means suggests that the same is true of the top achievers in these schools[18]. Additionally, average-achieving students who get the opportunity to shine in traditional public schools, due to the departure of SEP students, may benefit from their relative star status. Moreover, such an effect would make it less likely to observe a *relative* advantage for marginally admitted SEP students, even if their outcomes do improve. Together, these possibilities raise the question not only of SEP school effects on non-marginal admits, but of spillover effects at other schools throughout the district. This question of the *districtwide* effects of SEP schools, which has not yet been addressed, is where the current study enters.

Separate: Looking Beyond SEP Schools for Implications of High-Scope Differentiation

The literature discussed above suggests that elite schools may not be particularly well designed to impart benefits on all of their students. But why might this be and what implications might this have for SEP *districts*? If there are indeed district-wide implications of SEP schools, through what mechanisms might these effects take shape? Recall that what is unique about these

[16] The research cited here includes, but is not specific to, public schools.

[17] Exam schools are a specific sub-type of SEP schools that utilize exam scores as the sole (or primary) criterion in their admissions process.

[18] Some research has attempted to estimate effects for SEP students further from the margin, and reaches similar conclusions about the weak effect of SEP schools on enrolled students (Angrist and Rokkanen 2015).

schools is the high-scope and academic selectivity of differentiation. Put differently, what makes these schools distinctive is that they are, by definition, *separating* higher achieving students into contexts where they share no interactions or resources with lower achieving students. The importance of student interactions has been studied extensively in the peer effects literature. I will argue that if peer effects contribute to districtwide effects associated with SEP schools it is likely by shaping the expectations of teachers and district officials—and the decisions for resource allocation they make as a result—, rather than through direct effects on student competition or motivation.

Separating Students: Peer Effects

The theory of peer effects is based on the notion that the environment in which a student learns is important and that this environment is shaped by the behavior and abilities of the other students in that environment. Evidence of peer effects is quite mixed, largely due to different assumptions about the linearity and homogeneity of effects (for a review of the peer effects literature, see Sacerdote 2011). The linear-in-means model is the canonical peer effects model, and argues that all students, both higher- and lower-achieving, benefit as the average achievement of their peers increases. The boutique model, by contrast, holds that students benefit from being surrounded by similar ability students in homogenous classrooms. This can be thought of as the tracking model, and it suggests a heterogeneous effect of high achieving peers depending on the focal student's ability (Hoxby and Weingarth 2005). Both models predict that students in SEP schools should benefit because their learning environments are made up of both higher- and more similar-ability peers. However, these models have different implications for the effect of high-scope selective differentiation outside of SEP schools.

In principal, sorting high-achieving students into their own schools should produce a student body in traditional public schools that is both (1) lower-achieving on average, and (2) more academically homogenous than if high-achieving students were to remain in neighborhood schools. The linear-in-means model would therefore predict a negative peer effect, based on the lower average achievement in these schools, while the boutique model would predict a positive peer effect, as a result of the greater homogeneity. In practice, however, individual neighborhood public schools may only lose a handful of students to SEP schools, meaning perhaps just a student or two—or none at all—per classroom. Rather than a significant change to the average achievement in a student's peer group, then, the provision of SEP schools may more likely be associated with the loss of a "shining light" from traditional public school classrooms. Hoxby and Weingarth (2005) theorize "shining light" peer effects as a single outstanding student who can serve as a positive example and motivation for other students, thereby increasing average achievement, or, presumably, decreasing achievement by their absence. However, they find no support for this phenomenon (Hoxby and Weingarth 2005).

In order to assess whether any of these spillover peer effects might plausibly occur as a result of SEP schools, it is important to consider how exactly they might operate. I have suggested that a significant change to the composition of a student's peer group in traditional neighborhood schools may be unlikely. This means that a *direct* effect of average peer achievement seems an unlikely candidate to produce a significant effect. Importantly, theorists argue that peer effects may also operate indirectly by, for instance, allowing teachers to deliver more narrowly tailored curricula (Sacerdote 2011). While minor changes in enrollment in neighborhood schools may have little effect on classroom *needs* for different curricular content or delivery, the process of differentiating students may incidentally communicate such needs to teachers, whether they exist or not. In other words, teacher expectations or sense of efficacy may be shaped by district-level patterns of student differentiation, and this could have consequences for student outcomes both in and out of SEP schools.

In combination with findings from the regression discontinuity studies described above (Abdulkadiroğlu et al. 2014; Allensworth et al. 2016; Dobbie and Fryer 2014), this literature raises questions about the district-level effects of SEP schools on student *achievement*. Accordingly, Chapter 5 examines whether SEP districts exhibit different average *levels* of achievement or *inequality* of achievement. Both questions are important because the theories presented here suggest conflicting possibilities for effects on students towards the top and bottom of the achievement distribution, which could have implications not only for average achievement, but achievement gaps as well.

Separating Resources: Teachers and Instructional Tools

Importantly, high-scope selective differentiation of students could also affect teachers in other ways. Generally speaking, the process of teacher allocation is a topic of important concern. Teachers are typically given priority placements based on seniority, and these teachers tend to seek relocation into higher-income and higher-performing schools. This results in less experienced teachers being placed into schools with the greatest needs (Scott Krei 1998). It is quite plausible that SEP schools would constitute the highest prestige placement in their districts. If SEP districts narrow the number of schools seen as desirable placements, this could have implications for the proportion of students throughout the district with access to highly effective teachers. Finn and Hockett (2012) do indeed find preliminary evidence that teachers in SEP schools tend to have higher qualifications than the average public school teacher. This could be consequential, as studies estimate that differences in "teacher effectiveness" can account for 7-21% of the variance in student achievement, and that moving a student from a teacher at the 25th percentile of effectiveness to one at the 75th percentile is associated with over a third of a standard deviation gain in reading and nearly half a standard deviation in math (Nye, Konstantopoulos, and Hedges 2004).

Of course, districts distribute other resources between schools as well. In the U.S., funding is distributed on a per pupil basis, but these dollars may buy different resources in different schools. For instance, while one school may invest in tools for its vocational training, another may renovate its chemistry laboratories. A district with limited resources may not be able to replicate every resource in each school and may have to make strategic decisions about where to locate specialized resources. On the one hand, this might imply that SEP districts are actually able to provide resources more *efficiently*, by concentrating specialized resources into schools where students have demonstrated the relevant motivation and preparation to take advantage of them.

This potential for greater efficiency applies an economic logic to the structure of schooling. A similar logic is applied to promote current school choice policies, namely vouchers and charter schools. Proponents of these policies argue that increased competition will improve student outcomes in traditional public schools by motivating these schools to innovate the content or delivery of their curricula in order to continue to attract students. It is not clear, however, whether the proliferation of options increases efficiency in resource allocation. Charter schools, for instance, have been shown to increase fiscal stress in public districts (Arsen and Ni 2012; Cook 2018), but evidence that districts respond to this stress by reallocating resources is mixed. Some research finds very little effect of charter competition on resource allocation in general (Arsen and Ni 2012), while others find evidence of an increase in capital construction (Cook 2018). Neither find evidence of an increase in instructional (i.e., teaching) expenditures.

Charter schools, of course, differ from SEP schools in important ways, which have important implications for whether they would exert similar pressures on the allocation of instructional resources. First, SEP schools do not drain per pupil funding from the district as a whole. In fact, if they are able to attract wealthier families to stay in the district, they could improve public school funding by increasing the property tax base. This could imply that SEP schools have a positive influence on the overall level of district resources, irrespective of their relative distribution to different schools. On the other hand, there could be perverse incentives to concentrate advanced academic resources in SEP schools to draw wealthy families away from private schools or wealthier public districts. Second, charter schools are typically barred by law from using academically selective admissions procedures. Competing with charter versus SEP schools may therefore entail responses to different elements of instructions. To compete with charter schools, traditional public schools may be attuned to things like longer school days, unique instructional styles, or even "no excuses" discipline policies, while competition with SEP schools would seem to imply a narrow focus on rigorous academic offerings. It is possible, therefore, that SEP schools present more direct pressure for the "academic-oriented activities" for which Arsen and Ni (2012) found little evidence of pressure from charter competition.

Put simply, differences in resource allocation in SEP districts could be either a feature or a bug of high scope selective differentiation. In either case, it is important to understand what these differences might be. In Chapter 6, I focus on the distribution of college preparatory resources specifically, and evaluate (1) whether AP courses might arguably be distributed more efficiently in SEP districts, and (2) what effects this has on resources available to the district's *average* student.

Unequal: Selectively Exacerbating (Racial) Inequality

Significant as the implications of selective differentiation may be, it is not the separation of students by *academic performance* that draws the greatest attention to SEP schools. Rather, researchers and stakeholder focus on the fact that academically selective admissions procedures also separate students by race and socioeconomic status. Accordingly, this section will offer a brief overview of existing research on the exacerbation of segregation through selective assignment procedures, drawing on recent research on SEP schools specifically, as well as the longer history of research on more common forms of selective differentiation, namely tracking. While the prior section raises questions about the implications of SEP schools for (1) student achievement and (2) the distribution of resources to students throughout a district generally, this section suggests the need to approach these questions with additional consideration for inequalities, particularly racial disparities, as well as the need to consider (3) the implications for between-school segregation in SEP districts.

Racialized Gatekeeping and the Guise of Meritocracy

Above, I described the 1990s-era court cases, which raised concerns around the meritocratic legitimacy of considering a student's race in SEP school admissions. These cases, which appeared in cities like Boston, San Francisco, and Buffalo, NY, resulted in the retrenchment of integration efforts and the removal of individual student race as a factor in admissions decisions. With this reversal in policy has come a reversal in concerns around the (lack of) consideration of race in SEP school admissions. Today, local news coverage in these cities and elsewhere highlights the overwhelmingly white and Asian student bodies in SEP schools (Ebbert 2016; Irizarry 2017; Rey 2017; Tucker 2018). In Buffalo, these disparities sparked a federal civil rights complaint in 2014, which resulted in a formal investigation by the Office of Civil Rights (Orfield and Ayscue 2018). In New York City, Mayor de Blasio has challenged the exam-only admissions policy for the city's eight specialized high schools with the goal of improving access for Black and Latinx students, albeit without success (Shapiro 2019). Although these particular districts faced legal constraints to race-conscious admissions prior to the *PICS* decision in 2007, that ruling presents an additional challenge to designing procedures for more equitable access.

Indeed, findings from available research suggest that proposed reforms are unlikely to significantly improve existing disparities in SEP school access if admissions policies remain race-blind. Corcoran and Baker-Smith (2018) test the potential impact of various academically selective reforms to New York City's exam-only admissions policy for their "specialized" high schools. They test four admission rules that use only absolute achievement measures, including standardized tests and grades, one that includes attendance, and two that use relative achievement by requiring proportional representation at either the borough or sending-school level. Only the admission rule that selects students from the top 10% of each feeder middle school would substantially improve the representation of Black students and students eligible for free lunch in specialized high schools. This suggests that, although the SHSAT likely increases inequality somewhat over other academic criteria—perhaps by producing incentives for families with means to invest large amounts of time and money preparing for the test (Shapiro 2019; Yin 2017)—the root of the inequality in access appears to derive from disparities in information and preparation. In fact, additional experimental research on *non*-SHSAT high school admissions in New York City finds that students are more likely to match to higher-performing schools when they have better information[19] about their options (Corcoran et al. 2018).

Importantly, while the extreme competitiveness of the SHSAT may not necessarily translate to other contexts in which SEP schools are offered, the importance of disparities in information and preparation almost certainly does. For instance, decades of research on inequalities in tracking in the United States finds significant racial disparities in track placements. These inequalities are typically attributed largely to differences in socioeconomic status and preparation prior to differentiation, rather than to direct effects of race on placement (Gamoran 2010; Lucas and Berends 2002; Lucas and Gamoran 2004). Of course, this still means that Black and Latinx students face greater barriers to higher tracked classes, by virtue of their lower average socioeconomic status and the systemic inequalities that limit their preparation to and through early grades. However, research also shows that white and wealthier parents are more likely to advocate for higher placements for their child (Lewis and Diamond 2015) and that Black students may be vulnerable to "crowding out" by white students in racially diverse schools (Lucas and Berends 2007). Indeed, these kinds of processes are perfectly in keeping with white parents' agitation for access to SEP schools, which resulted in the legal removal of race from consideration in admissions in the first place.

It is important, therefore, to consider whether students in SEP districts experience different levels of segregation than students in non-SEP districts. This is the topic of Chapter 4.

[19] This does not necessarily mean *more* information, but rather targeted and digestible. However, even these targeted interventions may disproportionately benefit comparatively advantaged students.

Additionally, given these racial disparities in access to SEP schools themselves. Chapters 5 and 6 consider whether districtwide effects of SEP schools in terms of achievement and access to college preparatory resources differ for students of different racial and ethnic backgrounds.

Summary

In this chapter, I have offered a theoretical definition of Selective Enrollment Public schools based on the selectivity and scope with which they differentiate students. Specifically, SEP schools are whole schools in traditional public school districts that admit their entire student body using academically selective admissions criteria. I argue that selectively differentiating students between schools may result in concurrent differentiation of (1) expectations, (2) teacher effectiveness, and (3) resource allocation. This raises important questions about the effect of SEP schools not only for enrolled students, but for districtwide effects on student achievement and college and career preparation. These effects could be positive or negative. On the one hand, SEP districts may be better organized to deliver targeted curricula and resources, or to inspire innovation through school competition. On the other, key resources may be available to a smaller number of students, circumscribing the potential of students who do not demonstrate the relevant motivation or preparation at a relatively early stage in their educational careers. Moreover, I have established the potential for SEP schools to exacerbate racial inequalities in educational opportunity by differentiating students on a "meritocratic" basis that fails to account for systematic disparities in information and preparation. This is consequential, because research also shows that tracking widens the gap between top and bottom achievers (Gamoran 2010; Lucas and Berends 2002).

Thus, this chapter raises several important empirical questions about the impact of SEP schools on the entire district in which they operate, which I refer to as SEP districts. In the following analyses, therefore, I ask three key research questions, each of which examines implications for racial inequality. First, do SEP districts exhibit different patterns or trends of within- or between-district school segregation? Second, do SEP districts produce different overall levels of achievement growth and inequality of achievement? And, finally, are college preparatory resources distributed more narrowly in SEP districts and, if so, could this be characterized as greater efficiency? Before turning to these analyses, I will first introduce the original data that make it possible to address these questions for the first time.

III. ORIGINAL DATA COLLECTION: PROCEDURE AND PRODUCT

To date, data limitations have prevented analysis of the questions posed in this study. No existing dataset on the universe of U.S. public schools gathers data on selective assignment procedures. Thus, I use an innovative web scraping technique to produce an original dataset of SEP schools and the SEP districts that house them throughout the United States. Throughout the study, I combine these data with national datasets from the Common Core of Data (CCD), the Civil Rights Data Collection (CRDC), the Education Demographic and Geographic Estimates (EDGE), and the Stanford Education Data Archive (SEDA). Together, these data provide the unique ability to assess the role of SEP schools in shaping districts' overall ecology of educational opportunity. In this chapter, I will describe the steps taken and the data produced as part of my original data collection procedure. Secondary data sets will be introduced as they are used in subsequent chapters.

Web Scrape

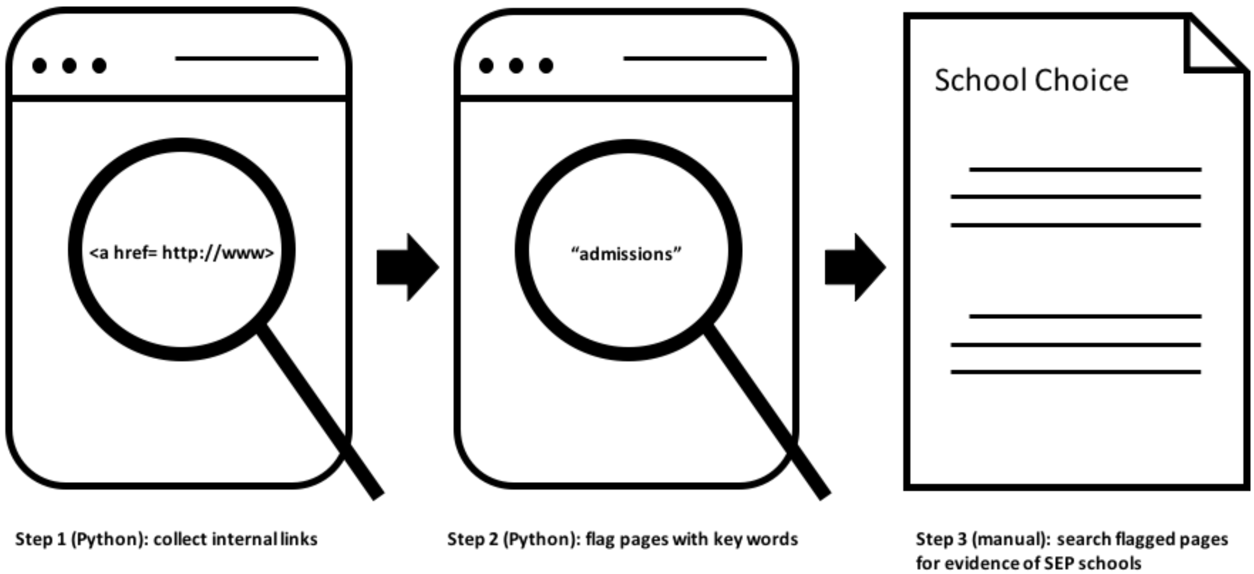
No formal list of Selective Enrollment Public schools exists. Information regarding public school admissions criteria is not collected by the Common Core of Data (CCD) or Civil Rights Data Collection (CRDC), and the one national survey that collects this information—the Schools and Staffing Survey (SASS)—has a small sample and insufficient detail for my purposes. Finn and Hockett (2012) begin the work of developing such a list, but only include schools that grant high school diplomas. Thus, I collected original data and compiled my own list of SEP schools spanning the full K-12 range. To do this, I used the programming language Python to conduct a systematic search for SEP schools by scraping public school district websites for information regarding admissions procedures.

The United States operates over 13,000 public school districts. The majority of these districts are very small, operating one or two primary schools, one high school, and serving fewer than 1,500 students. Districts of this size cannot support high-scope (between-school) differentiation. To make the web scrape both more feasible and more efficient, it was therefore important to limit the list of school districts whose websites I would search. Thus, I cast my original net to include the 400 largest school districts in the country, which serve at least 20,000 students. At this size, K-12 districts would serve about 6,000 high school students. The average high school in these districts serves about 1,200 students, which means that this net easily captures districts with as few as 5 high schools. Given that elementary schools tend to be smaller than high schools, districts of this size are large enough, a priori, to support specialized schools at both the primary and secondary level. I then tested for saturation of my search in two ways. First, I expanded my search to 86 additional districts with as few as 12,000 students but with at least one magnet school (as a signal of a favorable policy environment towards between-school differentiation). Of the remaining unsearched districts, only 3% operate more than three high schools, making it quite unlikely that these districts offer SEP schools. Nevertheless, to provide further confidence in my search, I conduct a second saturation check by drawing on lists of member institutions from the International Baccalaureate (IB) organization and the National Consortium of Specialized STEM Schools (NCSSS).

Before beginning the formal web-scrape, I manually searched through approximately 50 district websites to develop a sense of the language used to describe and provide information about SEP schools. This process generated a list of key words associated with student differentiation, which were used to flag potentially relevant information during the web scrape. With these terms established, the web scrape proceeded in three steps, illustrated in Figure 3.1.

Figure 3.1. Web Scrap Procedure

Panel A.



Panel B. Key Words for Step 2

academy	early college	optional school
academies	enroll	register
admission	entrance exam	registration
application	entrance test	specialty program
apply	exam	specialty school
attractor program	GATE	specialized program
attractor school	gifted	specialized school
AVID[i]	high ability	test[ii]
choice	lottery	transportation
charter	magnet	
criteria	optional program	

[i] All website text was imported and searched as lower case characters, so in actuality I searched for "avid" and "gate." I present these as upper case acronyms in the list for clarity.
[ii] For "test," I included a space prior to the word (" test") to avoid catching words ending in "test," i.e., contest, latest, protest, etc. I did this for AVID (" avid") and GATE (" gate") as well, to avoid catching the name David in directories, and to avoid words ending in gate, i.e., aggregate, mitigate, obligate, propagate, etc.

First, beginning from the home page of each district or school, I used the Python package BeautifulSoup to collect all of the links on the page and keep all those that were internal to the site. I then used this expanded list of web pages and repeated the procedure of extracting links, excluding duplicates. Second, I take this completed list of URLs, and again use BeautifulSoup to search through them. This time, I search the text on each page for a set of roughly thirty key words, including “admission”, “application”, “gifted”, “entrance exam”, etc., developed during the initial manual search (see Figure 3.1 for full list of keywords). Finally, I export the list of URLs for the pages with at least one key word, and search through these web-pages for evidence of SEP schools. During this step, I extracted over 400 documents, including primarily school choice handbooks, statements of admissions procedures, and actual application materials, and qualitatively coded this information, as well as each relevant webpage, for evidence of SEP schools.

As I combed through the district web pages that were flagged by my scraping procedure, I referenced Finn and Hockett’s (2012) list of exam schools[20], as well as schools reported by the CRDC to enroll at least 75% of their students in gifted programming. This served to draw my attention places where my scrape may have failed, and triggered me to make adjustments to the program. After fine-tuning my search procedure using the original list of 400 districts serving at least 20,000 students, I identified 87 districts with at least one SEP school. Following this, the saturation checks for medium size (12,000+) magnet districts and IB/NCSSS schools produced only an additional three SEP districts to add to my list.

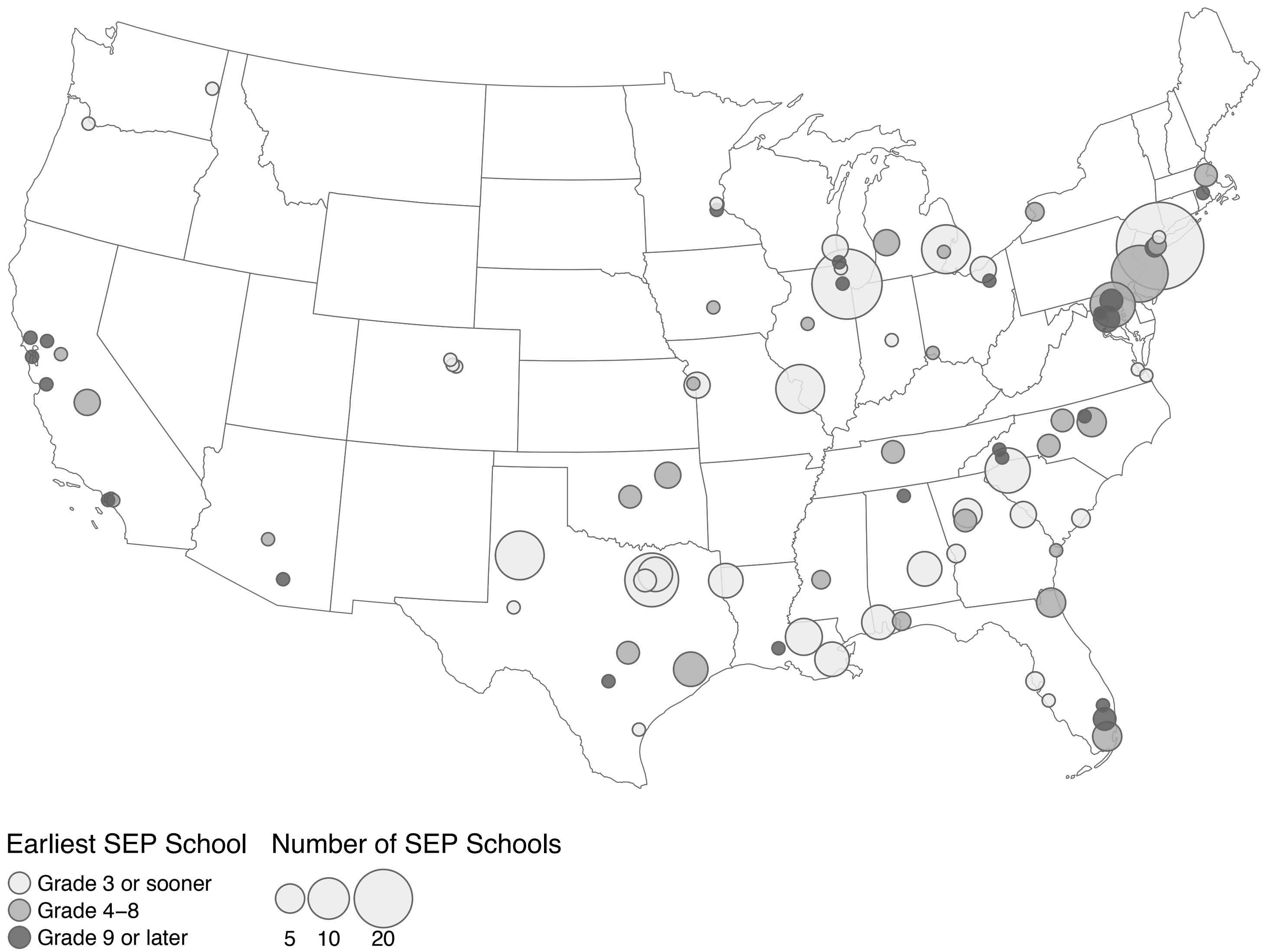
Where and When are SEP Schools Available?

In total, this approach resulted in the identification of 369 SEP schools at all grade levels across 90 districts. The location of these districts is mapped in Figure 3.2. The first thing to note is that, although SEP schools have been studied almost exclusively in New York City, Boston, and Chicago, they are not unique to the northern region of the country. In fact, they are especially prevalent in the South. Additionally, they are neither unique to nor ubiquitous in major cities. Although most of the largest cities in the country do in fact offer SEP schools, many do not. For instance, cities like Los Angeles, Phoenix, Fort Worth, and Columbus do not have SEP schools[21]. Meanwhile, midsize cities like Cleveland, Corpus Christi, and Buffalo, NY do offer SEP schools. The exception to the variation in the location of SEP districts is their noticeable absence throughout the Great Plains and Mountain West. However, this is not surprising. Districts in these regions often cover large but sparse geographies with very few schools, which significantly limits the feasibility of between-school differentiation. There are also no SEP schools in Alaska or Hawaii, which are omitted from the map for simplicity. Overall, SEP districts are relatively geographically widespread throughout the United States. This is important because it suggests that there is sufficient overlap in the characteristics of SEP and non-SEP districts to support statistical tests of differences between these districts as a function of the provision of SEP schools.

[20] My list differs significantly from Finn and Hockett’s because (1) they only include schools that offer high school diplomas, while mine includes primary schools, and (2) they include several schools run at the state level or by district consortia, which I omit to facilitate district-level analysis.

[21] However, there are SEP districts very close to Los Angeles (Anaheim and Long Beach, CA), Phoenix (Mesa, AZ), and Forth Worth (Dallas, Garland, and Grand Prairie, TX).

Figure 3.2. SEP Schools: Location, Level, and Prevalence



In addition to geographic variation, Figure 3.2 uses the size and color of each dot to illustrate the variance in grade level and prevalence of SEP schools. Among SEP districts, the median number of SEP schools is 2 and the mean is 4.2. This discrepancy is due to an extremely large number of SEP schools in a handful of districts, but especially in New York City (*n*=45) and Chicago (*n*=29). Although these districts are outliers in the absolute number of SEP schools offered, they are not particularly unusual in terms of the prevalence of SEP schools relative to the total number of schools in the district. Even so, the existence of more schools likely implies more complicated application and matching procedures in these districts. Overall, only a total of seven districts offer more than 10 SEP schools. It is important to consider, then, that New York City and Chicago—two of only three districts for which much is known about SEP school offerings and effects—may not be particularly reflective of SEP districts in the United States as a whole.

At the district level, the most common grade-levels to begin offering SEP schools are Kindergarten (*n*=22) or ninth grade (*n*=22). The average SEP district begins to offer SEP schools by fifth grade[22]. Altogether, half of SEP schools begin prior to seventh grade. In light of the existing research which focuses on high schools, it is noteworthy that my data identify a near-equal split between primary and secondary SEP schools. In the United States, typical processes for differentiating students look considerably different between primary and secondary school. In primary school, advanced instruction typically comes from Gifted and Talented programs, which are often delivered through pull-out instruction, whereas course-level tracking predominates in high school (National Association for Gifted Children & the Council of State Directors of Programs for the Gifted 2015). Because of these differences in advanced instruction in primary and secondary school, the addition of high-scope selective differentiation at these levels may have different implications for student outcomes. It is therefore important to study outcomes associated with SEP schools at both the primary and secondary level.

How Do They Admit Students?

The extent to which SEP schools shape educational opportunity at the district level may have a great deal to do with how these schools admit their students and the implications of these practices for racial equity in access. To illustrate the variation in criteria used for SEP school admissions, I define six levels of selectivity: schools using (1) screening for gifted and talented identification, (2) ranked admissions with (stated) advanced cutoff thresholds, (3) ranked admissions without a (stated) advanced cutoff, (4) lottery admissions among applications with at least a 3.0 grade point average (GPA), (5) lottery admissions among applicants testing above grade-level, and (6) schools using academic criteria that are selective but less rigorous than those already stated. Advanced cutoff criteria for categories 2 and 3 refer to the same standards that designate lottery admissions as more than minimally selective in categories 4 and 5: GPAs of at least 3.0 or test scores[23] reflecting above grade-level performance. Schools with "minimally selective" criteria, those in category 6, require lower GPAs or may have been missing adequate

[22] The average (mean) grade level for the onset of selective differentiation in SEP districts is 4.7, but the median is 6. Sixth grade is the next most common grade of differentiation (after Kindergarten and ninth) because this is a more common transition grade in the United States.

[23] Required test scores can come from state standardized tests or school-specific entrance examinations. Because proficiency could be determined through standardized tests alone, entrance exams typically appear to be designed for a higher level of selectivity.

information. Schools may use a combination of these criteria. For instance, schools may both require gifted identification and use ranked admissions, or schools may implement a lottery requiring both a 3.0 and grade-level proficiency. Each school is categorized into the first (i.e., most selective) category that characterizes its admissions procedures. In addition, schools in all categories may require application materials such as writing samples, letters of recommendation, or prerequisite coursework. Table 3.1 illustrates the criteria used by SEP schools to admit their students.

Table 3.1. Level of Selectivity

SEP District Grade Level [a]	All Grades	Begins 4th-8th	Includes 12th
Chapter Using Sample	Ch. 4	Ch. 5	Ch. 6
1. Gifted	9.21	10.91	6.31
2. Rank with Selective Minimum	31.44	30.91	33.33
3. Rank without Selective Minimum	14.36	3.64	15.62
4. Lottery with High GPA (>3.0)	17.34	25.45	18.32
5. Lottery with Tests Above Grade Level	4.07	9.09	3.60
6. Minimally Selective	23.58	20.00	22.82
N (Schools)	369	55	333

[a] These grade levels pertain to whole districts, not individual schools. The Chapter 5 sample in the second column includes all SEP districts that offer their first SEP school between 4th and 8th grade, but these districts may include high school SEP districts as well. The Chapter 6 column includes all SEP districts with SEP high schools, but these may also include elementary and middle schools.

Whether or not gifted identification is indeed a more selective admissions requirement than other ranked criteria likely varies between districts. However, as a criterion for SEP school admissions, gifted identification is typically used for elementary school differentiation. To have such a selective sorting criterion for such young students would seem to characterize an exceptionally rigid system of educational differentiation. Overall, only 9% of SEP schools—or 34 schools—require gifted identification among students. Roughly a third of SEP schools require that students meet advanced qualifications to enter the applicant pool *and* use rank-ordered procedures to admit only the *most* qualified students among applicants. A similar proportion of schools (about 35%) admits students either using a ranking without an explicit cutoff (category 3 in Table 3.1) or a high-performing cutoff without a ranking (i.e., a lottery, categories 4 or 5). The remaining 24% of districts are only minimally selective.

Altogether, this suggests wide variation in the level of selectivity among SEP schools. Yet, to date, research has focused on SEP districts with rank-ordered admissions, and particularly those with selective cutoffs[24]. Importantly, 40 of the 90 SEP districts identified here include no schools with this level of selectivity. To provide a complete understanding of the impacts of SEP schools, it is important not only to evaluate their districtwide effects—as has already been argued—but to consider the complete range of schools and districts that make up this sector of public education.

Who Do They Serve?

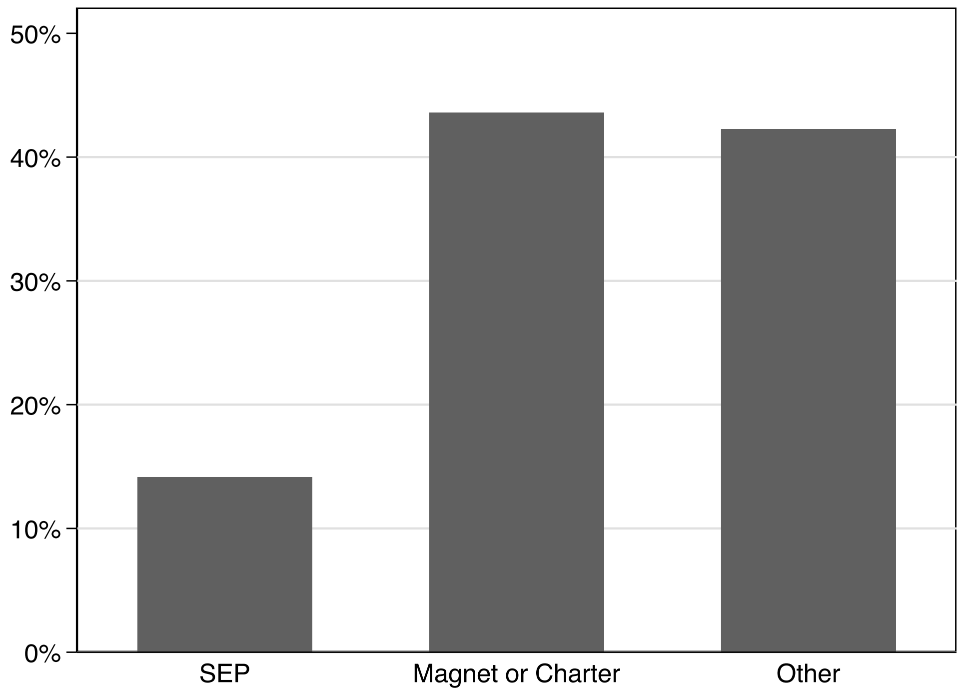
The foregoing sections describe a multifarious school sector—one which spans the K-12 system, both geographically and by grade-level, and which employs a broad range of selection criteria. Even so, there are only 369 SEP schools, located in just 90 SEP districts out of over 13,000 public school districts in the country. At first glance, then, it might seem that SEP districts and schools are not sufficiently prevalent to warrant significant attention from researchers. However, by considering prevalence slightly differently, we can see that SEP districts are indeed an important piece of the public school system in the United States. Panel A of Figure 3.3 shows that, overall, about 14% of public school students throughout the U.S. attend school in a SEP district. This is compared to 44% of students who attend a district that offers at least one magnet school (but no SEP school) and 42% of students who attend school in districts with none of these options. Thus, although other district configurations are certainly more common, SEP districts do impact the education of a significant proportion of American students.

Moreover, while SEP districts serve a significant proportion of the U.S. public school population, they serve an even larger share of public school students from underrepresented minority (URM) racial/ethnic backgrounds. This is illustrated in Panel B of Figure 3.3. Whereas only 14% of total public school students attend SEP districts, 22% of URM students attend SEP districts. In addition to serving a significant proportion of the total population, SEP districts may therefore exert a disproportionate influence on racial inequalities in education. In principal, if the provision of SEP schools helped to improve opportunities for URM students, then SEP districts might mitigate racial inequalities in educational outcomes. Such an expansion of opportunities would presumably require that URM students have at least equal access to SEP schools themselves. In practice, however, researchers and stakeholders have often pointed out that access to SEP schools is unevenly distributed to students of different racial/ethnic backgrounds, making it more likely that SEP districts exacerbate inequalities.

[24] New York and Chicago, which offer the most SEP schools in the country, include both highly selective and less selective rank admissions SEP schools. These ranking procedures are important to the identification of treatment effects in the regression discontinuity designs applied to these districts (and Boston). The fact that these admissions procedures are less common in other districts may have further ramifications for the broader policy implications of findings from this research, beyond those discussed in the preceding chapter.

Figure 3.3. Public School Student Coverage

Panel A. Percent of U.S. Public School Students Enrolled in Each Type of District



Panel B. Percent of U.S. URM Public School Students Enrolled in Each Type of District

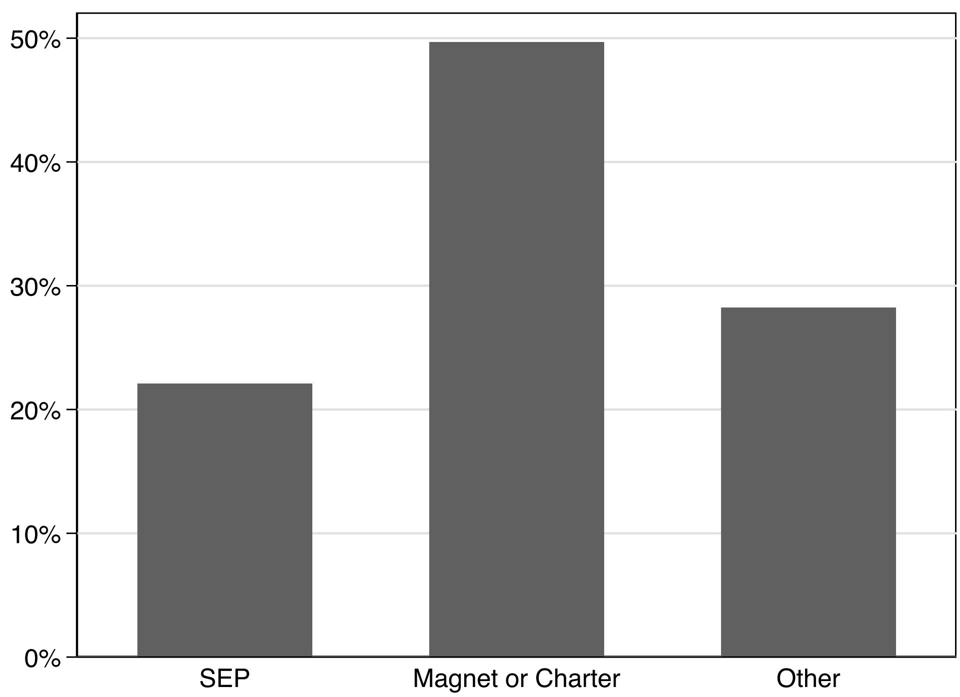
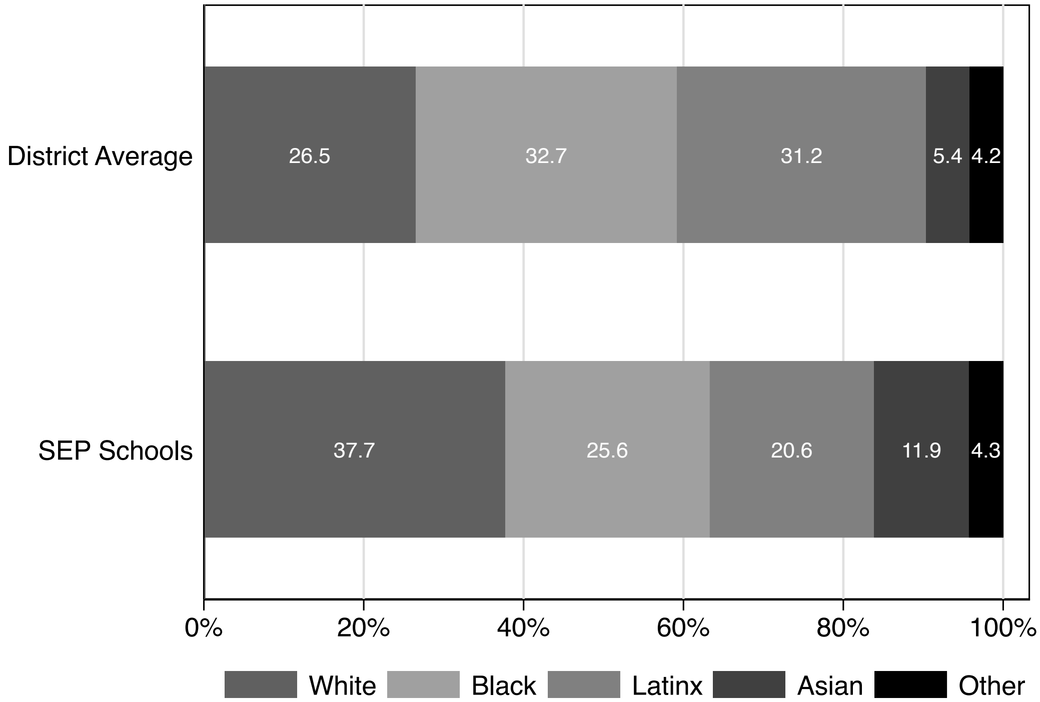


Figure 3.4. Racial/Ethnic Composition of SEP Districts



Up to this point, racial inequality in access to SEP schools has typically been assessed one district at a time. The new data presented here confirm that racial disparities in access exist in SEP districts throughout the United States. Figure 3.4 represents (1) the racial composition of the average SEP district and (2) the composition of SEP school enrollments in that district[25]. According to this figure, the average SEP district serves a student population that is only 26.5% white, yet 37.7% of students who gain admission to SEP schools in these districts are white. Asian students are even more over-represented, making up 11.9% of SEP school students, relative to only 5.4% of the total district student body. At the other end of the spectrum, this means that Black and Latinx students tend to be under-represented in SEP schools, relative to

[25] In other words, these are unweighted district-level averages (i.e., $Enroll_R = \sum_{i=1}^{n} R_i/T_i$, where R_i and T_i are the number of students of race R and the total enrollment, respectively, and i indexes SEP districts from 1 to 90). The fact that this is an unweighted district average is important because larger districts like New York and Chicago also serve a greater proportion of minority (particularly Black) students, such that calculating average school-level enrollments would make Black students appear over-represented. Note that this means the latter bar is *not* equivalent to the overall composition of SEP school *students* throughout the United States. Rather, it is the racial composition across SEP schools in the average SEP district, with each district weighted evenly regardless of size ($Enroll_R^{SEP} = \sum_{i=1}^{n} R_i^{SEP}/T_i^{SEP}$, where R_i^{SEP} and T_i^{SEP} are the number of students of race R and the total number of students enrolled in SEP schools in district i, respectively). The comparison of these district-level averages reflects the over/under-representation of students in SEP schools relative to their home district.

their share of the population in SEP districts. In the average SEP district, Black students comprise only 25.6% and Latinx students only 20.6% of students in SEP schools, compared with 32.7% and 31.2% of the district as a whole. There is reason for concern, then, that the provision of SEP schools may contribute to districtwide racial inequalities in educational opportunity. This possibility will be the focus of much of the analysis in subsequent chapters.

Summary

This chapter has introduced the original data that provides the foundation for analyses throughout the remainder of the study. I compiled these data using an innovative web scraping approach to collect a national list of SEP schools and the SEP districts that operate them. In total, I identify 369 schools across 90 districts. Whereas existing research has focused on individual districts (specifically New York, Boston, and Chicago) and/or on secondary schools, I find that these schools are neither exclusive to nor ubiquitous in extremely large districts and also serve students across the full K-12 grade range. Moreover, while SEP schools themselves are quite rare, SEP districts serve a significant proportion of public school students throughout the United States. However, as critics have pointed out, these schools tend to underserve Black and Latinx students. This provides sufficient variation to examine key district-level outcomes, as well as the impetus to examine racial differences in these outcomes.

IV. SELECTIVE SEGREGATION: FIGHT OR (WHITE) FLIGHT?

School choice has been an important policy intervention in the struggle to (de)segregate schools in the United States since *Brown v Board* in 1954. Magnet schools, designed to *attract* families from diverse backgrounds by offering high-quality and desirable specialized program, were formally accepted as a method of school integration via a federal court ruling in 1975[26]. In 1984, the Magnet Schools Assistance Program (MSAP) was developed to provide grant funding to support schools or districts using magnet programs to integrate students. However, as discussed in Chapter 2, the capacity of public school choice to achieve integration was eroded by a series of court rulings in the 1990s and 2000s. Many of these local challenges, including cases in Buffalo, Boston, and San Francisco, surrounded so-called "reverse-discrimination" in admissions to SEP schools. The precipitous dismantling of racial balancing through school admissions culminated in the 2007 Supreme Court decision of *Parents Involved in Community Schools (PICS)*, which ruled that individual student race could not be directly considered in school admissions. This history has spurred a tradition of research that examines the impact of school choice on student segregation. Yet, despite the centrality of SEP schools in many early legal challenges—and the fact that selective differentiation, by definition, produces relatively (academically) homogenous schools—research has yet to consider the association between these schools and segregation.

Importantly, the *Milliken v Bradley* decision (also discussed in Chapter 2), established the precedent that school districts only have the responsibility—and legal authority—to implement policies that redistribute students *within* district boundaries to achieve desegregation objectives. As such, parents dissatisfied with integrated district options can still exercise choice through exit. One important concern, then, is that policies intended to improve integration at the school level might contribute to white flight, thereby increasing between-district segregation. This is potentially consequential because the loss of more advantaged families may contribute to declining district resources, which, in the United States, are levied largely through property taxes. By offering parents an opportunity to bypass what they see as undesirable neighborhood schools, SEP schools might plausibly help to retain better resourced families who would otherwise have opted to exit the district. If this pull is sufficient to increase tax revenues and/or increase parent advocacy at the district level, then SEP schools might contribute to broad improvements to educational resources for all schools, even in the face of within-district segregation.

Accordingly, this chapter asks not only whether SEP schools might contribute to intra-district segregation by selectively sorting students into schools outside their neighborhoods, but also whether the racial/ethnic composition of overall district-level enrollments are associated with the provision of SEP schools. Although income segregation is also consequential, I focus on race for two main reasons: (1) the policy focus from stakeholders and the courts around school segregation has centered on racial segregation, and (2) although data on race are far from perfect, school-level data on student socioeconomic status are even more limited, based on (incomplete) reporting of free and reduced price lunch eligibility. Using data from the Education Demographic and Geographic Estimates (EDGE) and the Common Core of Data (CCD), I estimate measures

---

[26] This was not a Supreme Court ruling, but the Supreme Court effectively held up the decision by refusing to hear the appeal (Yu and Taylor 1997).

of both evenness and exposure among white and under-represented minority (URM) students, both between and within school districts.

Traditionally, studies of segregation evaluate historical trends to assess the impact of policy shifts. Given the lack of historical data on SEP schools, I instead draw leverage from the available demographic data sources. First, just as parents are more sensitive to school district boundaries than non-parents (Owens 2017), we would also expect traditional public school parents (i.e., those who do not or cannot opt into private or charter schools) to be more sensitive to the availability of school choice options than other parents as they make their residential decisions. If white and Black parents have different school choice preferences, or different abilities to act on these preferences, then this may result in different levels of segregation between district with different school choice options (i.e. SEP versus non-SEP districts). Thus, from EDGE, I exploit differences in the demographics of (1) public school children and (2) all school-aged children to evaluate whether SEP schools appear to exert an important influence on the residential choices of parents.

Second, structural school transitions (i.e., grade levels where students age out of one school and *must* transfer to another) produce natural changes in segregation, as schools—and the neighborhood catchment areas they serve—become larger and typically more diverse. Thus, trends in segregation over the course of a K-12 career offer useful variation to examine whether SEP schools are associated with any departures from "normal" levels of segregation. To evaluate within-district (between school) segregation, then, I leverage incremental changes in segregation by grade level, using data from the CCD, to evaluate whether SEP schools shape the allocation of students.

Overall, I find suggestive evidence of differences in segregation between SEP and non-SEP districts. First, I find that the racial composition across school districts is *more* uneven in Core Based Statistical Areas (CBSAs) with at least one SEP district than those without. Moreover, compared to CBSAs without SEP districts, segregation in these CBSAs is notably higher when estimated for public school students only than when estimated for all school aged children. This suggests that parents' evaluations of public school options are especially salient in these areas. However, this is a purely descriptive analysis and cannot establish whether SEP schools are a *cause of* or a *response to* this district-level segregation. Next, in terms of intra-district segregation, I find that SEP schools are associated with both the evenness of school racial composition and the exposure of in-group students to out-group peers at various points in their K-12 career, but particularly over the course of high school. Although there is evidence that segregation is lower in SEP districts for at least some grade levels, combined evidence from multiple measures of segregation suggest that URM students in SEP districts may be disproportionately likely to transfer out of predominantly white schools (such as SEP schools) during high school, thereby increasing segregation between white and URM students over this period.

Before turning to these results, I briefly discuss the importance of segregation to the experience and study of education. I then define the indicators of segregation used in the analyses: (1) the Dissimilarity Index, both for inter- and intra-district segregation, and (2) the Interaction Index, which I estimate both for white student exposure to URM peers and vice versa. Results are presented following a description of the fixed-effects approach to grade-level intra-district segregation.

School Choice and Segregation

Segregation is an important topic of education research for several reasons, namely (1) the connection between segregation and student learning, (2) the connection between school segregation and prosocial outcomes, and (3) the fraught legal history that these concerns have motivated, which continues to shape the landscape of public school choice in the United States. Overall, this research demonstrates neutral to positive effects of integration for educational and social outcomes (e.g., Mickelson and Nkomo 2012). The connection between school segregation and student learning has been a focus among researchers since the Coleman Report (Coleman et al. 1966). This report, which set out to examine resource disparities that might account for achievement gaps between predominantly white and predominantly Black schools, instead found little to explain away the relationship between segregation and student achievement. Yet, despite decades of intervening research, identifying a reliable mechanism for the relationship between segregation and achievement has proven relatively elusive, especially as the process and experience of segregation continues to evolve under new policy regimes (Reardon, Yun, and Eitle 2000; Reardon and Owens 2014). Nevertheless, research continues to identify a positive association between school desegregation and both short-term and long-term educational outcomes (Mickelson 2001). Although these effects are strongest for minority students, whose segregated school contexts are characterized by higher poverty and lower resources, research shows that desegregation also benefits—or at least does not harm—white students.

Research also identifies a positive relationship between desegregated schooling contexts and prosocial outcomes. This "contact theory" (Allport 1954) argues that exposure to diverse peers—under certain circumstances—reduces prejudice. Coming out of this tradition, perpetuation theory suggests that racial composition of one context is (causally) associated with racial composition of the next (Pettigrew and Tropp 2006). Thus, racially integrated high schools help to foster social cohesion and cross-race friendships (Braddock II and Gonzalez 2010; Mickelson and Nkomo 2012) and also reduce racial isolation in later contexts, such as the workplace (Pettigrew and Tropp 2006; Stearns 2010). These are important outcomes for the healthy functioning of a diverse democratic society.

Despite generally positive findings regarding the effects of school desegregation, recent court decisions, most notably *PICS*, limit the ability of districts to achieve racial balance across schools. Thus, despite initial decreases in segregation following the court orders of the 1960s and 1970s, research suggests that segregation has worsened over the last two or three decades, at least by some measures (e.g., Reardon and Owens 2014). Between the *PICS* decision and the expansion of charter schools—which do not share the same emphasis on integration that spurred the early proliferation of magnet schools—school choice is an important consideration in the re-segregation of U.S. public schools.

Of course, the first decision parents make when selecting schools for their children is where to live. Importantly, higher income families have more options to live in communities with higher property taxes and higher-resourced schools. Given racial disparities in income and wealth, this has contributed to growing segregation between districts in recent decades, even as segregation within districts has remained relatively constant (Clotfelter 2004; Reardon and Yun 2001; Reardon et al. 2000). Moreover, research shows that achievement gaps are larger in more (economically) segregated metropolitan areas (Owens 2018). Yet, desegregation plans may only be used to relocate students between schools *within* a school district, not between. Districts must therefore rely on voluntary decisions on the part of white and wealthy parents to locate in diverse districts. In this climate, intra-district choice, and SEP schools in particular, may be an important lever to shape between-district segregation.

The more common school choice model deployed to *attract* parents and reduce segregation is the traditional magnet school. Magnets are generally defined as schools or programs with a specialized curricular focus intended to draw students from throughout the district, rather than a local neighborhood catchment area. As discussed in Chapter 2, magnet schools proliferated in the 1980s in the wake of supreme court decisions in the 1970s (e.g. Keyes, Swann, Milliken), which defined the parameters of de facto segregation and the constitutional bounds for addressing it (i.e., within district boundaries). Since 1984, the Magnet Schools Assistance Program has been continuously funded[27] to support magnet schools, specifically their efforts to promote racial integration, which many (but not all) magnet schools explicitly state as part of their mission.

Despite enjoying wide support, evidence of positive effects of magnet schools—and school choice generally—on school segregation is mixed. Importantly, while magnet schools are often more diverse than their neighborhood counterparts, lower-income (Archbald 1996) and minority students often have limited access to these programs (Quiroz and Lindsay 2015). Evidence suggests that these disparities are not only worse in SEP schools, but that it is unlikely that these disparities could be remedied by changes to admissions procedures as long as they remain race-blind (Corcoran and Baker-Smith 2018). Moreover, while magnet schools themselves may be more diverse than neighborhood schools, they may contribute to segregation throughout the district. Saporito and Sohoni (2006), for instance, find that neighborhood schools are more segregated than the neighborhoods they serve. Specifically, they argue that between-school segregation is exacerbated by school choice options such as magnets and private schools because white students are especially likely to opt out of their neighborhood schools.

Unlike traditional lottery-based magnet schools, of course, SEP schools use academic admissions criteria to admit students. This process, as shown in the previous chapter, favors white students. It is therefore important to consider whether SEP schools exert a disproportionate influence on between-school student segregation. Moreover, by virtue of this non-random selection, which heaps advantage upon already advantaged families in the "fight" for seats in their desired schools, districts may view SEP schools as a stronger incentive than traditional magnets to curb white (or wealthy) flight. Thus, while we would expect SEP schools to increase *intra*-district segregation, we might also hypothesize a (possibly ameliorative) connection between SEP schools and *inter*-district segregation. It is therefore important not only to evaluate the relationship between SEP schools and segregation, but also to consider this relationship both within and between school districts.

Data and Methods

The following analyses evaluate segregation using two-group measures of the Dissimilarity Index and the Interaction Index, measures of exposure and evenness, respectively. Evenness refers to the similarity (or difference) in the distribution of groups between subunits (here, schools within districts or districts within CBSAs), whereas exposure measures the degree of potential interaction between groups (Massey and Denton 1988). In other words, evenness is measured relative to the overall population composition, which means that groups can be "evenly" distributed across subunits whether they comprise a small or large proportion of the

[27] It was most recently reauthorized as part of the *Every Student Success Act* (ESSA) in 2016.

total population. By contrast, if one group comprises a very large majority of the population, then exposure to that group will be high, regardless of whether different groups are distributed evenly. Thus, these two types of indices capture distinct dimensions of segregation. SEP schools might shape each for different reasons. For instance, if SEP schools successfully encourage more white families to enroll in the district, then URM exposure to white students might be higher simply because there are more white students to be exposed to. On the other hand, dissimilarity between schools may increase if white students enjoy significant advantages in SEP school admissions. It is therefore important to consider both exposure and evenness.

For simplicity, I estimate segregation between (1) white and (2) URM students. Under-represented minority students are defined as Black, Latinx, or Native American[28]. I choose to consolidate the URM population because the racial composition of districts differs considerably, such that the salient "minority" population is not consistent across the U.S. Using school enrollment data from the Common Core of Data (CCD) and the Education Demographic and Geographic Estimates (EDGE) of the American Community Survey (ACS), I generate these two-group segregation indices with STATA's *seg* package, written by Sean Reardon[29].

Measuring Segregation

The Dissimilarity Index (D) measures evenness by transforming the deviation of each school or district's proportion of in-group members from the overall in-group proportion, as defined in Equation 4.1.

$$D = \sum_{i=1}^{n} \frac{|p_i - P|}{2P(1-P)} * \frac{t_i}{T} \tag{4.1}$$

where p_i is the proportion of in-group students in subunit i. Depending on the analysis, subunits (indexed from i to n) may either be schools within districts or districts within CBSAs. P is the overall proportion of in-group students across all n subunits. Because I am measuring segregation between only two groups, D takes the same value regardless of whether P refers to white or URM students. Finally, t_i is the count of total students in subunit i and T is the overall unit population total. The Dissimilarity Index ranges from 0 (no dissimilarity) to 1 (maximum segregation) and can be interpreted as the proportion of students who would have to change schools in order to achieve a proportional distribution of white and URM students across schools.

The Interaction Index is a measure of the extent to which one group, x, is exposed to another group, y. It is an in-group-weighted average of the out-group share in each subunit. The "in-group" is the population whose experience of segregation is of interest. Typically, this is the "minority" group. Below, I examine both white exposure to URM students and URM exposure

[28] Asian students are omitted because heterogeneity in this group cannot be sufficiently addressed in the CCD. Whereas East and South Asian students are commonly over-represented in advanced academic contexts, Southeast Asian students are often underrepresented (Kao and Thompson 2003). These groups are not distinguished, so I do not include them in either group. There is, of course, also important heterogeneity among Black and Latinx students, but even Black and Latinx groups that tend to be higher achieving (i.e., Nigerian or Cuban Americans) do not experience similar "model minority" stereotypes that shape the experiences of Asian students.

[29] Specifically, D corresponds to d in *seg* and ${}_xP_y$ corresponds to x.

to white students. The Interaction Index is defined as follows:

$$ {}_xP_y = \sum_{i=1}^{n} \left[\frac{y_i}{t_i}\right]\left[\frac{x_i}{X}\right] \quad (4.2) $$

where x_i and, y_i are the number of in-group and out-group students in school i, respectively, and X is the total in-group student population throughout the district. Depending on the focal in-group, I will refer to ${}_wP_{urm}$ (the exposure of white students to URM peers) and ${}_{urm}P_w$ (the exposure of URM students to white peers) in the analyses below. The Interaction Index can be interpreted as the proportion of out-group students in the average in-group student's school. Like D, the Interaction Index ranges from 0 to 1. However, whereas an index score of 1 reflects maximum unevenness (i.e. high segregation) for the Dissimilarity Index, it reflects maximum exposure (i.e. low segregation) on the Interaction Index. Additionally, even in the two-group case, ${}_wP_{urm}$ differs from ${}_{urm}P_w$.

Analysis Strategy: Leveraging Demographic Variation at Different Levels

The first question of interest in this chapter is whether SEP schools are associated with between-district segregation. Given two key estimation constraints, I offer a simple descriptive analysis of this question. First, districts differ greatly in geographic and population size, and some are very large (i.e. whole counties versus individual suburbs). This not only means that crossing district boundaries constitutes a more drastic move in some CBSAs than others, but also makes it challenging to estimate a significant effect of one (or a handful) of schools on segregation between districts across an entire CBSA. Second, I have argued that SEP schools might be a reasonable tool to combat extreme white flight, making it both a potential *response to* and a *cause of* between-district segregation. Without historical data on the founding of these schools, it is impossible to rule out spurious causation. I therefore present purely descriptive relationships, but augment these by juxtaposing measures for the overlapping populations available from EDGE.

EDGE data are useful because they include estimates for (1) children enrolled in district public schools, as well as estimates for (2) *all* school-aged children that live within the boundaries of the school district, even if they are not enrolled in public schools. I use these data to compare between-district segregation using estimates of D for both public school (D_{enr}) and all school-aged (D_{res}) children[30]. This is useful because, if school choice options are indeed an important factor for parents making residential decisions, then we would expect this relationship to be strongest for parents who actually enroll their children in public schools. We would generally expect segregation to be greater for D_{enr} than D_{res}, but we might expect the relationship between these measures to differ depending on the availability of SEP schools.

Next, to get a better sense of whether and how segregation transforms in concert with the selective differentiation of students, I use school enrollment data from the 2015-16 CCD. Focusing on within-district segregation, I estimate district-level fixed effects models which predict segregation (D, ${}_wP_{urm}$ and ${}_{urm}P_w$) at each grade level, controlling for segregation in the prior grade. These models include only districts with at least 3 schools serving *each grade level* from second to twelfth grade (N=782), so that between-school segregation can be meaningfully

[30] Although public school enrollments can be found in greater detail from the CCD, I retain the EDGE estimates for this analysis to avoid any potential differences in measurement error between the two data sources.

calculated. The number of schools[31] is controlled for, as well as whether any magnets are offered in the district at the given grade-level. These models evaluate whether grade-level changes in within district (between school) segregation differ in SEP districts compared to non-SEP districts.

Generally speaking, we would expect *changes* in segregation to be most pronounced at grade levels that include structural school transitions. By structural transition, I mean that students change schools because they age out of their current school and *must* enroll in a new school, i.e. the transition between elementary and middle school or middle school and high school. In the United States, these are most common in sixth and ninth grade. In addition to a simple shuffling of students, we might expect significant changes—most likely decreases—in measured segregation at these grade levels because schools typically become larger at structural transition points and, therefore, serve students from a greater variety of neighborhoods. If SEP schools are particularly effective at attracting white students (in predominantly minority districts), we might expect these transitional changes to be especially pronounced in SEP districts.

Flight: Between-District Segregation

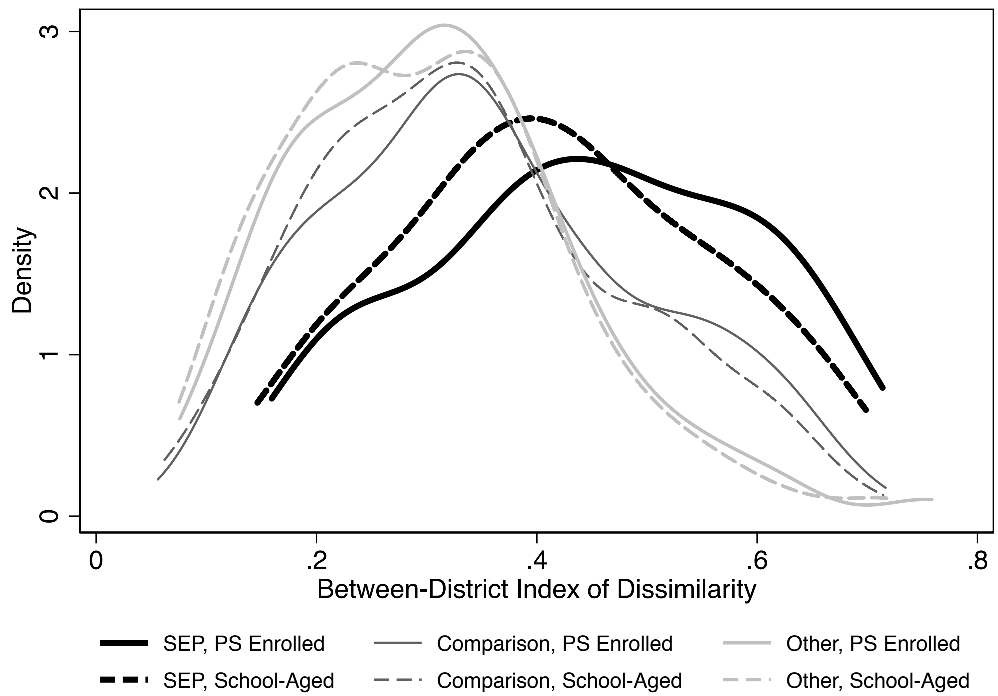
As discussed above, the first decision many parents make in the school choice process is where to live. Before choosing a school, parents must first choose a school district for their child. The literature reviewed above suggests that white parents (consciously or not) tend to avoid schools with high minority enrollments. Therefore, we might expect districts to be more successful in retaining white parents if they offer choice options that allow advantaged families to avoid "undesirable" schools. If this is the case, then, given that SEP districts are typically located in larger and more diverse districts, we might expect CBSAs with SEP districts to demonstrate more *even* racial distribution across districts. On the other hand, if SEP schools are founded in response to exceptionally severe white flight, then we might still observe greater dissimilarity across districts in these CBSAs, even if SEP schools do successfully attract some advantaged families.

Accordingly, this first descriptive analysis examines whether there is any evidence that the provision of SEP schools is associated with between-district segregation, measured using the Dissimilarity Index (D). Panel A of Figure 4.1 makes two key comparisons. First, it compares the distribution of D by CBSA type: CBSAs with at least one SEP district (black), comparison CBSAs (CBSAs with no SEP schools, but at least one magnet or charter school, in medium grey), and other CBSAs with at least 5 school districts[32]. This comparison shows that dissimilarity between districts is *greatest* in CBSAs with at least one SEP district and lowest in CBSAs with no high-scope school choice options. The second comparison made in Figure 4.1 is between estimates of D using enrollment (D_{enr}) versus residential district demographics (D_{res}). The difference between distributions of D_{enr} and D_{res} is most apparent in CBSAs with at least one SEP district. For these CBSAs, dissimilarity is greater when measured using enrollment demographics than using school-aged residential demographics. In other words, segregation is

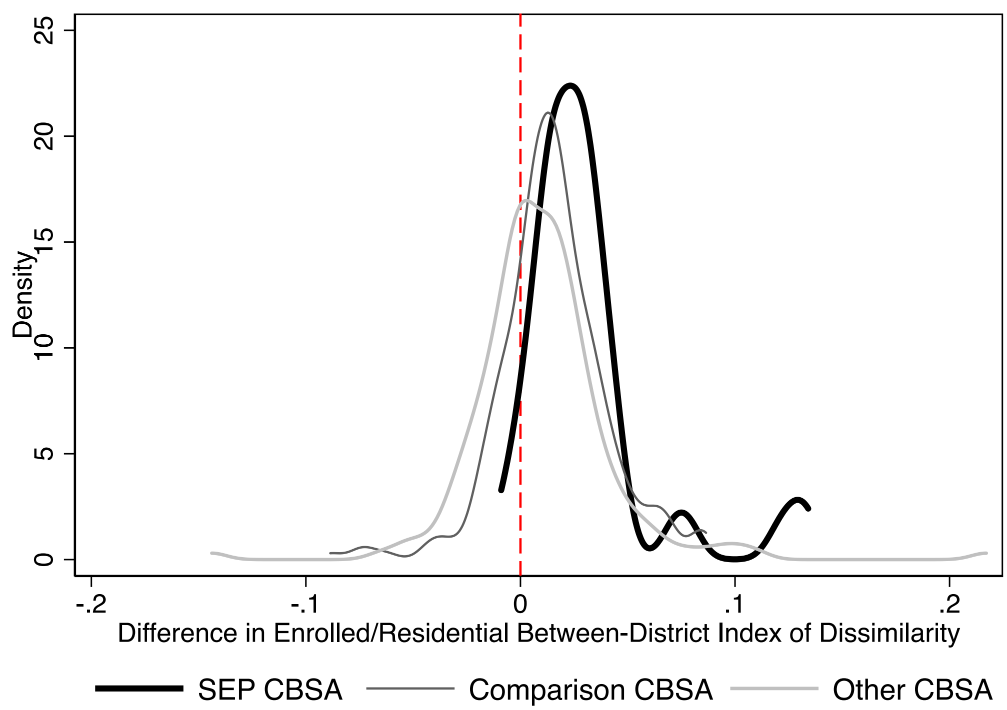
[31] I use a continuous count of schools, capped at 100.
[32] I restrict to CBSAs with at least five districts to ensure a meaningful measure of segregation. This threshold disproportionately drops Southern districts, which more often serve an entire county, rather than individual municipalities (the share of Southern districts drops from 40% to 25% of the sample).

Figure 4.1.

Panel A. CBSA-Level Between-District Dissimilarity Index (D)



Panel B. Difference in Between-District D: $(D_{enr_i} - D_{res_i})$



greater among public school students than among all school-aged children. This implies that white parents, particularly in CBSAs with SEP districts, are more likely to live in racially diverse districts if they can bypass the public school system.

Panel B clarifies this further by making *within*-CBSA comparisons of D_{enr} and D_{res} (i.e. $D_{enr_i} - D_{res_i}$). These distributions confirm that, on average, there is essentially no difference between D_{enr} and D_{res} in CBSAs without high-scope district-level choice, and very little difference in CBSAs with magnet or charter schools. In these CBSAs, it is by no means the case that dissimilarity is always greater between public school students than school-aged children as a whole. In CBSAs with at least one SEP district, by contrast, dissimilarity is essentially always more severe when measured based on district students rather than residents. While Owens' (2017) finds that between-district segregation is more severe for families with children than those without, this finding extends one step further, suggesting that, even when considering only families with school-aged children, segregation is more severe among families who do not (or cannot) opt out of public school. This finding also echoes (at the district-level) Saporito and Sohoni's (2006) finding that segregation is greater between neighborhood schools than between the residents of their catchment areas.

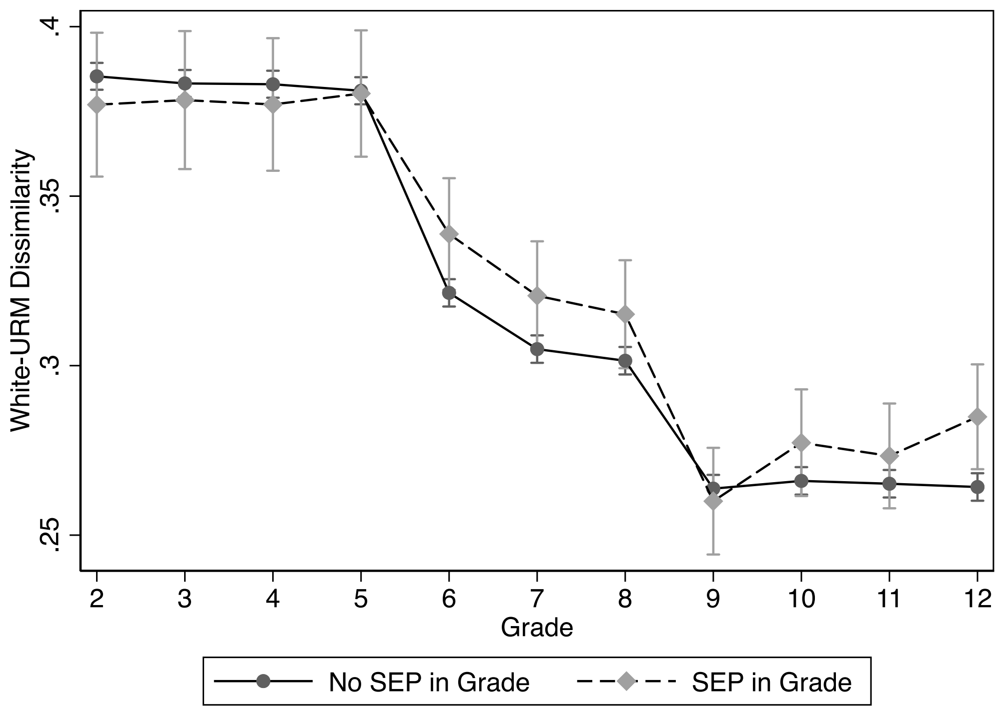
As noted above, these findings are purely associational and cannot establish causal direction. As such, greater observed segregation in SEP CBSAs does not necessarily contradict the hypothesis that SEP districts would reduce between-district dissimilarity within CBSAs. Rather, it could be that severe white flight is a key motivator for the implementation of SEP schools and that the observed dissimilarity would be even more stark in their absence. On the other hand, SEP schools enroll only very few students, meaning that they may be an insufficient incentive to overcome the "risk" white parents perceive in residing and enrolling their children in the large diverse districts where SEP schools tend to be located. Fully adjudicating between these possibilities would require historical data that is not available for the current analysis.

Fight: Within-District Segregation

The foregoing analysis offers important insight into the state of segregation *between* school districts that do and do not offer SEP schools. Desegregation plans in the United States, however, focus on school-level segregation *within* districts. Moreover, although data are not available for a historical analysis of segregation either within or between SEP districts, enrollment data from the CCD provide grade-level enrollments for each school, offering a degree of longitudinal variation. Given that relatively few students transfer between districts during their K-12 careers—but virtually all students change schools—, this grade-level variation offers greater potential to evaluate the connection between SEP schools and school segregation.

Figure 4.2 presents marginal estimates of expected levels of segregation based on regression models with district-level fixed effects, described above. This figure shows results for the Index of Dissimilarity, D, calculated using grade-level enrollment counts from the CCD. As expected, changes in between-school segregation are largest at structural transition years, declining steeply in sixth and ninth grade. There is also a slight decline in segregation in seventh grade, which is the next most common middle school transition grade (accounting for approximately 1 in 5 districts in the analytic sample). These decreases in segregation at school transitions do not differ significantly between SEP and non-SEP districts ($p>0.1$). However, whereas dissimilarity is consistent throughout the elementary *and* high school years in non-SEP districts, it increases throughout high school in SEP districts, such that by twelfth grade segregation is significantly greater in SEP than non-SEP districts. Given that very few districts include an additional structural transition between ninth and twelfth grade[33], it is unlikely that this can be attributed to a systematic reshuffling of students. This raises questions about differences in patterns of high school transfers and/or dropouts between SEP and non-SEP districts.

Figure 4.2. Within-District Dissimilarity



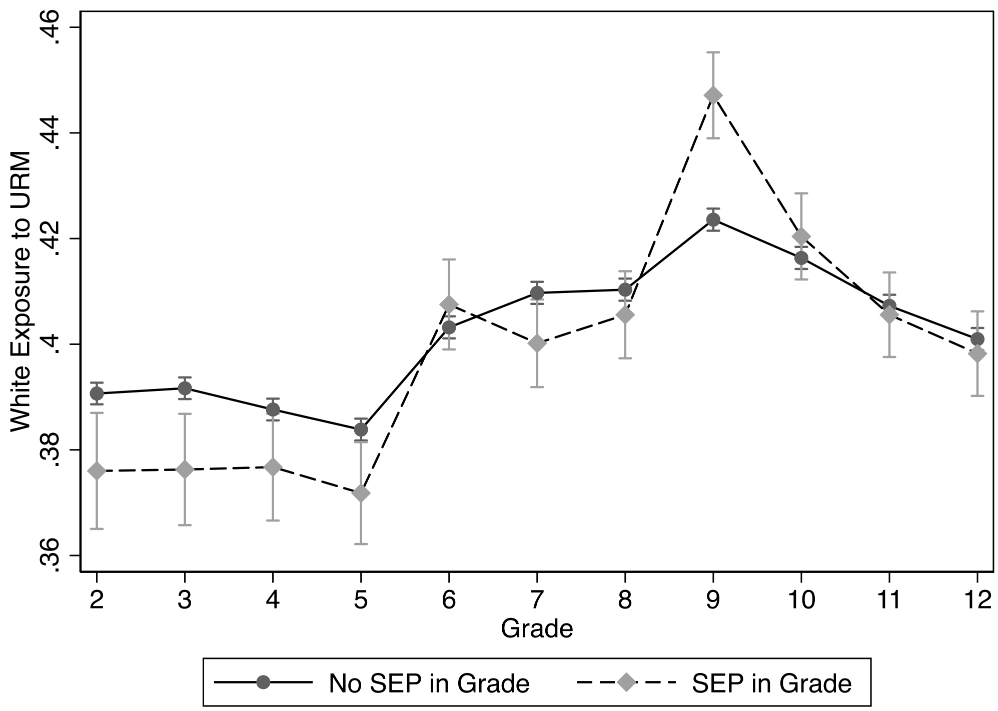
An increase in dissimilarity means that schools become more racially homogenous. In other words, this implies that students are more likely to exit schools where they are in the minority. To produce the pattern observed for high school in Figure 4.2, then, we might expect either that (1) white students in SEP districts are disproportionately likely to transfer or drop out of high school when they attend predominantly URM (likely *non-SEP*) schools, or that (2) URM students who attend SEP (or other predominantly white) schools are disproportionately likely to transfer (or drop) out of these schools before the end of high school.

Next, Figure 4.3 shows results for the same model using the white-URM Interaction Index (${}_{w}P_{urm}$) as the dependent variable. Recall that this index also ranges from 0 to 1, but that, whereas a higher score indicates greater segregation for D, it indicates less segregation for ${}_{w}P_{urm}$. Thus, the significant relative increase in ${}_{w}P_{urm}$ in sixth and ninth grades is consistent with the finding of declining segregation at structural transition years in Figure 4.2. However, there are noteworthy differences in the overall trajectory of these two indices. First, expected

[33] Ninth grade is the modal high school transition year in 99% of all districts throughout the U.S. 88% of high schools in the analytic sample span the full 9th to 12th grade range.

exposure is significantly lower in SEP districts than non-SEP districts in elementary school. Second, the increase in exposure (decrease in segregation) at the sixth-grade structural transition is significantly larger in SEP districts than non-SEP districts ($p<0.05$), such that the difference in average exposure between district types is eliminated in middle school. In ninth grade, there is another significant increase in white-URM exposure in SEP districts, resulting in exceptionally high exposure of white to URM students in SEP districts in ninth grade. Finally, whereas D diverges between SEP and non-SEP districts by the end of high school, ${}_{w}P_{urm}$ converges. Despite an apparent positive trend in white-URM exposure in SEP districts relative to non-SEP districts from elementary school to ninth grade, there is no discernable difference by the end of high school.

Figure 4.3. *White Exposure to URM Peers*

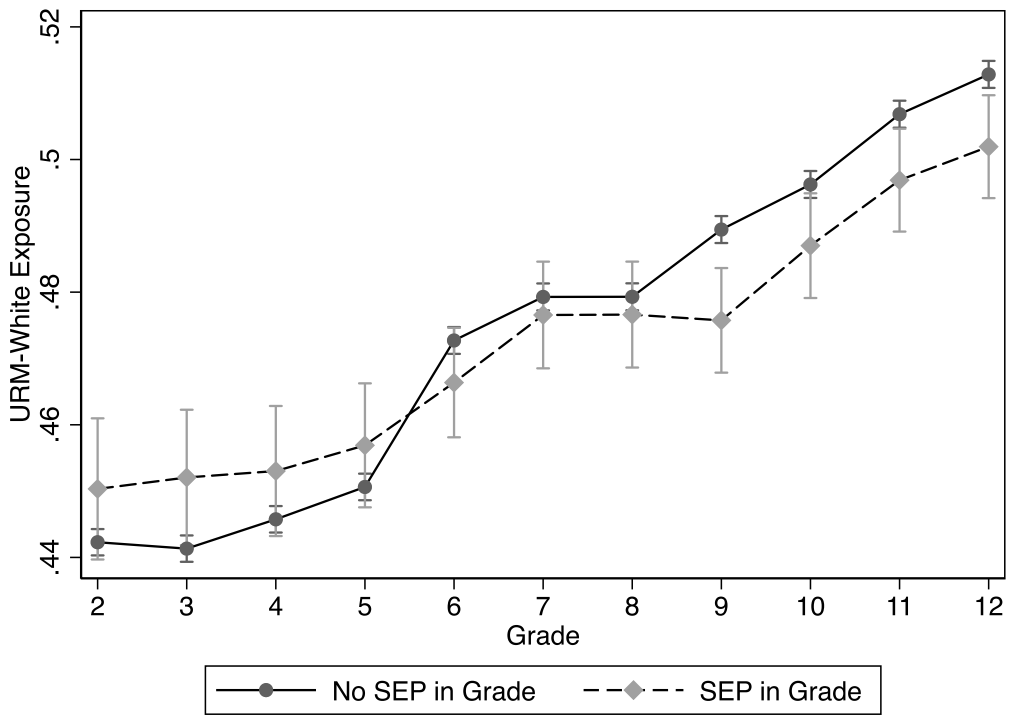


Two possibilities were suggested to explain the divergence of D between district types by twelfth grade in Figure 4.2—(1) disproportionate school exit rates among white students in non-SEP schools, or (2) disproportionate exit from SEP (or other predominantly white) schools among URM students. Importantly, the fact that ${}_{w}P_{urm}$ trends downward throughout high school, rather than dropping precipitously in twelfth grade, suggests that transfers may be more important than dropouts in contributing to the difference in segregation between SEP and non-SEP districts[34]. Nevertheless, because both explanations imply that students become more

[34] CCD enrollment figures reflect the student body on October 1 of the academic year, at which point tenth graders who have not experienced grade retention would not yet be sixteen years old. In other words, tenth graders are generally not of legal age to cease attending school, making

racially isolated, either would still be largely consistent with the decrease in ${}_{w}P_{urm}$ over the high school years. For a final piece of evidence to adjudicate between these possibilities, I replicate the models from Figures 4.2 and 4.3, this time using ${}_{urm}P_{w}$ as the dependent variable.

Figure 4.4. *URM Student Exposure to White Peers*



As discussed above, measures of exposure are sensitive to the population composition. This means that, whereas the two-group Dissimilarity Index is equivalent regardless of whether white or URM students are treated as the focal group, the Interaction Index will differ for white and URM students so long as one group makes up more of the population than the other[35]. The trend in URM exposure to white students (${}_{urm}P_{w}$), therefore, need not parallel the trend in white exposure to URM students. And, indeed, it does not. This is depicted in Figure 4.4. This figure shows that, for both white and URM students, exposure to outgroup peers is higher in middle school than in elementary school. However, key differences emerge in high school, both relative to ${}_{w}P_{urm}$ and between district types. First, the exposure of URM students to white students in SEP districts does not change between eighth and ninth grade. Strikingly, this is the only

dropouts a less likely culprit. Of course, students at the greatest risk of dropout are more likely to have experienced grade retention and be old for their grade. Nevertheless, dropouts are roughly three times as common in eleventh grade as in ninth or tenth grade (Current Population Survey 2018).

[35] One approach to account for this is to use the Normalized Exposure Index, which partials out population share. Findings for this measure (presented in Appendix A) closely approximate grade-level trends for D in Figure 4.2.

structural transition across segregation indices and district types that is not accompanied by a significant change in segregation. Second, this is the only index according to which segregation decreases across the high school years. These findings help narrow the possible processes at work.

Importantly, the increase in URM exposure to white students is not consistent with a disproportionate transfer/dropout of white students from predominantly URM schools in SEP districts, as this would *decrease* the exposure of URM students to white peers. Taken in combination with the trends in ${}_{w}P_{urm}$ and D, it would seem to be that predominantly white high schools in SEP districts, likely SEP schools themselves, become increasingly white throughout high school. Given that SEP schools require a formalized application process for admissions, and typically turn qualified applicants away because of capacity constraints, it is unlikely that transfers *into* these schools are an important factor. Rather, URM students in SEP districts may be disproportionately likely to transfer out of schools where white students are over-represented at the beginning of high school.

Discussion

This chapter presented preliminary analyses to address the question of whether SEP schools are associated with segregation among white and URM students. I extend our understanding from existing research on the impact of school choice on segregation by considering the effect of SEP schools specifically and by considering both inter- and intra-district segregation.

I began by examining the Dissimilarity Index among districts in CBSAs with at least five school districts and find that segregation, by this measure, tends to be especially high among CBSAs with SEP districts. I also find that dissimilarity is particularly severe in these CBSAs when measured based on school district enrollments versus school-aged residents. This implies that the quality of public schools is particularly salient in residential decisions in these CBSAs. It also implies that, within these CBSAs, districts with small white residential populations have even smaller white public school populations, relatively speaking. Put differently, we would expect white students to tend to be under-represented in public schools relative to their share of the residential population in high-minority districts[36]. The previous chapter showed that SEP districts tend to be among these disproportionately minority-serving districts compared to the U.S. public school system as a whole. Thus, taken together, we would expect that white students in SEP districts are more likely to opt out of public schools altogether[37]. Importantly, this analysis does *not* imply that SEP schools *cause* intra-district segregation to be higher. Further research is needed to determine (1) whether this phenomenon predates SEP schools (perhaps even motivating their creation) or is a response to the (perceived) quality of education these schools help districts to achieve and (2) if SEP schools do in fact shape intra-district segregation, whether this is associated with any downstream effects, such as districts' property tax base or level of parental involvement.

---

[36] Or, vice versa, that districts with large white populations have even relatively larger white public school populations. However, this would imply that URM students are opting out of public schools at a higher rate than white students, which, given structural inequalities in financial capital for private schools or the capacity for transportation to charter or other non-neighborhood options, is less likely.

[37] Figure A2 in the appendix shows that this is indeed the case.

Next, I showed that intra-district segregation, in terms of both evenness and exposure, changes across the K-12 career, controlling for segregation in the prior grade-level. By and large, trends are relatively similarly in SEP and non-SEP districts, particularly in terms of evenness. In terms of exposure, structural school transitions in SEP districts bring relatively larger increases in white student exposure to URM peers and relatively smaller increases in URM exposure to white peers ($p<0.05$). Additionally, over the course of high school, segregation in SEP districts departs significantly from segregation in non-SEP districts, resulting in greater dissimilarity (by twelfth grade), greater white-to-URM exposure (in ninth grade only), and less URM-to-white exposure. Given that the divergence in exposure begins relatively early in high school, I argue that these trends are most consistent with a phenomenon whereby URM students who attend predominantly white (e.g., SEP) schools in SEP districts are disproportionately likely to transfer out of these schools compared to URM students in predominantly white high schools in non-SEP districts.

Overall, then, there does seem to be a relationship between SEP schools and segregation. The observation that structural school transitions bring significantly different changes in levels of exposure in SEP and non-SEP districts suggests that this is at least in part due to the way that school admissions procedures shape student sorting. Importantly, however, differences at the point of admission do not seem to explain all of the difference. Rather, fluctuations in the student population throughout high school are greater in SEP districts. Although it is unlikely that this can be (fully) attributed to differences in dropout, it does suggest that certain students—likely URM students in predominantly white schools—face additional challenges in SEP districts. One possibility is that SEP high schools attempt to push lower achieving students out early on in order to boost their "stats," such as four-year graduation rates, four-year college acceptance rates, or exam scores, in order to maintain their prestigious reputations. If URM students are disproportionately likely to be among lower-achieving admits, this might contribute to the eventual "whitening" of SEP schools during high school. A different but related possibility is that URM students simply receive less support at SEP schools than their white peers and are more likely to "opt" to transfer. These possibilities are largely in keeping with a (racialized) winner-take-all high school phenomenon (Attewell 2001) and suggest the need to evaluate whether these experiences are particularly damaging for underrepresented minority students.

V. DISTRICTWIDE MATH ACHIEVEMENT: SLOWING GROWTH AND GROWING GAPS[38]

Evidence from existing research fails to produce a consensus regarding the impacts of selective differentiation or "ability-grouping" on achievement. In general, proponents of school choice argue that competition or specialization between schools should increase system efficiency, thereby improving student achievement. Meanwhile, detractors voice concerns that sorting students might result in a drain on resources from traditional public schools or diminished potential for positive peer spillovers. Importantly, both sets of theories imply achievement effects beyond the walls of choice schools themselves. However, little research addresses districtwide effects of student sorting on achievement, and none has explored these questions as they pertain to SEP schools specifically.

Thus, this chapter asks: do SEP districts produce different (1) overall levels of achievement growth or (2) inequality of achievement than non-SEP districts? And, given the documented concerns around equity in access, (3) do these effects differ for students of different racial/ethnic backgrounds? By focusing on districtwide achievement, this chapter moves beyond a micro-level understanding of the effects of between-school student differentiation on the individual students and families who actively participate in school choice, in order to evaluate the meso-level impact of district organization on achievement outcomes.

To date, data limitations have prevented any such analysis. In addition to the original data collected for this study, nationally normed data on district-level student achievement were unavailable until recently. These data are now available from the Stanford Education Data Archive (SEDA, Reardon et al. 2018). From SEDA, I draw on measures of district-level mean math achievement from third to eighth grade to evaluate achievement growth, as well as standard deviation of math achievement and racial achievement gaps as measures of inequality. Together, these data provide the unique ability to assess the role of SEP schools in shaping districts' overall ecology of educational opportunity.

Analyses employ a difference-in-differences (DiD) design, using grade-level as the longitudinal dimension, in order to maximize the analytical leverage available from SEDA. Accordingly, the "treated" sample consists of 30 SEP districts that begin offering SEP schools between fourth and eighth grade, to allow for a third grade "pre-treatment" observation. To evaluate whether the "dosage" of treatment matters and leverage the variation in these districts, I test for differences not only between SEP and non-SEP districts, but between districts with different *prevalence* of SEP schools between fourth and eighth grade. Overall, I find that district-wide math achievement growth is slower in SEP districts than non-SEP districts. This result holds for white, Black, and Latinx students, but not for Asian students. Additionally, I test the effect of SEP schools on district-level inequality of achievement using measures of standard deviation of achievement, as an indicator of total variance, as well as racial achievement gaps. Findings for each of these outcomes is mixed: I find marginal evidence of an increase in standard deviation only when using the dose treatment specifications, and I find evidence of an increase racial/ethnic achievement gaps only between white and Latinx students, not white and Black students.

The remainder of the chapter is organized as follows. First, I introduce research on the effects of between-school student differentiation on system-wide achievement outcomes, in order

[38] A modified version of this chapter has been published in *Sociological Science* (Irwin 2020).

to generate hypotheses for the subsequent analyses. This is followed by a discussion of my data collection procedure, secondary data from SEDA, and my difference-in-differences modelling approach. I then present results, before turning to a discussion, which highlights the questions that remain about mechanisms involved in this process.

System-Wide Effects of High-Scope Differentiation

Chapter 2 describes recent research that examines the effect of SEP schools on the achievement of enrolled students. Findings from rigorous regression discontinuity designs suggest no effect of SEP schools on marginally admitted students (Abdulkadiroğlu et al. 2014; Allensworth et al. 2016; Dobbie and Fryer Jr. 2011). In this discussion, I also point out that these marginally admitted students may be the least likely to benefit from SEP schools, either because they struggle with the advanced curricula and/or insufficient remedial attention or because the marginally rejected students against whom they are measured benefit from their new top-of-the-class status in traditional public schools. Moreover, the null finding for these students does not necessarily imply null effects across the achievement spectrum or throughout the district. Given the potential consequences for SEP schools on peer effects, resource allocation and—as shown in the preceding chapter—segregation, it is important to consider districtwide effects of SEP schools on student achievement.

To the best of my knowledge, no existing research has examined district-level effects of high-scope *selective* student differentiation in the United States. However, the proliferation of charter schools has motivated research on the effect of "competition" on traditional public school students. By law, charter schools are generally forbidden from using academic criteria in their assignment procedures[39], but they produce between-school (high-scope) differentiation on the basis of organizational, instructional, or disciplinary preferences. Research has evaluated what effect charter "competition" has on achievement in traditional public districts, but evidence is mixed, with studies finding positive (Booker et al. 2008; Jinnai 2014), negative (Ni 2009), or null (Zimmer and Buddin 2009) effects. Although these studies take the important step of testing system-level achievement outcomes, they face geographic limitations because charter schools typically constitute their own "district" and do not have strict geographic attendance boundaries. This makes it difficult to define the differentiated *system* of which they are a part, and is likely a contributing factor in producing such equivocal findings.

In contrast to the charter literature, which is challenged by a lack of clear school system boundaries, stratification researchers evaluate the effect of selective school system differentiation at the national level, thereby precluding this issue. Scholars in this field find that differentiated school systems are more likely to (re)produce inequality (Hanushek and Wößmann 2006; Pfeffer 2012; van de Werfhorst and Mijs 2010), while also producing some evidence of lower average achievement (Hanushek and Wößmann 2006; Pfeffer 2012). Further, studies find that the more *rigidly* systems are differentiated, that is, the more difficult it is to switch between tracks once sorted (Pfeffer 2008), and the earlier differentiation happens (Horn 2009), the greater the inequality.

SEP districts can be considered relatively rigid in the American context, because earning access to the top "track" in these districts is not a matter of changing classes within a single comprehensive school, but rather of gaining admissions to a SEP school in a specific grade. While SEP districts are certainly not as rigidly differentiated as countries like Germany (e.g.

[39] However, academic *self*-selection into charter schools is possible.

West and Nikolai 2013), for instance, they nevertheless present a contrast to what is otherwise considered a relatively open system throughout the United States[40]. Based on this literature, then, we might expect SEP districts to demonstrate (1) higher inequality and (2) similar or perhaps somewhat lower levels of achievement than non-SEP districts.

Data and Analysis

Beyond the lack of data on SEP districts, the ability to answer the questions posed in this chapter has been limited to date by insufficient nationwide and nationally-normed district-level achievement data. This gap has recently been filled by the Stanford Education Data Archive (SEDA), which normalizes annual standardized test performance for students in grades three through eight across districts throughout the United States. Together with my original data, I use SEDA to estimate difference-in-differences (DiD) analyses of district-level average math achievement and inequality of achievement between third and eighth grade, in SEP versus non-SEP districts.

SEDA: Dependent Variable and Controls

The key dependent variables utilized here come from the district-level achievement measures in SEDA version 2.1 (Reardon et al. 2018). Every district in the United States is required to test students annually in grades three through eight in math and English Language Arts (ELA) and to report this achievement in the national *EdFacts* data system; however, states establish their own tests and proficiency levels. Reardon and colleagues compile district-level data from these tests (for 2009 to 2015) and benchmark these against the National Assessment of Educational Progress (NAEP) by state, in order to create measures that are comparable for districts across states. This requires transforming achievement on state tests, which are often reported as coarse proficiency categories, onto the continuous NAEP scale, to generate estimates of both mean and standard deviation of achievement. After benchmarking to NAEP, SEDA then scales these estimates for interpretability. The following analyses use the Cohort Standardized (CS) scale, which is calculated by mean-centering and dividing by the national grade-specific standard deviation from the original NAEP transformation[41] (Fahle et al. 2018; Reardon, Kalogrides, and Ho 2017). This scale is somewhat confusing to interpret for achievement *growth* because average achievement on this scale does not increase from one grade to the next. However, the CS scale is well-suited to evaluate changes in inequality across grade levels, because it is standardized by the amount of variation in each grade. Moreover, by representing achievement in standard deviation units, results using this scale can be interpreted as effect sizes. For consistency, therefore, all model results are presented using the CS scale. Where appropriate, I will reference parallel model results using grade-equivalent units to aid in interpretation of achievement growth[42].

---

[40] i.e., traditional public school students in SEP districts are not structurally precluded from selective four-year colleges, as is the case for vocational school students in Germany.

[41] At each grade level, every cohort is mean-centered and standardized relative to a single cohort (at the appropriate grade-level) to allow for analysis of absolute differences in achievement levels over time. The reference cohort is the cohort that participated in NAEP in fourth grade in 2009.

[42] SEDA provides an additional scale, called the Grade Cohort Scale (GCS), which is standardized relative to the change in NAEP scores from one grade to the next, each within a

Table 5.1. Pre-Treatment (3rd Grade) Dependent Variables, Averaged 2009-2015

Variable	Data Source	Range		Non-SEP [a]	SEP	Total
N				755	30	785
Dependent Variables, Pre-Treatment (3rd Grade)						
Mean Achievement						
All		-1.310	1.127	-0.030	-0.209	-0.037
White (N=744)		-0.801	1.259	0.247	0.195	0.245
Black (N=626)		-1.189	0.286	-0.424	-0.502	-0.428
Latinx (N=739)		-1.107	0.564	-0.260	-0.293	-0.261
Asian (N=539)		-0.947	1.776	0.479	0.385	0.474
Inequality of Achievement						
Standard Deviation		0.777	1.324	0.967	0.979	0.967
White-Black Gap (N=612)		0.097	1.680	0.693	0.698	0.693
White-Latinx Gap (N=708)		-0.108	1.434	0.506	0.488	0.506

Notes:
[a] Non-SEP districts signifies all comparison districts, including future SEP districts

In addition to overall district-wide achievement, SEDA provides district achievement estimates by race. Because of the racially unequal access to SEP schools described in Chapter 2, it is important to test for potential differences in outcomes associated with SEP districts for students of different racial backgrounds. Accordingly, in addition to overall district mean achievement, I assess district mean math achievement for each race available from SEDA—white, Asian, Black, and Latinx—separately. I also examine inequality using SEDA's measures of white-Black and white-Latinx math achievement gaps, which are constructed by differencing race-specific mean achievement. Descriptive statistics for these variables (pre-treatment) are presented in Table 5.1.

Analytic Approach: Difference-in-Differences for Math Achievement

This chapter evaluates the effect of SEP schools on district-level mean and standard deviation of math achievement. I focus on math achievement for three reasons. First, math is a better indicator of the effect of school interventions because math is learned primarily in school, while language acquisition and reading skills are more heavily reinforced in the home (Bryk and

given cohort. This scale produces grade-equivalent units, which increase from an average of 3 for students achieving at grade-level in third grade to 8 in eighth grade. On this scale, the average standard deviation of math achievement increases with each grade-level, making it somewhat difficult to compare inequality between pre- and post-treatment. For questions of achievement growth, however, the GCS is quite useful because scores increase by approximately one unit with each additional grade, providing an intuitive measurement for achievement *growth*. Models for achievement growth are replicated in Appendix C using the GCS scale.

Raudenbush 1988; Burkam et al. 2004; Murnane 1975; Parcel and Dufur 2001). This is reflected in the fact that summer losses tend to be greater in math than in reading, and this is increasingly true as students progress through school (see Cooper et al. 1996 for review). Given the potentially diffuse effects of SEP schools on achievement throughout a district, it is prudent to focus on an outcome that is more tightly coupled with schooling. Second, over the period from third to eighth grade, math instruction itself also becomes more differentiated, with students progressing at different speeds—and in different learning groups—through subject areas like General Math, Pre-Algebra, Algebra, or even Geometry by the end of eighth grade. Although differentiation in English/Language Arts certainly occurs, this tracking primarily distinguishes the *pacing* of classes—between "regular," "honors," or Gifted, for example—as opposed to distinct curricular content. Given the structure of these subjects during the focal grade-levels for the present analyses, then, we might expect SEP schools to exert a more direct influence on the organization of math instruction[43].

To evaluate the relationship between district-level selective differentiation and student achievement outcomes in math, I deploy a difference-in-differences (DiD) design using grade level as the longitudinal dimension. DiD analyses compare the difference between treated and control units before treatment (the first difference) to their difference after treatment (the second difference) to estimate whether the change (if any) in the treated units is distinguishable from that of the control units and therefore attributable to treatment. Typically, treatments pertaining to policy changes or medical trials, for example, begin in a particular month or year, such that the periods "pre" and "post" treatment are defined by standard time. This is not possible for the present analyses, because the SEDA data are too new to measure achievement before SEP schools were founded. Instead, I use grade as the longitudinal dimension. For the case of SEP schools, I argue that this approach actually offers some advantages over the traditional design. First, SEP schools are not the product of a targeted policy intervention that took effect in a single year, meaning that the particular temporal contexts of their founding could have different implications for their immediate impacts across districts. Moreover, whereas traditional longitudinal designs offer a single observation of the point of treatment, using grade level as the longitudinal dimension means that I observe the transition into treatment up to 7 times for each SEP district in SEDA, which provides achievement data from 2009 to 2015. By averaging grade-level achievement across years, I am able to greatly reduce the risk that results are influenced by idiosyncratic shocks. Models, described below, therefore include one pre-treatment (third grade) and one post-treatment (eighth grade) observation per district[44].

Drawing on the grade-level variation in the provision of SEP schools, I define three treatment measures. First, I model treatment as a binary indicator of SEP district status, where 1

[43] Due to this tracking structure, it is also easier to identify advanced coursework in math than in reading using available national data. In fact, while the Civil Rights Data Collection (CRDC) provides relatively detailed information on mathematics courses, beginning with Algebra in seventh grade, it contains no information about English courses. This information on student participation in advanced coursework is tested as a mediator in Appendix B.

[44] Averaging across years also avoids the issue of states with missing data in eighth grade for multiple years because of the use of end-of-course (as opposed to end-of-year) tests, which could therefore not be benchmarked to state-level NAEP scores. For instance, California has only one year of complete eighth grade data and Texas only 2 (out of 7), so running regressions with one observation for each year could importantly under-weight these large states.

indicates at least one SEP school present in the district during the observation period, and 0 indicates none. Districts are only considered treated if they offer their first SEP schools between fourth and eighth grade (N=30), so that third grade functions as a true pre-treatment observation for all districts in the analytic sample. Then, to test whether the "dosage" of treatment is important, I leverage the fact that different districts begin to offer SEP schools at different grade levels, and with greater or lesser prevalence. I operationalize the dosage of treatment in two ways: (1) the number of grades between fourth and eighth with at least one SEP school, ranging from 0-5, and (2) the percentage of fourth to eighth grade students throughout the district enrolled in a SEP school (with an interquartile range of 2-7%), based on enrollment data from the Civil Rights Data Collection (CRDC). In an ordinary least squares (OLS) framework, the effect of interest is operationalized as the interaction between the relevant treatment measure and grade-level, as represented by β_3 in the following equation.

$$Y_{ig} = \alpha + \boldsymbol{\beta X_{ig}} + \beta_1 post_g + \beta_2 SEP\_Dist_i + \beta_3 DiD_{ig} + \epsilon_{ig} \qquad (5.1)$$

where,

Y_{ig} is the achievement outcome for district i in grade g.

α is the intercept, which can be interpreted as the expected achievement outcome in 3rd grade in comparison (i.e., non-SEP) districts.

$\boldsymbol{X_{ig}}$ is a set of mean-centered district-(grade-)level covariates.

$post_g$ is an indicator for whether the observation occurs in 8th grade (equal to 1 in 8th grade, 0 in 3rd).

SEP_Dist_i is one of three measures of the existence/prevalence of SEP schools in the given district, described above, and takes on the same value in 3rd and 8th grade for each district.

DiD_{ig} is the product of SEP district status and grade level ($SEP\_Dist_i * post_g$)

ϵ_{ig} is the Huber-White heteroscedasticity-robust error term, clustered at the district level to account for non-independence of errors at this level[45].

Identification of a treatment effect in these models rests on the assumption that, in the absence of treatment, treated and control units would follow parallel trends. I address this assumption by carefully considering the definition of my treatment and comparison groups. For the comparison group, I restrict to 'peer' districts by including only districts with some form of (non-selective) high-scope differentiation, which I define as districts with at least 5 schools serving students in each grade between grades 4 and 8, at least one of which is a traditional magnet *or* charter. There are 772 such districts with math achievement data from SEDA. These districts present an analytically useful comparison group because they are most similar in size and structure to SEP districts. In this way, I am comparing SEP districts to the districts they would be most likely to resemble in the absence of SEP schools. While this necessarily limits the purview of the analysis and the conclusions that can be drawn from it, I argue that these bounds strategically sharpen the test of the mechanisms of interest—the scope and selectivity of student

[45] Additionally, to account for potential state-level differences in proficiency standards or other policies that may influence student test outcomes, models are estimated with state fixed effects by including a state indicator as a factor variable in the models, representing 42 states and the District of Columbia; I set the largest state—California—to be the base category.

differentiation—in two important ways. First, it avoids comparing SEP districts to districts that logically cannot support high scope (between-school) differentiation by virtue of their small size[46]. And second, by restricting the comparison sample to districts with intra-district school choice, I focus my analysis specifically on the *selectivity* of high-scope differentiation[47].

Still, selecting a strong comparison group does not fully address the assumption of parallel trends. To this end, as mentioned above, I omit districts that begin offering SEP schools before or during third grade. This reduces the likelihood of divergent trends emerging prior to the observation period—at least as a function of treatment—but does not permit an empirical test of this assumption, due to the lack of achievement data prior to third grade. Instead, I draw on the variation in the grade-level of treatment to generate a falsification test. Several SEP districts in my data offer SEP schools for the first time during or after 9th grade. This means that, during the fourth to eighth grade observation window for the current analysis, these districts are operating under a pre-treatment district structure. Because these districts are likely to be the most similar to the analytic treatment sample, they offer useful leverage to evaluate the assumption that it is treatment itself, not the predisposition or preparation for treatment, that produces divergent trends between SEP districts and comparison districts. Thus, in addition to the 30 *treated* SEP districts, analyses below include 22 *future* SEP districts. These future SEP districts enter the model as an additional set of DiD terms, as expressed in equation 5.2, below. In this setup, the falsification test asks whether future SEP districts differ from non-SEP comparison districts over the treatment period. In other words, confidence in the effect measured by β_3 is enhanced if β_5 is *not* statistically significant.

$$Y_{ig} = \alpha + \boldsymbol{\beta X_{ig}} + \beta_1 post_g + \beta_2 SEP\_Dist_i + \beta_3 DiD_{ig} + \beta_4 FutureSEP\_Dist_i + \beta_5 (FutureDiD_{ig}) + \epsilon_{ig} \quad (5.2)$$

Finally, each model includes a set of control variables ($\boldsymbol{X_{ig}}$), shown in Table 5.2. Control variables are the same in all models, with the exception of measures of poverty and socioeconomic inequality, for which race-specific measures are used in the models for race-specific district mean achievement. I include both percent Black and percent Latinx in all models to account for the percent of students from racial/ethnic backgrounds that tend to be underserved by SEP schools (r = -0.347 in analytic sample). Control variables from Census Education Demographic and Geographic Estimates (EDGE) are estimates for all school-aged children, regardless of public enrollment status. This is useful because families' choice to opt in or out of public schools may be influenced by whether the district offers SEP schools. All control variables are mean-centered in the models, besides urbanicity, which is comprised of a set of binary variables (the reference category is suburban). Accordingly, model intercepts (shown only in Table 5.3) can be interpreted as the third grade math achievement in a (suburban) district with average characteristics[48].

[46] In addition to being substantively meaningless to attribute achievement differences in a district with one school to the lack of SEP schools, these small districts, if included, would also receive disproportionate weight by virtue of the district-level structure of the analysis.
[47] Additionally analyses, not shown, define the comparison group solely by district size and produce substantively similar results.
[48] Such an "average" district may be more or less typical for students of different racial backgrounds.

Table 5.2. Descriptive Statistics for Control Variables, Pre-Treatment (3rd Grade)

Variable	Range		Non-SEP[a]	SEP	Total
N			755	30	785
Total Expenditures (log)	8.74	10.49	9.31	9.39	9.31
Total Teachers (log)	4.14	10.30	6.75	7.95	6.80
% Free or Reduced Lunch	0.04	0.96	0.49	0.66	0.50
% Black students	0.00	0.98	0.16	0.37	0.17
% Latinx students	0.00	1.00	0.32	0.30	0.32
% of students in charters	0.00	0.50	0.18	0.25	0.18
50/10 income ratio	1.73	9.93	3.82	4.17	3.83
Urban (0/1)	0	1	0.41	0.80	0.42
Rural (0/1)	0	1	0.11	0.00	0.11
Number of Charters in Geographic District	0.00	223.29	3.95	20.17	4.57
% school-age children who are foreign born	0.00	0.20	0.05	0.06	0.05
% school-age children in female-headed households	0.01	0.65	0.28	0.37	0.29
% of school-age children in owner-occupied homes	0.13	0.91	0.60	0.50	0.59
Median owner-occupied home value (log)	10.47	13.82	12.15	11.95	12.14
Total school-age population (log)	6.91	13.56	9.87	11.09	9.92
FRL-eligible student exposure to FL-eligible peers [b]	0.04	0.97	0.55	0.72	0.56
Race-Specific SES					
% of white 5-17 year olds in poverty	0.00	0.58	0.12	0.16	0.13
White 50/10 income ratio	1.10	17.68	3.65	4.04	3.66
White exposure to FL-eligible peers	0.04	0.96	0.43	0.53	0.43
% of Black 5-17 year olds in poverty	0.00	1.00	0.30	0.36	0.30
Black 50/10 income ratio	1.04	35.04	4.54	4.70	4.54
Black exposure to FL-eligible peers	0.03	0.95	0.52	0.70	0.53
% of Latinx 5-17 year olds in poverty	0.00	1.00	0.28	0.36	0.29
Latinx 50/10 income ratio	1.05	21.00	3.65	3.34	3.64
Black exposure to FL-eligible peers	0.04	0.97	0.54	0.70	0.55

Notes:

[a] Non-SEP districts signifies all comparison districts, including future SEP districts

[b] FRL stands for Free or Reduced Lunch; FL stands for Free lunch. In districts where schools offer ‘community coverage’ for school lunches, meaning that 100% of students are provided free lunches, SEDA imputes the percentage that would be eligible based on traditional eligibility requirements.

Where there are an insufficient number of students (fewer than 20) with reported test data, SEDA does not provide race-specific achievement data. This results in a different number of observations in models for different racial groups. Further, SEDA does not produce the race-specific poverty and socioeconomic segregation control measures for Asian students, so I rely on the full district measures for this group. Finally, of the 772 comparison districts with math achievement data, 17 are missing covariate data and are dropped from the analytic sample. No SEP districts are missing covariate data, resulting in a final analytic sample of 785 districts.

Results

Before turning to model results, I present a graphical depiction of grade-level math achievement from third to eighth grade to establish the empirical phenomenon to be explained. Figure 5.3 shows the raw (i.e., uncontrolled) trend in average math achievement in SEP districts that begin offering SEP schools in fifth, sixth, or seventh grade[49], as well as comparison districts. Overall, achievement in both SEP and comparison districts appears relatively low (below 0 on a mean-centered scale) and appears to be trending downward (in relative terms). Note that this trend does not imply negative *absolute* growth from one grade-level to the next[50]. An additional trend line is provided for smaller and/or undifferentiated districts (i.e., "other" districts that are omitted from the analytic sample) in order to demonstrate that, although comparison districts in the analytic sample are relatively low-achieving, their math achievement trajectory parallels that of districts throughout the United States. In other words, the apparent downward achievement trend is an artifact of the SEDA scaling[51], rather than a strange feature the analytic sample. The importance of Figure 5.3 lies in the relationship between these trend lines, not the specific slope of any individual line.

With this in mind, Figure 5.3 depicts relatively large drops in mean achievement at the grade level when SEP schools are first offered, with generally (but not always) flatter slopes at other grade-levels. This could in part reflect an effect of school transitions, as research has shown that achievement tends suffer initially when students enter a new school, before bouncing back (Grigg 2012). However, although nearly all students in the U.S. experience a school transition between third and eighth grade, the overall math achievement trend is steeper and more negative in SEP districts than in comparison districts on average across grade levels. Further, the line of best fit through the pre-treatment trend in SEP districts (i.e. through only

---

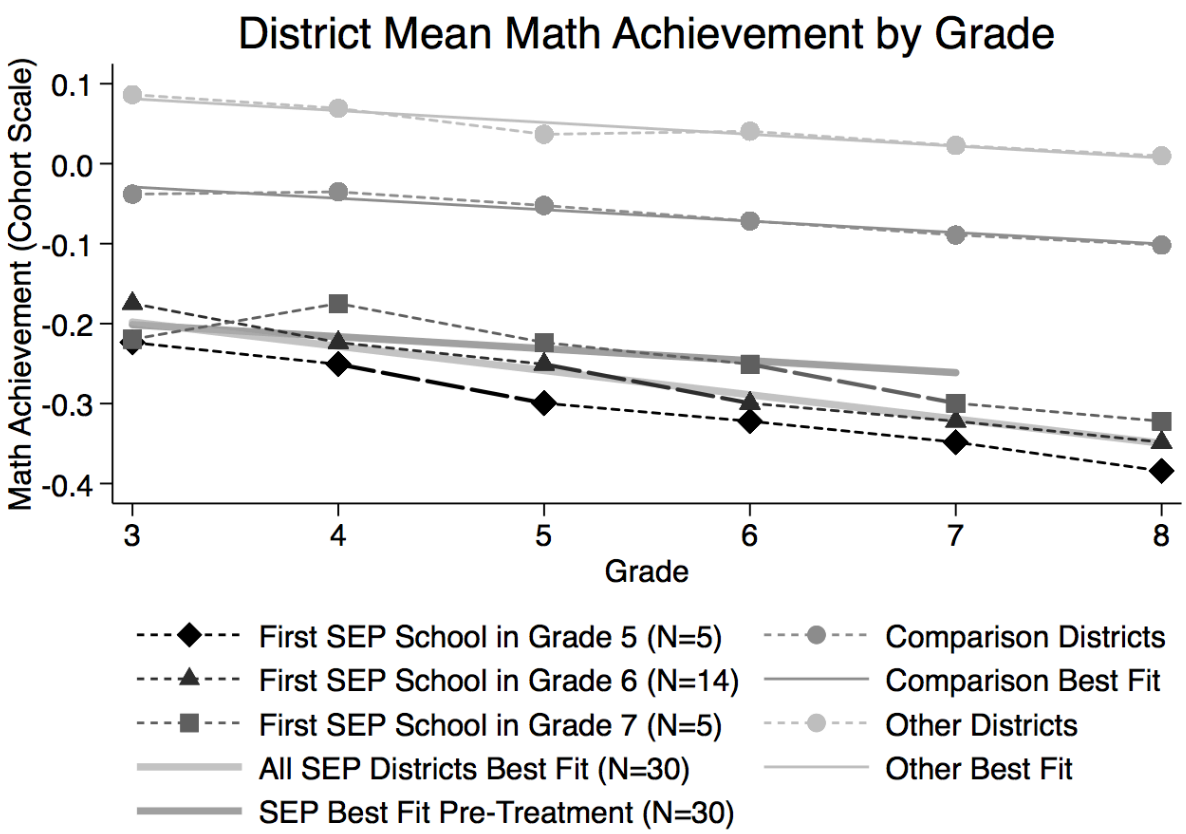
[49] Three districts each begin offering SEP schools in fourth and eighth grade. Individual trend lines for these sets of districts are omitted to avoid over-crowding the graph, but are included in the average lines of best fit.

[50] Rather, it may imply relatively *faster* growth in the NAEP-participating cohort, against which other cohorts in the data were standardized, producing an artifact of relatively slower growth overall.

[51] This downward trend from one grade to the next persist for all cohorts *except* the reference cohort of 2009 fourth grade NAEP participants. This suggests there is some artifact of benchmarking other cohorts against this cohort that produces this trend. This artifact is unique to this cohort-mean-standardized scale and does not affect model results. Results for models using the grade equivalent scale (which increases 1 unit per grade-level on average) are presented in Appendix B.

those district-grades in which SEP schools are *not yet* offered[52]) shows that before SEP districts begin to offer SEP schools, their achievement follows a very similar trend to comparison districts on average. This suggests that SEP schools may indeed be implicated in the lower average math achievement growth in these districts. By controlling for district characteristics, and considering the prevalence of SEP schools, the DiD regression models, below, further test whether this divergence in district math achievement is associated with the provision of SEP schools.

Figure 5.3. Grade-Level Achievement in SEP Districts vs. Other Districts



Note: for Pre-Treatment SEP district sample, the sample size at each grade-level is as follows: (3^{rd} = 30, 4^{th} = 27, 5^{th} = 22, 6^{th} = 8, 7^{th} = 3)

In each table below, the coefficient of interest is called “SEP District DiD”. This is the interaction term between a district’s SEP treatment status (or dosage), and whether the observed achievement level is measured ‘pre’ (third grade = 0) or ‘post’ (eighth grade = 1) ‘treatment’. This is the additional difference between SEP and non-SEP districts that arises between third and eighth grade, beyond any difference that might have existed in third grade. Table 5.3 also includes coefficients for (1) the main effect of SEP district treatment status/dosage, which can be interpreted as the difference between treated and comparison districts in third grade, and (2) an

[52] i.e. there is a diminishing sample from one grade to the next, with 30 in 3^{rd} grade and only three districts that are still “pre-treatment” by 7^{th} grade.

indicator for eighth grade (third grade = 0, eighth grade = 1), which can be interpreted as the expected change in non-SEP districts between third and eighth grade. These coefficients are omitted from subsequent tables for brevity. Results for the falsification test are represented as the coefficient for *future* SEP DiD.

Slowing Growth in SEP Districts

First, I examine the association between high scope selective differentiation and achievement growth at the district level. Model 1 shows the uncontrolled estimate where SEP district treatment is operationalized as a binary status. In this naïve model, the DiD coefficient is -0.084. Recall that, because achievement is measured in standard deviation units, this represents a negative 8.4% effect size. This translates to the equivalent of nearly half a grade-level less growth over the third to eighth grade period in SEP districts relative to non-SEP districts (see Appendix Table B1 for results using SEDA's grade-equivalent scale). There is evidence that SEP districts tend to be different from non-SEP districts prior to treatment (the coefficient for the third grade difference is significant in Model 1), but this difference is fully accounted for by the district-level control variables introduced in Model 2[53]. However, adding these controls has no impact on the estimated effect of district SEP status on achievement growth between third and eighth grade.

To further test that there are no unmeasured characteristics that increase *both* the propensity of districts to offer SEP schools *and* math achievement growth, Model 3 includes the additional difference-in-difference test for districts that will begin selectively differentiating students into SEP schools at the high school level. The DiD for these *future* SEP districts in Model 3 is not statistically significant, meaning that there is no expected difference in math achievement trends between these districts and other comparison districts. Moreover, there *is* a statistically significant difference between the achievement trends in future SEP districts and current treated SEP districts (p=0.044, Wald test comparing DiD coefficients not shown). Together these findings lend confidence to the assumption of parallel trends prior to treatment and imply that the divergence in trends between SEP districts and comparison districts between third and eighth grade is not being spuriously driven by the *propensity* for districts to offer SEP schools, but rather by the logistics and practice of providing SEP schools[54].

Models 4 and 5 further assess the relationship between SEP offerings and district-wide achievement by operationalizing treatment as the district-level "dosage", or prevalence of SEP schools over the fourth to eighth grade treatment period. In Model 4, SEP districts' treatment dosage is defined as the number of grades in which at least one SEP school is offered, ranging from 0 to 5. In Model 5, SEP school dosage refers to the percentage of fourth to eighth grade students enrolled in SEP schools in the district. In Model 4, each additional grade for which

---

[53] Note that in Appendix Table B1, which presents model results using the GCS (grade-equivalent) scale, the difference between SEP and comparison districts in third grade remains significant in Models 2-5. It is possible that the collinearity between grade level itself and the GCS scale may artificially produce this relationship. However, this does not affect the DiD coefficients, which are consistent both in significance and relative effect size in models using the GCS and CS scales.

[54] By examining districts that are likely to be most similar to the treated districts in this analysis, this test also provides reassurance that results are not likely to be significantly affected by the omission of any districts that the scrape may have failed to identify.

districts offer at least one SEP option is associated with a 0.022 decrement in math achievement, or a -2.2% effect size. Among SEP districts, the average number of treated grades is 3 (i.e., SEP schools begin in sixth grade), meaning that this translates to an average expected decrement of 0.066, or the equivalent of approximately one third of a grade-level of learning (see Model B3 in the appendix). Similarly, Model 5 estimates a growth penalty of 0.014 standardized achievement units for each additional percentage of fourth to eighth grade students enrolled in a SEP school. With an average SEP school enrollment of 4.77% in SEP districts, this again equates to an average effect size of negative 6.6%.

Table 5.3. OLS DiD Regression of SEP District Status on Mean Math Achievement

	(1)	(2)	(3)	(4)	(5)
District controls		x	x	x	x
Future-SEP			x	x	x
SEP Scope Definition	binary	binary	binary	grade count	% enrolled
SEP District DiD	-0.084**	-0.085**	-0.085**	-0.022*	-0.014**
	(0.027)	(0.028)	(0.028)	(0.009)	(0.005)
SEP District	-0.171**	0.027	0.025	0.008	0.006
(3rd grade difference)	(0.057)	(0.023)	(0.024)	(0.008)	(0.005)
Post/8th Grade	-0.064***	-0.089***	-0.089***	-0.090***	-0.090***
(ref, 3rd Grade)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
Future-SEP District DiD			-0.014	-0.013	-0.013
			(0.024)	(0.024)	(0.024)
Constant	-0.039**	-0.284***	-0.283***	-0.283***	-0.285***
	(0.012)	(0.029)	(0.029)	(0.028)	(0.029)
Observations	1604	1570	1570	1570	1570
R-squared	0.021	0.788	0.788	0.788	0.788

Huber-White heteroscedasticity-robust clustered standard errors in parentheses
+ p<.1, * p<.05, ** p<.01, *** p<.001

Overall, the estimated the DiD coefficient for SEP districts is quite consistent across models in Table 5.3. However, note that while this increases confidence in the robustness of the result, it does not necessarily imply that SEP school dosage exerts a truly linear effect. Due to the relatively small number of treated districts (N=30), it is difficult to interrogate these relationship in detail. One possibility is that these patterns reflect a “tipping point,” rather than a linear effect. For instance, estimated effects may be driven by districts with above-average prevalence of SEP schools (i.e., 4^{th}-8^{th} grade district enrollment of 5% or higher). Additional regressions, not shown, suggest that this may be the case. Such an effect is more difficult to evaluate for the grade-level of differentiation, because of clustering of treatment onset at sixth grade.

Nevertheless, these findings present a high degree of confidence that a relationship exists between district provision of SEP schools and overall math achievement.

Race-Specific Achievement

Next, I test whether this effect differs by race. Each panel in Table 5.4 parallels Models 2-5 from Table 5.3 for white (W), Asian (A), Black (B), and Latinx (L) students, separately. Where relevant, I will refer to these models collectively as race-specific (RS) models.

Table 5.4. DiD of SEP Status on Math Achievement, by Race

	binary	binary	grade count	% enrolled
District controls	x	x	x	x
Future SEP		x	x	x
Panel 1: White (N=1488)	(2-W)	(3-W)	(4-W)	(5-W)
SEP District DiD	-0.091**	-0.091**	-0.029*	-0.016*
	(0.034)	(0.034)	(0.013)	(0.007)
Future-SEP District DiD		-0.005	-0.005	-0.004
		(0.030)	(0.030)	(0.030)
Panel 2: Asian (N=1078)	(2-A)	(3-A)	(4-A)	(5-A)
SEP District DiD	-0.033	-0.033	-0.004	0.000
	(0.041)	(0.042)	(0.013)	(0.008)
Future-SEP District DiD		0.012	0.013	0.014
		(0.064)	(0.064)	(0.064)
Panel 3: Black (N=1250)	(2-B)	(3-B)	(4-B)	(5-B)
SEP District DiD	-0.087*	-0.088*	-0.020+	-0.010+
	(0.036)	(0.036)	(0.011)	(0.005)
Future-SEP District DiD		-0.040	-0.039	-0.038
		(0.026)	(0.026)	(0.026)
Panel 4: Latinx (N=1478)	(2-L)	(3-L)	(4-L)	(5-L)
SEP District DiD	-0.110***	-0.110***	-0.030***	-0.017***
	(0.026)	(0.026)	(0.009)	(0.005)
Future-SEP District DiD		-0.003	-0.002	-0.002
		(0.026)	(0.026)	(0.026)

Huber-White heteroscedasticity-robust clustered standard errors in parentheses
+ p<.1, * p<.05, ** p<.01, *** p<.001

Models 2-W, 2-B and 2-L show that the provision of SEP schools is associated with lower levels of math achievement growth for white, Black, and Latinx students. The substantive story is the same for each of these three racial/ethnic groups, although the relative effect size appears somewhat larger for Latinx students (11%, compared to about 9% for white and Black students). Model 2-A shows no observed effect of SEP district status on the math achievement of Asian students.

Like the pooled race models, Models 3-RS include a second DiD test for districts that offer SEP school beginning in high school. Again, the coefficients for the future SEP DiD are not significant. This suggests that it is unlikely that some predisposition towards selective differentiation is spuriously driving the findings in this analysis. One important caveat is that, although the DiD for future SEP districts is not itself significant for Black students, it is also not significantly distinguishable from the treated SEP DiD. This may suggest that some part of the achievement penalty for Black students in SEP districts begins to arise prior to treatment. I will return to this possibility in the discussion.

Models 4-RS and 5-RS test the operationalization of treatment as continuous indicators of the prevalence of SEP schools throughout the district. For white and Latinx students, the average implied treatment effect over the third to eighth grade period remains consistent and statistically significant. For Black students, the estimated effect falls just below traditional levels of significance ($p=0.057$ and $p=0.064$ for Models 4-B and 5-B, respectively). Overall, however, both the magnitude and significance of these effects is very similar across model specifications for all racial/ethnic groups. In fact, in supplemental models (see Appendix B for replication of key coefficients from Tables 5.3 and 5.4), the effect for Black students remains statistically significant ($p<0.01$) for both dose treatment specifications. It is especially striking that these effects are not only consistent across model specifications, but are also quite consistent for students of different racial/ethnic backgrounds.

Importantly, just because effects are similar between students of different racial/ethnic backgrounds on average *across* districts does not mean they are necessarily the same *within* districts. This is a separate question, which requires evaluation of district-level racial achievement gaps. I turn to this next, along with the question of overall inequality of achievement.

Mixed Evidence of Increasing Inequality

I evaluate the association between SEP district status and inequality of achievement at the district level, using standard deviation and racial achievement gaps as measures of inequality, for a total of three inequality outcomes: white-Black achievement gaps (WBG), white-Latinx achievement gaps (WLG), and district-wide standard deviation of achievement (SD). As a reminder, the scale used for these analyses is standardized by the national grade-level standard deviation, and therefore accounts for the total amount of variation at each grade. Again, Table 5.5 includes models using both the binary status and dose treatment measures of district provision of SEP schools.

Overall, findings regarding the association between SEP schools and district-level inequality of achievement are quite mixed. There is no evidence of an effect of SEP schools on standard deviation of achievement using the binary treatment specification in Model 6-SD. However, there is marginal evidence of an increase in inequality when using the dose specifications. Model 7-SD predicts an average effect size of 2.4% (0.008 x 3), while Model 8-SD predicts an average effect size of 1.7% (0.003 x 4.77). These effects achieve only marginal significance (p=0.068 and p=0.072, respectively), but it is interesting to note the difference from Table 5.3, where greater significance was achieved using the binary than the dose treatment specifications. This could suggest that the appropriate functional form (and perhaps the relevant mechanisms) are different for standard deviation of achievement than for average achievement.

Table 5.5. DiD of District-Level Achievement Inequality on District SEP Status

	binary	grade count	% enrolled
District controls	x	x	x
Future SEP	x	x	x
Panel 1: District-Wide Standard Deviation (N=1570)	(6-SD)	(7-SD)	(8-SD)
SEP Status/Scope DiD	0.022	0.008+	0.004+
	(0.016)	(0.004)	(0.002)
Future SEP DiD	0.026	0.026	0.026
	(0.020)	(0.020)	(0.020)
Panel 2: White-Black Achievement Gap (N=1222)	(6-WBG)	(7-WBG)	(8-WBG)
SEP Status/Scope DiD	0.024	-0.001	-0.002
	(0.034)	(0.013)	(0.006)
Future SEP DiD	0.058+	0.057+	0.057+
	(0.033)	(0.033)	(0.033)
Panel 3: White-Latinx Achievement Gap (N=1416)	(6-WLG)	(7-WLG)	(8-WLG)
SEP Status/Scope DiD	0.041**	0.013**	0.006*
	(0.014)	(0.005)	(0.003)
Future SEP DiD	0.003	0.003	0.002
	(0.030)	(0.030)	(0.030)

Huber-White heteroscedasticity-robust clustered standard errors in parentheses
+ p<.1, * p<.05, ** p<.01, *** p<.001

Additionally, there is mixed evidence of worsening racial/ethnic gaps. There is evidence that the achievement gap between white and Latinx students increases between third and eighth grade in SEP districts relative to comparison districts. This effect is moderate in magnitude, on the order of a 4% effect size on average based on results from both Model 6-WLG and 7-WLG[55].

[55] In grade-equivalent terms, this reflects an increase in the white-Latinx gap of approximately one eighth (~.124) of a school year, relative to an average gap of 1.30 grade-equivalent units in third grade and 1.65 in eighth grade, holding all other control variables at their means.

The average effect appears somewhat smaller in Model 8-WLG, at about 2.9% (0.006 x 4.77). For white and Black students, by contrast, there is no evidence of a change in the achievement gap in SEP districts relative to comparison districts between third and eighth grade. Not only is the DiD non-significant, but the point estimate is actually negative (but essentially zero) in the dose treatment models. However, the *future* SEP DiD for Black students is marginally significant (p=0.079-0.087 across WBG models) and positive. While not too much should be made of these coefficients themselves, this does suggest the need for caution in interpreting results causally for Black students, and to consider the potential for effects of "preparing" for differentiation, particularly for these students.

Discussion

Existing research on the impact of SEP schools generally suggests neutral or slightly negative effects on achievement levels, whether that research focuses on enrolled students (Abdulkadiroğlu et al. 2014; Dobbie and Fryer Jr. 2011) or the entire differentiated school system (Hanushek and Wößmann 2006; Horn 2009). The present study estimates a slowing of math achievement growth by roughly 6.6-8.5% of a standard deviation in SEP districts relative to comparison districts, overall. This is a substantively large effect, equivalent to nearly half a grade-level's worth of learning over five years. This effect is relatively consistent for white, Black, and Latinx students, on the order of a negative 9-11% effect size using the binary treatment specification. The exception to this rule is for Asian students, for whom there is no discernable relationship between SEP schools and district-level achievement growth. Although these results are somewhat in keeping with existing research, the consistency and magnitude of the finding that SEP districts produce significantly *less* achievement growth between third and eighth grade than non-SEP districts is nevertheless striking.

Next, the relatively weak evidence of a relationship between high-scope selective differentiation and inequality of achievement at the district-level in the United States presents a contrast to the dominant findings from the nation-level social stratification literature, discussed above. While there is some suggestion of a marginal increase in the district-wide standard deviation of achievement as the prevalence of SEP schools increases throughout the district, the only consistent evidence of an effect on inequality of achievement is for the achievement gap between white and Latinx students. In interpreting this finding, it is worth noting both the advantages and disadvantages of the present analyses relative to existing work. On the one hand, unlike the stratification literature, whose units of analysis are whole countries with distinct educational regimes, the present analysis has the advantage of observing SEP and non-SEP districts within same states. This may allow the current study to better account for other policies considered important to student achievement, such as testing standards and the inclusivity of (public) postsecondary education (Ayalon and Gamoran 2000; Kerckhoff 2001). On the other hand, the present study may be limited by the fact that SEDA's estimates of standard deviation rely heavily on coarse proficiency categories, which might obscure achievement differences at the extreme. Estimates of standard deviation may therefore be conservative. In other words, while these estimates present important improvements over existing work, they also face important limitations that warrant investigation in future research.

Another point of consideration around the analysis of inequality is that the measures used here, particularly standard deviation of achievement, may obscure differences in the meaning of inequality in districts that have different average levels of achievement. For instance, if mean achievement is high, then a relatively high standard deviation may not in itself be 'bad', as it may simply imply exceptionally high achievement among students at the top of the distribution. Similarly, the same white-URM achievement gap would signify different educational experiences in a district where the average white student achieves at levels equivalent to the national average compared to one where they achieve a full grade-level above the mean. Using SEDA's standardized scaling helps to address issues of comparability by adjusting each dependent variable using national grade-level standard deviations, but does not address differences in districts' average levels of achievement. Yet, even with these differences and limitations, there is strong evidence of growing inequality between white and Latinx students in SEP districts relative to non-SEP districts. Given that mean achievement is simultaneously *decreasing* in SEP districts, it is unlikely that this inequality could be characterized as a "good" thing.

These findings were produced with a unique DiD design that uses grade-level as the longitudinal dimension. This was a useful approach to extract analytic leverage from the data, but may also raise concerns based on the inability to evaluate pre-treatment trends. This is important because a divergence between SEP and non-SEP districts prior to treatment might suggest unobserved differences in districts that have the *propensity* to offer SEP schools, or perhaps that there is some effect of "preparing" for selective differentiation. If either were the case, we would reasonably expect similar processes to characterize the trend in *future* SEP districts—those that have not begun offering SEP schools in the grade range of the current analysis, but which offer SEP schools in high school. Analyses test this possibility, and the estimated DiD coefficients for future SEP districts do *not* reach conventional levels of significance ($p<0.05$) in any of the models presented above. Overall, then, analyses in this paper present strong evidence of slowing of math achievement growth overall and a widening white-Latinx achievement gap in SEP districts relative to similar non-SEP districts.

Importantly, however, this DiD test for future SEP districts also raises some questions about the effect of SEP districts on Black students. Although future SEP status itself is not significantly predictive of math achievement growth relative to comparison districts, there is overlap in the estimated growth trend for Black students in treated SEP and future SEP districts. Moreover, for the white-Black achievement gap, future treatment is marginally predictive ($p<0.1$) of an increase in the achievement gap, while there is no evidence of a treatment effect of SEP districts from third to eighth grade. Rather than indicating something about the *type* of district that offers SEP schools (i.e., a spurious effect, which would imply an effect in both current *and* future treated districts), this may imply an effect of *preparation* for selective differentiation. For instance, it could be that middle schools that typically send very few of their students to SEP high schools are less proactive about preparing their students for the admissions process (or, vice versa—SEP "feeder" schools may be especially proactive), which could have spillovers for achievement on standardized tests. If there are systematic inequalities in who these middle schools serve, whereby SEP feeder schools disproportionately underserve Black students, then this could result in relatively lower achievement outcomes for Black students prior to differentiation. In order to establish whether these findings may indeed be evidence of negative externalities in the *preparation* for differentiation, particularly for Black students, future research must consider the mechanisms through which the provision of SEP schools generate districtwide effects.

Finally, it is important to acknowledge that SEP schools currently operate under a system of unequal opportunity. Under these less-than-ideal conditions, SEP districts may not be producing their best potential achievement outcomes. Of course, achieving equal access is no doubt a sticky issue (e.g. Orfield and Ayscue 2018), and the likelihood of positive results under such a regime is made somewhat questionable by the evidence from this study that SEP schools currently hinder average districtwide achievement growth even among white students, who tend to be *over*-represented in SEP schools. However, this chapter far from closes the debate on the merits of SEP schools. In particular, although this chapter represents an important starting point in our understanding of the potential spillover effects of SEP schools, the analyses here are still limited by their inability to identify the mechanisms at work. As discussed in the introduction, SEP schools are likely to shape students' educational experiences by shaping the peers they share classrooms with, the curricula taught in those classrooms, and the expectations teachers hold for them, both for students who attend SEP schools and those who attend non-SEP schools in SEP districts. Districts may also make strategic decisions about how to allocate resources based on this process of student differentiation, perhaps concentrating college preparatory resources in a smaller number of schools. Such mechanisms are not easy to measure, particularly using datasets that cover the universe of public schools[56]. In the next chapter, I offer a first look at these questions, using data on the Advanced Placement program from the CRDC.

[56] Appendix Table B2 provides results for an attempt at such a mediation analysis, using available data from the CRDC and SEDA— (1) the percentage of students in the district attending a school where Algebra is offered in seventh grade, as an indicator of the (un)evenness of curricular rigor throughout the districts, and (2) student exposure to low-income peers, based on free and reduced price lunch status, as an indicator of the concentration of exposure to more/less privileged peers. Information on advanced math courses is the only available data on curricular differentiation from the CRDC prior to high school. One might reasonably expect a school's advanced math offerings to be directly shaped by its selective standing within a district, and for this in turn to shape math achievement, making it a reasonable mediator. Likewise, SEP schools might shape the distribution of more/less privileged students throughout the district, in turn shaping things like PTA resources or teacher expectations at different schools, which might then shape achievement. Student socioeconomic segregation is therefore also tested as a mediator, although it may be more an indicator of other mechanisms than a true mediator in itself.

VI. PLACING ADVANCED PLACEMENT IN SEP DISTRICTS: EFFICIENCY AT A THE EXPENSE OF EQUALITY

Decades of research on tracking in the United States has concerned itself with racial inequalities that arise from selectively differentiating students between classrooms within the same school. These studies find that Black and Latinx students are disproportionately likely to be placed in lower-tracked classes relative to their white and Asian peers, and that these placements exacerbate existing inequalities (Gamoran 2010). In SEP districts, students are not only "tracked" within schools, but between schools as well. This is important because between-school differentiation of students may encourage or even require the selective apportionment of finite district resources to schools. On the one hand, this could allow districts to strategically target resources to the same number of students in fewer schools, thereby avoiding costly duplication. On the other, divvying opportunities between schools could exacerbate the racial and socioeconomic inequalities that have long been the focus of research on tracking. Research is therefore needed to understand whether and how SEP districts organize course offerings, and whether these processes reflect the racial inequalities identified in research on within-school tracking. Specifically, this chapter asks whether SEP schools produce a trade-off between the efficiency and inequality of educational opportunity in a district.

This study is guided by the understanding that selective admissions criteria shape the student body at both SEP *and* non-SEP schools throughout SEP districts. It follows that the provision of SEP schools may reasonably be associated with the distribution of resources for advanced academic options *as well as* resources designed to meet needs that selective admissions procedures weed out of SEP schools. For instance, students who meet reading and writing requirements to gain admissions to SEP schools may be less in need of resources for English as a second language, which may in turn become more concentrated in non-SEP schools. Put simply, if schools become more homogenous in terms of students' demonstrated academic preparation, then the breadth of resources deemed essential at any given school may narrow. However, to reflect the resources that are more likely to be explicitly considered as a function of academically selective differentiation—and in keeping with much of the tracking literature—I will focus specifically on college preparatory (college "prep") resources. Accordingly, the analyses below examine whether the provision of SEP schools is associated with districtwide (1) allocation of college prep resources between schools and/or (2) the overall participation and success in college prep throughout the district. Given findings from the tracking literature, I examine whether (3) these effects differ for students of different racial/ethnic backgrounds. Finally, based on this evidence, I will evaluate whether SEP districts appear to either *efficiently* provide resources and/or *exacerbate* inequality.

As indicators of college preparation, I use data on Advanced Placement (AP) enrollments, exam passing, and course offerings from the Civil Rights Data Collection (CRDC). The AP program offers a useful indicator of college preparatory opportunity for a number of reasons. First, it has been the target of state and federal funding for the expansion of college preparatory opportunities, with specific goals of increasing minority access (Schneider 2009). In addition to simply making these courses more widespread and thereby facilitating their study, this indicates that AP courses occupy a relatively important position in the current landscape of college preparation in the United States. Additionally, because they are regulated by the College Board, the curricula in AP courses is likely to be more standardized than courses designed to meet state or local requirements, making them more suitable for nationwide evaluation. Finally, AP courses are typically not required for high school graduation and may demand relatively high resource inputs. As a result, AP courses may be especially vulnerable to resource constraints and discretionary practices. Districts' decisions about how to allocate AP resources in light of these constraints may reasonably be shaped by whether or not students are selectively differentiated between schools. Thus, the AP program provides a useful potential site of discretion to examine the association between SEP schools and the districtwide distribution of resources.

Analyses use Seemingly Unrelated Regression (SUR) to estimate the effect of district structure on AP outcomes for students of different racial/ethnic backgrounds. Because this chapter focuses on high school outcomes, SEP districts are limited to those that offer at least one SEP school at the high school level. I test two measures of resource distribution—(1) evenness of per pupil resources and (2) average course exposure—and two measures of (successful) uptake—(1) the percentage of district students enrolled in one or more AP and (2) the percentage of AP-enrolled students who pass one or more AP exam. I find evidence that AP resources are indeed distributed more narrowly in SEP districts than in non-SEP districts, but that this results in lower course exposure for Black students only. Not only do Black students have access to fewer AP courses than their same-race peers in non-SEP districts, but the white-Black gap in course exposure is also largest in SEP districts. On the other hand, despite relatively similar course exposure and enrollment rates for white and Latinx students across district types, passing rates for these students are *higher* in SEP districts than in certain non-SEP districts (those that also do not offer magnet schools). Once again, these patterns produce a larger white-Black gap in (successful) AP uptake in SEP districts, as well as larger white-Latinx gaps in AP enrollment. Taken together, these findings may suggest that AP courses are distributed more efficiently in SEP districts, but at the cost of greater inequality.

SEP Districts: Efficient Specialization or Effectively Maintained Inequality?

Research on tracking has placed considerable emphasis on racial differences in track placements. This research has produced relatively disparate results in terms of whether under-represented minority (URM) students are indeed under-represented in college preparatory courses after controlling for prior achievement (e.g. Lucas and Berends 2007; Mickelson 2001). These inconsistencies may in part reflect differences in tracking practices between different types of schools. For instance, research has suggested that patterns of tracking look different in private schools or in schools with different levels of diversity (Lucas and Berends 2002). Importantly, SEP schools differ from traditional course-level tracking in the scope of their differentiation (i.e., that they are whole schools). This means that students in SEP schools spend all of their instructional time exclusively with peers who have been similarly selected. Beyond sorting students, this process also sorts parents and resources. This may have implications not only for the resources available to students (of different racial/ethnic backgrounds) within schools, but between schools as well.

Decisions regarding resource allocation in U.S. public schools are made in the context of finite, and often limited, funding. Duplicating resources for which few students have need may therefore present an impractical or even unworkable financial burden. Importantly, the same per pupil expenditures need not buy the same resources in different schools, and discrepancies are likely to reflect the perceived needs of students in each school. For instance, schools serving more advantaged students may have fewer needs in terms of non-instructional staff, such as teacher aids or on-site social workers. This might allow these schools to spend a greater share of their budget on new books and science lab materials, or to support the higher salaries of a more experienced teacher corps. If school differentiation is executed in such a way that all students with the preparation and motivation to take advantage of these resources still have access to them, then this could be a *feature* of SEP districts, rather than a bug, allowing them to provide college preparatory resources more efficiently.

However, rather than operating in a purely "efficient" manner, research suggests that these processes of resource distribution are likely to be shaped by the efforts of white and/or socioeconomically advantaged parents to hoard opportunities and maintain advantages for their children. Effectively Maintained Inequality posits that advantaged parents will pursue advantages for their children wherever they are available (Lucas 2001). An important implication of this theory is that as the quantity of a resource increases and becomes more widely available (e.g., as a high school education becomes more universal) privileged parents will pursue advantages in quality (e.g., advanced college preparatory resources and track placements). These processes exist at the individual student level—for instance, as persistent parents seek entry for their (average-performing) children in advanced courses (e.g., Lewis and Diamond 2015)—but have also been shown to shape school-level resources. For instance, Klugman (2013) shows that as the AP program became more common throughout California, schools with a greater middle class presence maintained their advantage by increasing the variety of AP courses they offered. Further, research suggests that this kind of opportunity hoarding can be exacerbated by criteria-based school choice policies, which can be gamed through additional investments of time or money, such as tutoring or test preparation for exam-based admissions (Sattin-Bajaj and Roda 2018).

At minimum, this research suggests that parents in SEP districts may be exceptionally well situated to hoard opportunities for their children, simply through the process of school choice and selective admissions. Importantly, the concentration of privileged parents in SEP schools may also make it easier to expand the opportunity gap beyond admissions. Schools with better resourced parents benefit from greater capacity for organizing and private fundraising, which can increase discretionary income beyond per pupil public funding. Bolstered by the guise of meritocratic admissions, this may especially enhance the pressure SEP schools exert on district decision-making processes regarding the allocation of college preparatory resources. Rather than simply allowing districts to distribute resources more *efficiently*, then, SEP schools may exacerbate patterns of EMI.

In order to evaluate whether the allocation of college prep resources in SEP districts reflects either greater efficiency and/or EMI, it is important to consider what resources might allow measurement of these phenomena. An ideal resource would have at least three qualities. First, it would be nationally standardized. This is important so that presence of the resource indicates the same quality of college prep across districts and schools. For reasons discussed above, per pupil funding does not meet this criterion. Second, there should be adequate variation in the provision of the resource at different schools within a district. In other words, districts should have some discretion in where and to what extent the resource is provided. Finally, particularly to evaluate EMI, an ideal resource should offer a mark of distinction for students, such that privileged parents are likely to seek it as an advantage. The AP program meets all of these requirements. It also has the added appeal of having been the target of equity initiatives, such that we might expect districts to be cognizant of inequalities in the distribution of these resources between schools. I turn next to a brief history of this program.

Placing Advanced Placement

The Advanced Placement (AP) Program was developed in the 1950s to offer opportunities for challenge and acceleration to the country's "best and brightest" students in the hopes that they could more quickly take on leadership roles and advance U.S. interests during the Cold War (Kolluri 2018; Schneider 2009). In its early days, the AP program was offered exclusively at the most elite high schools in the country, but it has since grown dramatically. As demand for a college education increased, students began to pursue additional marks of distinction to set them apart in the application process. However, early expansion of the program remained centered in private schools and elite suburban public schools. As the Civil Rights movement got underway, critics began to raise important concerns about racial equity in access to this new resource, but expansion of the program continued primarily under the impetus of demand from privileged families until the 1980s (Schneider 2009).

By the 1990s, state and local education agencies were spending considerable resources to supplement the expansion of the AP program. Over the course of the decade, the federal government spent about $2.5 million per year on the proliferation of the AP program, with an emphasis on expanding access for underserved students (Kolluri 2018; Schneider 2009). Even as the policy focus shifted towards proficiency with the passage of No Child Left Behind in 2002 and the Javits Act[57] struggled to remain funded, AP proliferation received continued support in the form of both exam fee subsidies and grants to fund studies on AP expansion in high poverty schools (Klopfenstein 2004). As of 2016, twenty states offered financial incentives to support AP course offerings or success (Education Commission of the States n.d.), and eight states and the District of Columbia had requirements for the provision of AP courses at the district or school level (Education Commission of the States n.d.)[58]. As a result, participation in AP exams increased from 330,000 students in 1990 to 2.2 million in 2013.

This push for expansion—as a product of both student demand and incentive programs—clearly indicates the importance placed on the AP program as a gateway to postsecondary opportunities. This is reflected in national rankings, issued annually by *Newsweek* and *The Washington Post*, which rate schools exclusively by the ratio of AP, IB and Cambridge tests to graduating seniors[59,60] (Matthews 2017; Schonfeld 2014). Major state university systems have also placed a premium on AP coursework. For instance, the University of California guarantees admissions to the top 9% of applicants based on an admission index that assigns extra points for AP (or IB) coursework (*UC Statewide Guarantee* n.d.). Even as 'holistic reviews' become more common in elite university admissions, whereby student achievement is evaluated relative to the challenges and resources available to them, access to a wider variety of AP courses is likely beneficial in allowing students both to demonstrate rigor *and* to package a coherent "story" about their interests and experiences for admissions committees (Stevens 2007). To many students and parents, then, access to AP courses may seem high stakes.

It is worth noting that existing research offers little evidence that AP courses actually improve college performance (Klopfenstein and Kathleen Thomas 2009). Nevertheless, access to

[57] See Chapter 2 for brief history of the Jacob K. Javits Gifted and Talented Students Education Act of 1988.

[58] An additional six states have requirements for some form of advanced coursework, including APs, International Baccalaureate, Advanced International Certificate of Education (AICE), Cambridge, or dual enrollment.

[59] AP exams are by far the most common of these assessments.

[60] Newsweek added a second ranking system in 2014.

these courses may be important for the reasons just described. Moreover, the availability of AP courses may be associated with the provision of other college prep resources that are more difficult to measure, including additional rigorous coursework or even students' expectations about whether and where they should apply to college. Finally, the effects of AP coursework can only be studied in schools where AP courses are offered. Although the AP program has expanded significantly since the 1990s—and is available in the vast majority of schools considered in the following analyses—it is still not available in all schools[61] and the breadth of the program varies widely across schools. It is therefore important to consider how this resource is distributed between schools at the district level, and whether this is associated with the provision of SEP schools.

Data and Analysis

To evaluate differences in the distribution of AP resources between SEP and non-SEP districts and for students of different racial/ethnic backgrounds in those districts, I produce a series of Seemingly Unrelated Regression (SUR) estimates. These models draw on AP data from the Civil Rights Data Collection (CRDC), which include course counts, enrollments, and exam passing by race in each high school in the United States. Because AP courses are a high school resource, these analyses include only SEP districts that offer at least one SEP school at the high school level. In total, I identified 212 SEP high schools across 62 districts[62]. My modelling approach, which includes controls for eighth grade resource distribution and achievement, will be described in more detail below. First, I describe the dependent variables used in these analyses.

Dependent Variables: AP Allocation, Uptake and Success

This paper asks whether offering one or more SEP high school is associated with (1) the distribution of college prep resources across schools and to students of different racial/ethnic backgrounds, and (2) whether any potential concentration of resources might be characterized as an increase in either efficiency or opportunity hoarding. To answer these questions, I draw on three sets of variables from the 2015-16 CRDC about the Advanced Placement Program at each school: (a) the number of AP courses offered, (b) the number of students enrolled in at least one AP class, by race, and (c) the number of students passing at least one AP exam, by race.

The number of AP *courses* refers to a count of unique curricular offerings (i.e., AP

---

[61] Overall, about a third of high schools in the U.S. offer no AP courses. This rate is cut by half in the analytic sample for this chapter.

[62] Consistent with my definition of SEP schools as whole schools operated by a traditional public school district (i.e., one that includes neighborhood schools), I omit selective schools operated by state-level or consortium districts. Several states offer public boarding schools for highly capable students, often with a focus in math or science. Similarly, consortium districts, such as Monmouth Vocational District in New Jersey and the Capitol Region Education Council in Connecticut, connect several traditional districts to offer a system of specialized options. Although these schools are selective, their districts do not include traditional neighborhood-based public schools, such that the question of selective resource distribution posed here would not carry the same meaning. As a result of the omission of these districts, my list of SEP high schools includes only about 100 of the 165 "exam schools" identified by Finn Jr and Hockett (2012). Of the 60 districts in my sample, 45 were identified in Finn and Hockett's list.

Calculus AB counts as one course, regardless of whether there is one class section or five). The College Board offers 36 subject area courses, as well as 2 research-based "Capstone" courses, for a maximum of 38 courses. For schools that offer the International Baccalaureate program, I add an additional 8 courses[63] to a school's total AP course count. I use this count to define two key measures of resource distribution, which are motivated by classic measures of segregation (e.g. Massey and Denton 1988). The first is an indicator of exposure, and is modelled after the Interaction Index. It is an enrollment-weighted average of the *count* of AP courses offered at each high school in the district. The second is an indicator of evenness, based (loosely) on the Dissimilarity Index, which measures the resources available at the average student's school as a *percentage of the maximum* resources offered at any single school in the district.

The traditional Interaction Index (as measured in Chapter 4) is a population-weighted average of one group's exposure to another[64]. Here, I calculate a population-weighted average of one group's exposure to the *number* of AP courses in their school. I define district-level AP course exposure as:

$$ {}_rP_{AP} = \sum_{i=1}^{n} AP_i * \frac{r_i}{R} \tag{6.1} $$

where,

i indexes schools in the district from 1 to n
AP_i is the number of AP (and/or IB) courses offered in school i
r_i is the number of 9th graders of race r in school i, and
R is the total number of 9th graders of race r in the district.

I weight by ninth grade enrollments[65], rather than total enrollments, to avoid bias from differential dropout rates across districts or by race.

This measure can be interpreted as the number of AP courses offered in the average student (of race r)'s school. This is extremely useful as a baseline measure of students' opportunity to take AP classes. However, as an enrollment-weighted average, it presents a few challenges. First, it is likely at least somewhat mechanically related to the size of a student's school. While core classes like math, science, English, and history must be offered at every grade

[63] The IB Diploma Program curriculum includes six subject areas, plus "Theory of Knowledge". Schools may or may not provide multiple options for students to meet these six area requirements, but eight courses is likely a conservative estimate of the IB courses offered at each school with a certified Diploma Program. Only 892 schools reported IB enrollments to the CRDC in 2015-16 (as of 2020, the IB organization reports 948 Diploma Program schools in the U.S.).

[64] The interaction index is defined by Lieberson as: ${}_xP_y = \sum_{i=1}^{n} \left[\frac{x_i}{X}\right]\left[\frac{y_i}{t_i}\right]$, where x_i, y_i, and t_i are the numbers of the minority, majority, and total populations in unit i, respectively, and X is the total population of minority group X (Lieberson 1981, Massey & Denton 1988).

[65] In about 6% of districts, there is an additional opportunity for school transitions after ninth grade, such that AP courses are offered at schools that do not have any ninth graders. In these districts, I weight instead by enrollment in the final possible grade of transition into a school with AP courses, approximately 4% in tenth grade and 2% in eleventh.

level, making it relatively easy to support AP courses in these areas regardless of the number of students, schools with more students will have a greater ability to support a wider array of electives. To account for this, models control for school size. A second challenge is that greater AP exposure may not only reflect differences in patterns of resource *distribution*, but also difference in overall *districtwide* levels of AP resources (i.e., the total pool of resources being distributed). Thus, I define a second measure of access to APs that more directly captures the difference in resources between schools in the same district.

The second measure of college prep resource allocation is inspired by measures of evenness. Traditional evenness measures, like the Dissimilarity Index (also used in Chapter 4) or the Entropy Index, compare the minority *proportion* in a given subunit to the *average* minority proportion in the larger geography[66]. My approach differs for two key reasons. First, as with the measure of ${}_rP_{AP}$, the distribution I am interested in is not the proportion of a group, but rather a quantity (count) of a resource. Second, I choose to benchmark a school's AP resources against the *maximum* AP resources offered at any single school in the district, rather than the average. I do this so that the direction of the deviation is evident in the estimate (every school has either the same amount or *fewer* AP courses than the school with the most). I find this particularly helpful because even some of the largest districts in the country operate only a handful of high schools, meaning that a normal distribution above and below the mean is unlikely. Traditional measures of evenness would take on the same value whether the distribution were skewed to the right or the left. This is not ideal for the purposes of understanding whether districts provide better or worse access to AP courses. Thus, I define district-level AP evenness as:

$$E_r = \sum_{i=1}^{n} \frac{AP_i/t_i}{\chi} * \frac{r_i}{R} \quad , \qquad \chi = \max\left(\{AP_i/t_i : i = 1, \dots, n\}\right) \tag{6.2}$$

where,

t_i is the total number of 9th graders[67] in school i

I scale by the total number of students here because this index directly compares resources in one school against those in another. This adjustment is important because SEP high schools—which are likely to have a more targeted college preparatory focus—tend to serve fewer students than the average high school in their district, and raw course counts may therefore make for a poor apples-to-apples comparison. I interpret this index as the per-pupil AP resources available in the average r-race student's school as a percentage of the greatest per-pupil AP resources available at any single school in the district.

Finally, the CRDC includes school-level data on students who (1) enroll in at least one AP course, and who (2) pass at least one AP exam. I use these data to measure the effectiveness

[66] E.g., dissimilarity is given by, $D = \sum_{i=1}^{n} \frac{|p_i - P|}{2P(1-P)} * \frac{t_i}{T}$, where t_i is the total population in subunit i, T is the total population overall, p_i is the minority proportion in subunit i, and P is the total minority proportion. And the entropy index is $H = \sum_{i=1}^{n} \frac{(E-E_i)}{E} * \frac{t_i}{T}$, where E is the total entropy, given by $E = (P)\log[1/P] + (1-P)\log[1/(1-p)]$ and E_i is the comparable measure for unit i (Massey & Denton 1988).

[67] See footnote 65.

of a districts' college prep offerings. Rather than a distribution of resources, these measures are intended to capture a district's overall performance in providing college prep resources. Accordingly, I use simple measures of prevalence. By summing across schools, I generate district-level measures of (1) the percentage of students enrolled in at least one AP, and (2) the percentage of AP-enrolled students who pass at least one exam. Data on IB enrollments are not sufficiently detailed in the CRDC, so these measures focus exclusively on the AP program.

$$\%AP\_enr_r = \sum_{i=1}^{n} AP\_enr_{ri}/R \tag{6.3}$$

$$\%AP\_pass_r = \sum_{i=1}^{n} \frac{AP\_pass_{ri}}{AP\_enr_{ri}} * \frac{r_i}{R} \tag{6.4}$$

Analytic Approach

The following analyses test whether SEP districts shape patterns of resource allocation (${}_rP_{AP}$ and E_r) and whether these patterns reflect an increase school system efficiency and/or EMI. In particular, I evaluate whether white parents are better positioned to maintain advantages over Black and Latinx students in SEP districts. To answer this question, outcomes must be evaluated for each racial/ethnic group. In other words, this requires repeated measures of the same outcome for each district (e.g., ${}_{white}P_{AP}$, ${}_{Black}P_{AP}$, ${}_{Latinx}P_{AP}$). Because models for these redundant dependent variables would have highly correlated error terms, analyses are conducted using Seemingly Unrelated Regression (SUR) to increase efficiency. In addition to increasing efficiency by taking into account the correlation of errors between outcomes, SUR is well-suited to these analyses because it allows formal hypothesis tests (Wald tests) for coefficients across race-specific equations (Felmlee and Hargens 1988).

Because two of the four outcomes of interest are measures of between-school distribution of resources, I limit the sample to districts with at least five high schools[68], so that there is sufficient potential for variation[69]. Further, in order to facilitate the investigation of meaningful racial differences in outcomes, the sample is restricted to districts with at least 200 high school students of each racial/ethnic background. The logic here is the same as the school count restriction—to ensure sufficient numbers of students of each racial/ethnic group in at least a handful of schools, so that measures can be calculated with reasonable precision. Because dropping districts with fewer than 200 Asian students would greatly reduce the sample[70], I omit Asian students from these analyses. Analyses therefore focus exclusively on white, Black, and Latinx students.

---

[68] I exclude alternative schools, special education schools, virtual schools, and juvenile justice centers, as well as schools with fewer than 20 students.

[69] This is slightly different from my approach to analyzing treatment with the DiD in the prior chapter because (1) schools tend to be larger (and thus fewer in number) at the high school level, making five schools a more restrictive criterion at this level, because, (2) I do not have similarly good data on the presence of charter schools in high school, so I omit this filter, and (3) because I do not have "future" SEP districts as a falsification test, I use districts with magnet schools as my next-most-similar districts, and therefore need non-magnet districts to compare these against.

[70] Additionally, race-specific demographic controls are not available for Asian students.

The key independent variable in the models below is a binary indicator of whether a district offers at least one SEP school at the high school level. The comparison group, as just described, is any district with at least five high schools. To better understand what qualities of SEP districts might shape decisions regarding college prep resource allocation, I further divide the comparison sample into (1) magnet (*n*=143) and (2) non-magnet (*n*=120) districts. The magnet district sample includes districts with whole school magnets and/or magnet programs. These magnets may offer specialized curricula, but because their admissions process is non-selective and/or low in scope[71], there are fewer opportunities for privileged parents to spend time and money to hoard advantages for their children. Therefore, this series of district comparisons is particularly relevant to establishing whether differences in AP availability and success reflect a process of Effectively Maintained Inequality. If selectively sorting students between schools facilitates EMI, we would expect inequalities to be greater in SEP districts than magnet districts, relative to non-magnet districts.

Models include a set of district-level controls from the CRDC, Education Demographic and Geographic Estimates (EDGE), and Stanford Education Data Archive (SEDA). From EDGE, I include controls for district geography (urban, rural, suburban) and demographic characteristics of all school-aged children including percent in owner-occupied homes, logged median value of owner-occupied homes, and percent foreign born. Controls from SEDA pertain to eighth grade district characteristics and include logged total expenditures, percent of students eligible for free or reduced lunch, percent of students in charter schools*, percent of 5-17 year olds living in poverty*, the 50/10 income ratio*, percent whose parents had a BA or more*, and percent of students living in the same house as the previous year*. Race-specific versions of the controls with asterisks are included in the race-specific equations that compose each SUR system of equations. Race-specific controls (from SEDA) also include each group's proportion of the public school population in eighth grade, exposure to white students, and student teacher ratios. At the high school level, I control for the number of high schools and the average school size, using data from the CRDC.

Finally, models include relevant controls for eighth grade achievement and resource exposure. All models include overall and race-specific measures of mean and standard deviation of eighth grade achievement (pooled for math and English Language Arts). Models predicting resource distribution (${}_{r}P_{AP}$ and E_r) include race-specific measures of eighth grade exposure to peers passing end-of-course exams in Algebra. Models predicting average AP participation and success ($\%APenr_r$ and $\%APpass_r$) control for race-specific eighth grade Algebra pass rates. Data on Algebra passing come from the CRDC and are the best available measure of (successful) advanced course completion prior to high school. These are important controls because research suggests that AP course offerings are strongly correlated with the math achievement of rising ninth graders (Iatarola, Conger, and Long 2011). I use these two different measures of the

[71] Note that I identified an additional 20 districts with selective *programs* at the high school level, which I have omitted from my list of SEP districts because they do not include *whole schools* with selective assignment procedures. These districts are not omitted from the sample of magnet districts. However, these programs are actually roughly evenly split between magnet and non-magnet districts in the sample (not all of the schools with these programs self-report as magnet programs in the CRDC), so I would expect their presence across the two groups to be offsetting and therefore do not expect the existence of these programs to significantly impact findings.

prevalence of Algebra passing in order to parallel the construction of the dependent variables as closely as possible.

Results

This section presents both descriptive statistics and SUR model results. Model results are presented in graphical form for ease of comparing expected outcomes for different racial/ethnic groups across different district types. Coefficients and additional Wald tests for SUR models can be found in Appendix Tables C1 and C2. These tables present findings for multiple model estimates for each outcome. The first includes only district demographic controls, and the second adds controls for eighth grade measures of achievement and prevalence of success in Algebra[72]. The figures presented below represent model results from estimates including controls for eighth grade achievement and Algebra. Results are not significantly altered by the addition of these controls.

Concentrated College Prep in SEP Districts

The first research question asks whether SEP districts distribute college prep resources differently than non-SEP districts. To begin to answer this question, I first evaluate whether SEP schools themselves appear to benefit from a greater supply of college prep resources than other schools in their district. Figure 6.1 shows the average number of AP/IB courses taught in SEP and non-SEP schools *in SEP districts*. Unsurprisingly, SEP schools appear to enjoy a relatively large allocation of college preparatory resources relative to non-SEP schools in the same district, in terms of the number of AP and IB courses offered. The average SEP school, irrespective of school size, offers 11 AP/IB courses. By contrast, the average non-SEP school in SEP districts offers only about 7 AP/IB courses. This relationship is essentially unchanged if we consider only AP courses[73]. This demonstrates what we would expect intuitively—that SEP schools appear to have a strong emphasis on providing college preparation. However, this still leaves a few questions unanswered. First, although Figure 6.1 suggests that SEP schools offer more AP/IB courses on average, it does not directly test whether SEP schools tend to be the *best*-resourced schools in their districts. And second, it does not address whether similar disparities exist between more- and less-resourced schools in non-SEP districts.

SEP schools do in fact tend to have the *most* AP/IB resources among high schools in their own district. Appendix Figure C1 shows that, of the 62 high school SEP districts in this analysis, 47 offer the most AP/IB courses in SEP schools, either in per pupil or absolute terms. Of the remaining SEP districts, 5 offer "early college" SEP high schools that offer no AP or IB courses, but enroll students directly in college courses. This leaves 9 of 62 districts for which SEP schools

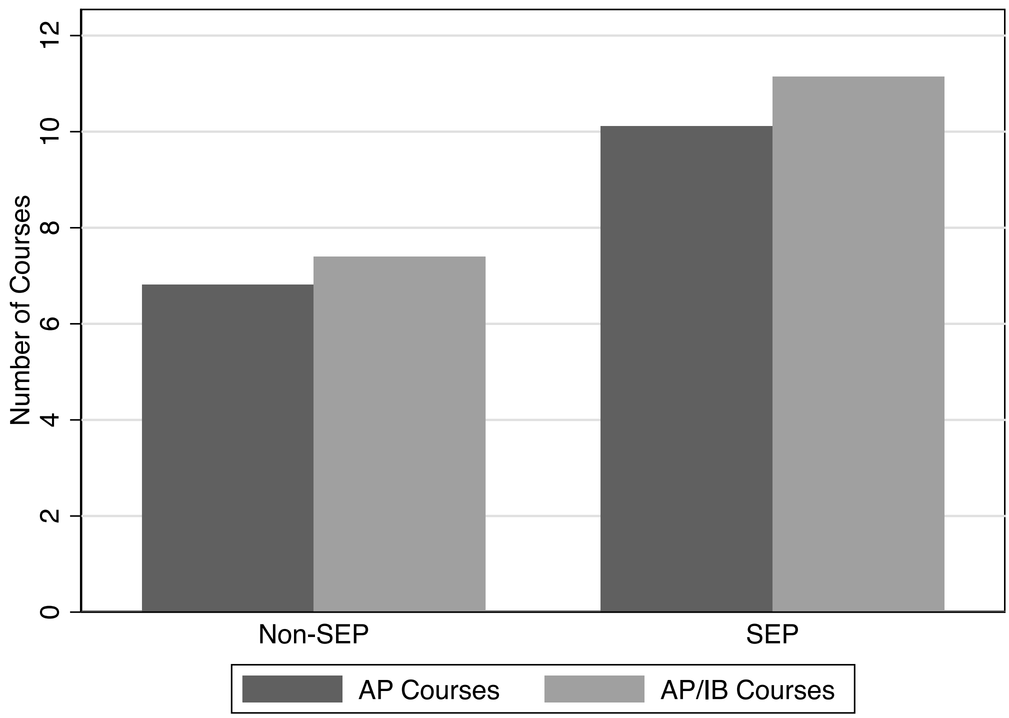
---

[72] A third set of models splits the magnet district comparison group into districts that offer only partial-school magnet programs (*n*=37) and those that offer whole school magnets (*n*=87). These results are described briefly in the discussion.

[73] The vast majority of advanced coursework comes from the AP program. Even among SEP schools, only 1 in 8 offer the IB program. Thus, while it is useful to consider the *total* variety of coursework offered (AP and IB) to generate a complete picture of resource differences, the models that estimate overall district-level participation and success in college prep (AP) classes are unlikely to be significantly affected by the lack of data for the small percentage of students who take/pass one IB class/exam but *no* AP classes/exams.

do not offer exceptional resources (and 2 with missing data), according to the CRDC[74]. Thus, while not uniformly true, SEP schools do typically offer the most per pupil college prep resources in their district.

Figure 6.1. Average AP/IB Courses offered in SEP Districts by School Type

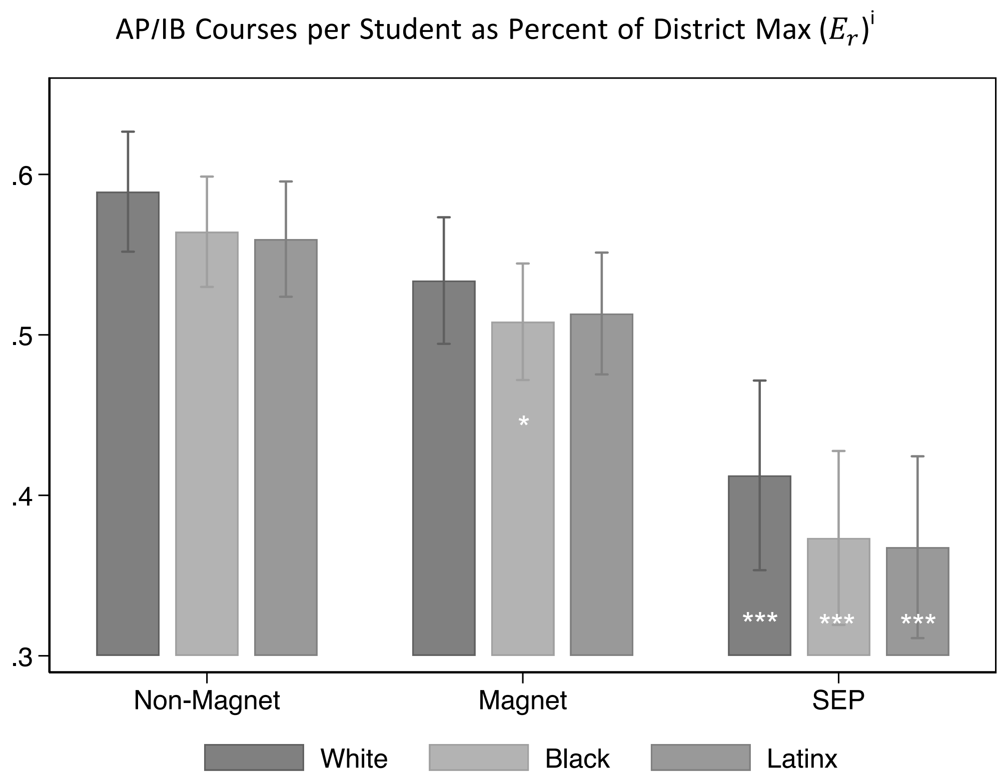


The next question is whether these between-school disparities are unique to SEP districts, or whether they exist in non-SEP districts as well. For this, I turn to the SUR estimates of per pupil AP/IB resources evenness (E_r). Figure 6.2 depicts the marginal estimates of AP/IB evenness for students of different racial/ethnic backgrounds in non-magnet, magnet, and SEP districts. Recall that this measure can be interpreted as the *per pupil* AP/IB resources in the average student's school as a *proportion* of those in the best-resourced school in the district. Model coefficients and additional Wald tests can be found in Appendix Table C1. This figure shows that all districts have resource imbalances between the average school and the best-resourced school. However, this imbalance is worse in SEP districts. In SEP districts, the average student attends a school with roughly 40% of the per pupil resources that are offered in the best-resourced schools. This is significantly different from the expected resource evenness (around

[74] In these districts where SEP schools do not offer the most resources, the "best-resourced" schools tend to be exceptionally small, driving up the ratio of AP courses to students. Moreover, some of the best resourced non-SEP schools in SEP districts are *partially* selective. For instance, highly-resourced iPrep academy in Dade (Miami) is a K-12 school that does not have academic admissions criteria in kindergarten (and so is not considered a SEP school) but *does* have selective admissions criteria for students applying in middle or high school.

55%) in non-magnet districts. Importantly, this relative unevenness of *per pupil* resources in SEP versus non-magnet districts is experienced relatively similarly (i.e. there is no statistically distinguishable difference) by students of all racial/ethnic backgrounds.

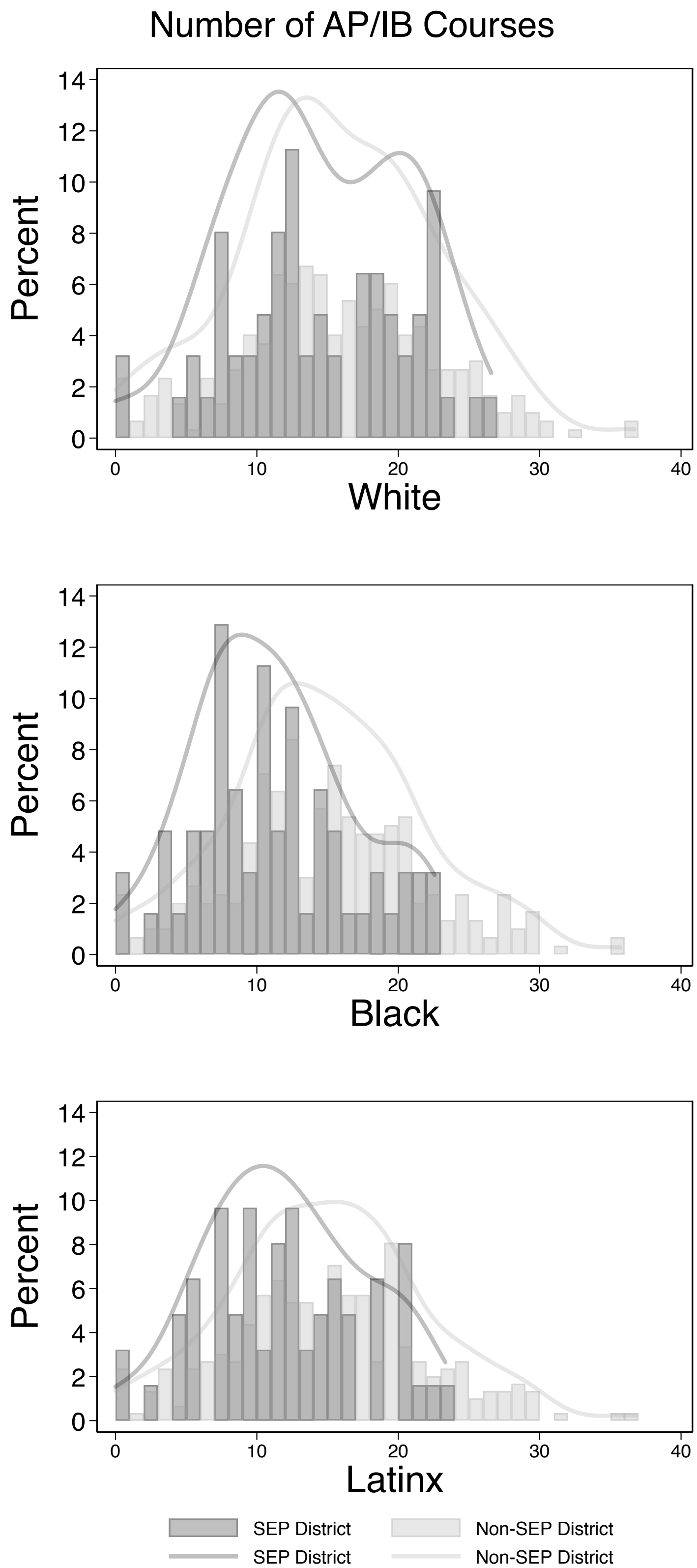
Figure 6.2. Marginal Estimates for SUR of Average AP Exposure



[i] Note: y-axis is truncated for ease of visually comparing confidence intervals.
Wald Tests for Magnet/SEP(race)=Non-Magnet(race): * p<0.05, ** p<.01, *** p<.001
Wald Tests for SEP(Black/Latinx)=SEP(white): † p<0.05, †† p<.01, ††† p<.001

In keeping with Figure 6.1, this finding likely primarily reflects the fact that SEP districts offer at least one school with *exceptional* resources, against which other schools appear relatively poorly supported. Note that there is no evidence that this same allotment of exceptional resources occurs in magnet districts. The evenness of AP resources in magnet districts is statistically indistinguishable from non-magnet districts for white and Latinx students and significantly higher than in SEP districts for all racial/ethnic groups (additional Wald tests presented in Appendix Table C1). This suggests that specialized programming alone does not garner the same exceptional investment of college preparatory resources if students are not selectively differentiated into these schools. Importantly, this does not elucidate the quantity of resources available in these "exceptional" schools, or by extension, whether the average resource exposure in SEP districts actually differs from that in non-SEP districts. To get a better understanding of this, I evaluate student exposure to total AP/IB course offerings.

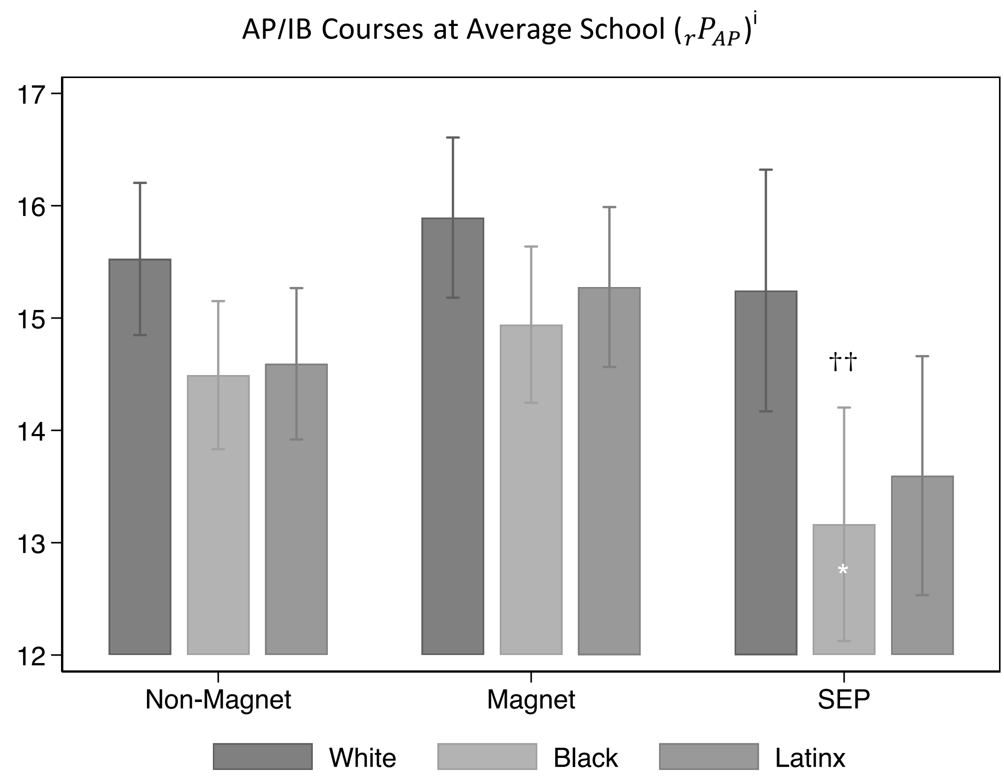
Figure 6.3. Number of AP/IB Courses in Average Student's School, by Race



Average College Prep Resource Exposure

Figure 6.3, above, illustrates the unadjusted AP/IB course exposure (${}_{r}P_{AP}$) in SEP versus non-SEP districts for white, Black and Latinx students. I combine magnet and non-magnet districts for simplicity. These distributions illustrate wide variation in AP course offerings across districts overall, and significant overlap in the distributions for SEP and non-SEP districts. However, this overlap appears somewhat stronger for white students than for Black and Latinx students. For white students, the distribution takes on a different shape, but generally centers around the same point for SEP and non-SEP districts. For minority students, by contrast, average exposure to AP courses appears to be somewhat lower in SEP districts, as evidenced by the shift of the entire distribution to the left of those in non-SEP districts. This implies that it is not just the evenness of resources—or the *relative* quantity in the average school—that differs between SEP and non-SEP districts. Rather, the abundance of resources in SEP schools may come at the expense of the *absolute* quantity of resource in traditional public schools, at least for minority students. SUR estimates examine whether this pattern holds after controlling for district characteristics, including average school size, eighth grade achievement, and eighth grade exposure to advanced coursework.

Figure 6.4. Marginal Estimates for SUR of Average AP/IB Course Exposure



[i] Note: y-axis is truncated for ease of visually comparing confidence intervals.
Wald Tests for Magnet/SEP(race)=Non-Magnet(race): * p<0.05, ** p<.01, *** p<.001
Wald Tests for SEP(Black/Latinx)=SEP(white): † p<0.05, †† p<.01, ††† p<.001

Whereas Figure 6.2 suggests relatively similar effects of SEP schools for students of different racial/ethnic backgrounds in terms of the evenness of per pupil resources, Figure 6.4 (see Model 5 in Appendix Table C1) suggests important racial differences in absolute resource exposure. Recall that ${}_rP_{AP}$ can be interpreted as the number of AP/IB courses offered at the school attended by the average student of race r. One key finding from Figure 6.4, then, is that there is evidence that Black students in SEP districts attend schools with 1 to 2 fewer AP courses than the average Black student in non-magnet or magnet districts. This effect is statistically significant and, given an average offering of about 15 AP courses (for a high school with 1600 students) in non-SEP districts, it is substantively large as well. To further scale this effect, note that this is about the equivalent difference in AP course exposure associated with an increase in average school size of 300 students, or greater than a one standard deviation increase eighth grade exposure to peers passing Algebra, which are themselves strong predictors of AP course exposure (not shown).

For white and Latinx students, there is no statistically significant difference in the exposure to AP resources between same-race peers in SEP versus non-magnet district[75]. These differential within-race effects have further implications for between-race inequalities. Specifically, in addition to the fact that Black students experience lower AP exposure relative to same-race peers in non-SEP districts, the gap between Black and white students in AP exposure is significantly larger in SEP districts. The white-Black gap in average AP course exposure is about 1 course in non-SEP (non-magnet and magnet) districts and about 2 courses in SEP districts. The white-Latinx gap is only marginally larger ($p<0.069$) in SEP districts (about 1.5 courses) than in non-magnet districts (about one course), but is significantly larger than in magnet districts (which enjoy the greatest average AP exposure). Importantly, because the *average* AP course exposure is lower for Black and Latinx students in SEP districts, these gaps represent an even greater disproportionality in resources for white students in SEP districts.

Effectiveness of Resource Distribution

Findings from Figures 6.2 and 6.4 suggest that AP resources are distributed more unequally in SEP districts than non-SEP districts. In the remaining analyses, I consider whether the lower AP exposure for URM students in SEP districts arises because these districts divide similar resources differently, or whether they may (also) deploy fewer total resources towards the AP program[76]. If *only* the latter were the case, we would likely expect to see disadvantages in AP exposure for white students in Figure 6.4 as well. Although we do not see this, findings from Figure 6.4 cannot completely rule out the possibility that SEP districts invest less overall in the AP program. To gain more purchase on this question of total AP investment, this section evaluates overall district *uptake* of AP courses.

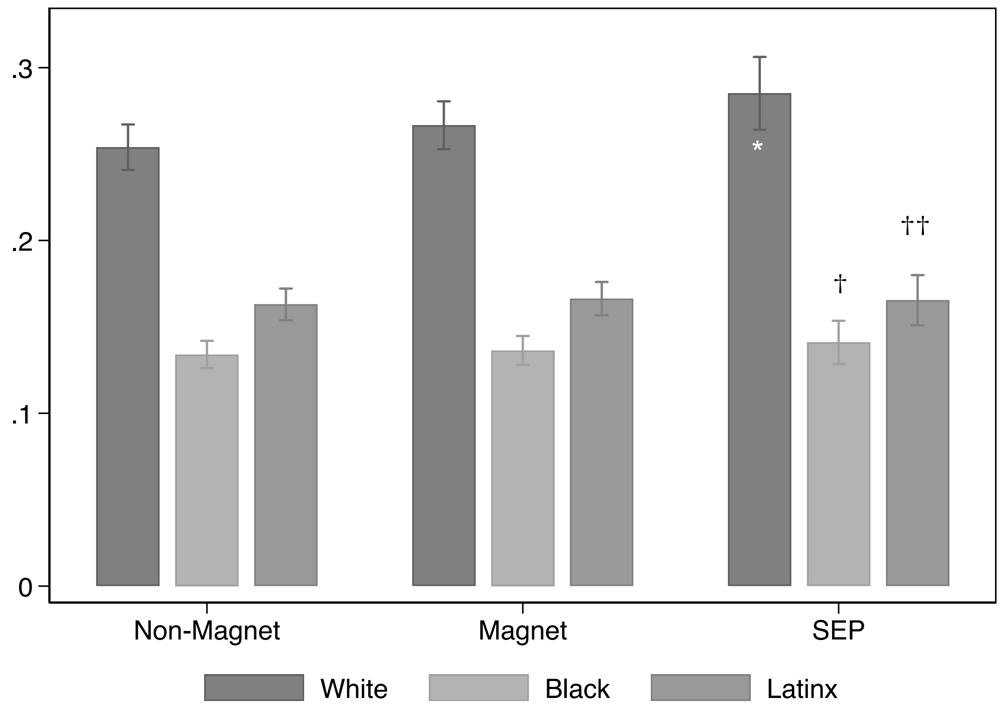
---

[75] However, due to a small increase in AP exposure in magnet districts (not statistically distinguishable from non-magnet districts), there is a significant (negative) difference in exposure for Latinx students in SEP relative to magnet districts.

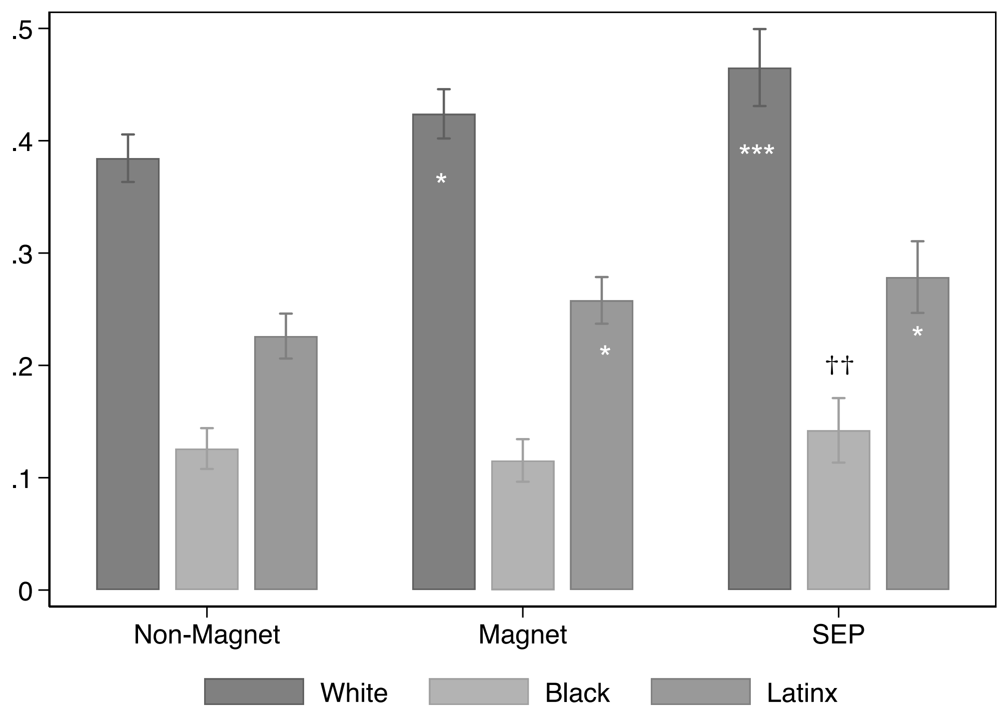
[76] This is distinct from the question of a district's *net* fiscal resources (which are controlled using median home value, total eighth grade expenditures, and other district income/poverty measures), and is specifically a question of resources put towards the AP program.

Figure 6.5. Marginal Estimates for SUR of District AP Efficiency

Panel A: AP Enrollment



Panel B: AP Exam Passing Among Exam-Takers



Wald Tests for Magnet/SEP(race)=Non-Magnet(race): * p<0.05, ** p<.01, *** p<.001
Wald Tests for SEP(Black/Latinx)=SEP(white): † p<0.05, †† p<.01, ††† p<.001

Panel A of Figure 6.5 presents results for the total district-wide enrollment in at least one AP course, net of eighth grade achievement. This figure shows a 3 percentage-point increase in overall AP enrollments for white students in SEP districts relative to their same-race peers in non-magnet districts ($p<0.05$). This enrollment bump occurs despite the fact that AP exposure is the same for white students in SEP districts as in non-SEP districts. Given an average AP enrollment around 25% in non-magnet districts, this reflects over a 12% increase in AP enrollments for white students. There is no equivalent increase in AP participation for Black or Latinx students; however, there is also no penalty. In other words, although Black students have less exposure to AP courses than same-race peers in non-SEP districts, their overall rates of enrollment are no different. Altogether, this suggests that despite apparently targeting AP courses more narrowly, SEP districts are still able to achieve the same—or even higher—participation. This might imply that SEP districts are in fact distributing both students and resources strategically, in order to achieve the same outcomes. To establish whether this is in fact more efficient, it is important to consider how students perform once enrolled.

Panel B of Figure 6.5 shows that SEP districts produce a greater proportion of white and Latinx AP exam passers (among enrolled students) than non-magnet districts, again holding eighth grade achievement constant. White students who enroll in at least one AP course are 7.7 percentage points more likely to pass at least one exam and Latinx students are 4.9 percentage points more likely than same-race peers in non-magnet districts. These increases represent a substantively large relative advantage over non-magnet districts of around 20%.

There are at least three reasons we might expect exam passing rates to improve: (1) AP courses are taught more effectively, (2) students are able to better match to AP courses for which they have an interest or are most prepared, or (3) higher levels of preparation are required for entry into AP courses in the first place. Based on Panel A, there is no evidence of the latter—that SEP districts are more selective overall in their AP student body. In fact, a *larger* proportion of white students enroll in AP courses in SEP districts, net of average eighth grade achievement. This offers further evidence that the resource distribution in SEP districts might be strategically efficient.

However, neither of these SEP district advantages—either in enrollment rates or rates of exam passing—is shared by Black students. Black students perform the same on these indicators in SEP and non-SEP districts. One could argue that, because SEP districts are able to achieve the same outcomes for Black students even while targeting resources more narrowly, this might still be characterized as improved efficiency. It is important to consider, however, what costs this efficiency might have for inequality. Importantly, even though no individual group appears to be harmed relative to same-race peers in other districts, differences in these effects by race translate to larger white-Black *and* white-Latinx gaps for AP enrollments in SEP districts relative to non-magnet districts. The white-Black gap also increases for passing rates among AP-enrolled students in SEP districts relative to non-magnet districts. In other words, while it is plausible that efficiency of resource distributions is increasing for all groups, the benefits of this efficiency do not appear to be shared equitably. Rather, white students appear relatively better situated to take advantage of the resources to which they have access. This pattern may suggest that SEP schools exacerbate EMI at the district level.

It is worth noting, however, that all of the significant findings reported from Figure 6.5 pertain only to the difference between SEP districts and non-magnet districts, not the difference between SEP districts and magnet districts. The one near-exception is the passing rate of white students, which is higher in SEP districts than in magnet districts ($p=0.067$). However, it is also the case that there is no significant increase in racial enrollment gaps in magnet districts relative to non-magnet districts. Thus, to the extent that white parents leverage student differentiation to collectively advocate to maintain their advantage, they may command greater legitimacy and effectiveness when they can point to district-sanctioned markers of their children's above-average abilities and needs.

Discussion

This chapter asked whether AP courses are distributed differently between schools in SEP districts than in non-SEP districts and, if so, whether this difference might reflect either greater efficiency and/or a process of Effectively Maintained Inequality. Overall, I find some evidence of different distributional choices in the provision of AP courses, in terms of both evenness and exposure, such that AP resources appear to be distributed more narrowly (E_r is lower) in SEP districts. Findings suggest that the average school in SEP districts offers fewer AP/IB courses per student relative to the best resources school in the district. This likely reflects a pattern whereby SEP districts are more likely operate at least one school with *exceptional* resources. This finding is largely consistent across racial/ethnic groups. However, despite similarities in (un)evenness, only Black students are exposed to fewer total AP courses in SEP districts than their same-race peers in non-SEP districts. Both white and Latinx students in SEP districts attend schools where, on average, they are exposed to the same number of AP courses as their same-race peers in non-SEP districts. By contrast, Black students attend schools with 1-2 fewer AP courses than their same race peers in magnet or non-magnet districts. This produces a white-Black gap of 2 AP/IB courses in SEP districts, which is significantly larger than the average 1-course gap in non-SEP districts.

Despite being distributed more narrowly at the school-level, these resources appear to reach a similar proportion of beneficiaries throughout the district. What is more, white and Latinx AP students are more likely to pass at least one AP exam in SEP districts than non-magnet districts. Together, these findings suggest that the distribution of AP resources may in fact be more efficient in SEP districts than non-magnet districts. However, this efficiency may come at the cost of increased racial inequalities. This is evidenced by larger advantages for white students over their URM peers in SEP districts. Although white students are more likely to enroll in AP courses than their Black or Latinx peers across district types, these gaps are largest in SEP districts. Additionally, the gap in white-Black exam passing rates is larger in SEP districts than in non-magnet districts. This is not because Black students perform worse in SEP districts than non-SEP districts, but rather because white students (and their parents) are able to more effectively maintain their advantage.

It should be noted that available data limit these analyses in important ways. The CRDC treats all students as equivalent who take *at least one* AP course or who pass *at least one* AP exam, regardless of how many courses they take or how many exams they pass. This may have implications for my findings regarding overall participation in the AP program, particularly for Black students. Black students in SEP districts attend schools where they are exposed to fewer AP/IB courses, both relative to their same-race peers in non-SEP districts and relative to their white peers in SEP districts. As a result, my estimate of AP participation gaps in SEP districts may be conservative relative to the inequalities in *intensity* of AP participation.

On the other hand, findings for exam passing are less likely to be conservative. I find that a greater proportion of white and Latinx students who take at least one AP course pass at least one AP exam in SEP districts than non-magnet districts. If more students in SEP districts are passing *one* exam because more of them are attempting *several*, then this may not reflect greater efficiency at all. While I cannot completely rule out this possibility, it does not seem particularly likely, given that exposure to AP courses is the same or lower on average in SEP districts compared to non-magnet districts. Furthermore, because of the uneven distribution of AP courses in these districts, the students most likely to participate heavily in the AP program are those who actually attend SEP schools. Because these students are also the most likely to have passed at least one AP exam in the absence of SEP schools, it is unlikely they are driving up the average passing rates. Therefore, it is unlikely that the estimated advantage in exam passing rates in SEP districts is being driven by unobserved variation in intensity of participation.

A final consideration pertains to comparison groups. Of course, both the non-magnet and magnet comparison groups are heterogeneous collections of districts. For the magnet group in particular, it is important to acknowledge that this sample includes both partial-school programs and whole-school magnets. The districts with partial-school programs are important to consider for two reasons. First, these programs may be selective. My definition of SEP schools purposefully includes only schools that admit their *entire* student body using selective criteria. Beyond the 62 SEP districts included in this analysis, my search identified an additional 20 districts with selective high school *programs*. In many cases, these programs may operate as schools-within-a-school. The second reason this is important is that, whether magnet programs are selective or not, the breadth of AP courses available in these buildings may not truly be available to all the students considered by the CRDC to be a part of the school. As a result, AP exposure may be overestimated in these districts. Additional models (see Appendix Tables C1 and C2) do in fact predict higher levels of exposure in these districts, but they also predict higher overall enrollment, suggesting that *over*estimation of exposure is not likely to be an important concern. However, these models do indeed reveal important differences between districts with only partial magnets versus those with whole-school magnets. Nevertheless, major take-aways from the analyses presented above hold: (1) AP/IB courses are more narrowly distributed in SEP districts in terms of both evenness and exposure than in either (kind of) magnet districts or non-magnet districts, (2) the white-URM enrollment gaps are largest in SEP districts, and (3) the white-Black passing gap is larger in SEP districts than non-magnet districts, but statistically indistinguishable from magnet districts.

More research is needed to understand the district-level decision-making processes that shape the distribution of college preparatory resources throughout a district and whether or how the differentiation of students into different schools (or programs) fits into these strategies. This chapter provides an important starting point and offers important policy implications for non-magnet districts that may be considering adding differentiated school options as part of their traditional public school choice menu. If the primary goal is to deploy fewer resources towards college preparation without hampering average participation and success rates, then offering SEP high schools may be one possible way to do this. However, this option comes with the risk of exacerbating EMI and increasing opportunities for white students at the expense of minority students, particularly Black students. If districts were to choose this option, they should carefully consider policies to equalize access, both to SEP schools as a whole and to AP courses within SEP schools. On the other hand, if a district's goal is simply to improve participation and success in the AP program, then magnet schools appear to offer a structure for this that does not share the same repercussions for inequality. Overall, this analysis is an additional piece of evidence to suggest that policymakers should pay careful attention not only to *access* to or absolute *levels* of resources, but to the potential for policies to exacerbate inequalities.

VII. CONCULSION:
(INE)QUALITY OF EDUCATIONAL OUTCOMES IN SEP DISTRICTS

This study was motivated by the fact that, to the extent that selective assignment procedures shape the educational experiences of students in SEP schools, they also shape the educational experiences of the students shut out of them. Rather than isolated institutions then, it is important to consider SEP schools as part of a selectively differentiated school system—a SEP district. Accordingly, the intent of this dissertation was to evaluate district-level outcomes associated with high-scope (i.e., between school) selective differentiation, rather than the effect of SEP schools on enrolled (or excluded) students separately. Specifically, I evaluated the impact of SEP districts on (1) school segregation, (2) math achievement in primary grades, and (3) the allocation of Advanced Placement courses in high school. In addition to encompassing a range of different *types* of outcomes, these analyses also span across primary and secondary school levels. Together, findings from this study provide a detailed picture the overall experience of educational opportunity and achievement in SEP districts.

This broad set of analyses is made possible by combining existing national datasets with the original list of SEP districts and schools compiled here. I collected these data using an innovative approach that drew on web scraping tools in Python to track down relevant information from school district websites pertaining to selective school admissions criteria. This search produced a list of 369 schools in 90 districts across the K-12 spectrum. These 90 districts are responsible for educating a sizeable 14% of the total public school population in the United States and 22% of the URM population. Yet, this selectively differentiated district structure has been given no consideration in existing research. This is particularly noteworthy in light of significant concerns around racial inequity in these districts, concerns which are validated by these new data. Put simply, SEP schools disproportionately serve white and Asian students in districts that disproportionately serve black and Latinx students. Given these racial disproportionalities, SEP districts could potentially exert inordinate influence on racial inequalities in educational opportunities in the United States. This motivates not only the analysis of racial school segregation, but also analyses of the race-specific implications of SEP districts for achievement and college preparation.

Overall, evidence from this study suggests that SEP schools have mixed effects on district-level educational opportunities and outcomes. In this final chapter, I briefly recount and synthesize these findings and argue that the risk of negative externalities should be a primary consideration for any policymaker considering the implementation or improvement of SEP schools.

Shuffling Students: The First-Order Question of SEP Schools and Segregation

In Chapter 4, I ask whether patterns of school segregation, both between and within districts, differ as a function of the presence of SEP schools. The question of intra-district segregation is part of a well-established tradition of research that examines the role of school choice in exacerbating school segregation (e.g. Saporito & Sohoni 2006). Meanwhile, the question of inter-district segregation is motivated by the possibility that advantaged families may be more inclined to enroll their children in relatively under-resourced districts if they can rely on a non-random procedure (i.e., academically selective admissions) to give them an edge in securing a seat in their desired school. In contrast to this hypothesis, a preliminary descriptive analysis suggests that dissimilarity in racial composition across districts is *greater* in CBSAs that include at least one SEP district. However, the present study cannot determine the direction of causality in this relationship—that is, whether SEP districts contribute to an increase in segregation or whether SEP schools are, for instance, a response to white flight. Accordingly, although it is important to consider the potential effects of SEP schools on district-level choice, this chapter focused the bulk of its analysis on within-district school segregation.

Over the course of a K-12 career, most students experience multiple changes in their school context as they make transitions from elementary to middle to high school. These changes typically bring with them larger and more diverse groups of peers. In other words, between school segregation changes across grade levels. Overall, segregation is higher in elementary school than in middle school and higher in middle school than in high school. However, these trends differ somewhat for different indicators of segregation and between SEP and non-SEP districts. I evaluate segregation in terms of evenness (the Dissimilarity Index, D) and exposure (the Interaction Index ${}_xP_y$,), including both the exposure of white students to URM peers and the exposure of URM students to white peers.

Overall, differences in the level of segregation between SEP and non-SEP districts are fairly consistent in between structural transition years (i.e., within the primary/middle/high school levels). However, the *change* in the Interaction Index at structural transition years—that is, sixth and ninth grades—differs significantly between SEP and non-SEP districts. Whereas white student exposure to URM students (${}_wP_{urm}$) increases significantly *more* at transitions in SEP districts, URM exposure to white students (${}_{urm}P_w$) increases significantly *less*. This suggests that application procedures do indeed shape the school choices families (are able to) make. Importantly, given the potential social implications of contact theory (Allport 1954), the relative increase in white students' exposure to URM peers at school transitions is among the more positive SEP district outcomes identified throughout this study. Unfortunately, the increase in ${}_wP_{urm}$ in ninth grade is immediately reversed in tenth. Moreover, together with the trends in the other two segregation indices, (D and ${}_{urm}P_w$), I argue that URM students may be particularly likely to transfer (or drop) out of predominantly white high schools in SEP districts. Contrary to suggesting positive social outcomes, this may imply that URM students face additional barriers to receiving full support in SEP districts, and perhaps especially in SEP schools themselves.

Although this is an important finding in terms of the lived experience of support and upheaval for high school students in SEP districts, it is also important to note that, overall, there are relatively few significant differences in the level of racial segregation between SEP and non-SEP districts by grade-level. These null findings actually provide a strong foundation for the analyses that followed—which evaluate the association between SEP schools and district-level academic outcomes—by suggesting that it is not racialized sorting, but performance-based sorting that drives differences between SEP and non-SEP districts.

What Students Learn: Quality and Inequality of Academic Outcomes

Chapters 5 and 6 evaluate academic outcomes in SEP districts relative to non-SEP districts. Each chapter assesses both the quality and inequality of outcomes and offers reasonable grounds for causal interpretations. Chapter 5 uses a grade-level Difference-in-Differences design to estimate changes in math achievement from third to eighth grade and compares the treatment effect in SEP districts to that in *future* SEP districts as a falsification test to reject spurious causation. The same grade-level longitudinal design is not possible at the high school level. However, analyses in Chapter 6 control for eighth grade achievement outcomes and compare high school SEP districts to their most similar peer districts—those with traditional magnet schools. Using Seemingly Unrelated Regression, Chapter 6 examines whether the allocation and uptake of AP coursework differs by district type. Findings from these analyses suggest mixed implications for SEP districts' impacts on quality and inequality of academic outcomes.

Quality of Achievement in SEP Districts

By *quality* of achievement, I refer to the overall (average) performance a district is able to elicit from its students. Across Chapters 5 and 6, I measure this as (1) average math achievement growth from third to eighth grade, (2) participation in AP courses, and (3) AP exam passing rates, each evaluated separately by race. Whereas middle school math achievement outcomes present a decidedly negative picture of educational quality in SEP districts, AP outcomes suggest possible increases in college preparation, at least for white and Latinx students.

In terms of average math achievement, I find that, on average, students in SEP districts demonstrate significantly slower math achievement growth from third to eighth grade, with effect sizes ranging from -9% to -11% for different racial/ethnic groups. This is a substantively large effect, which equates to approximately half a grade-level of learning. Asian students' math achievement is the exception to this rule. There is no significant difference in math achievement for Asian students between SEP and non-SEP districts. One possible explanation for this is that Asian students are exceptionally over-represented in SEP schools, such that the benefits accrued to students who actually attend SEP schools is sufficient to outweigh the penalties to students in non-SEP schools. However, white students are also over-represented in these schools and still experience significantly slower math achievement growth on average than white students in non-SEP districts. Thus, although these analyses do not foreclose the possibility that students benefit from enrollment in SEP schools[77], they do strongly suggest that average districtwide math achievement growth is slower in districts that offer SEP schools between fourth and eighth grade.

By contrast, Chapter 6 shows that more white students enroll in AP courses in SEP districts than non-SEP districts and that white and Latinx students who *enroll* in at least one AP course are more likely to pass at least one exam. This would seem to suggest that SEP districts *improve* the quality of opportunity of achievement for some students, albeit not for all. It is worth keeping in mind, however, that these analyses control for eighth grade achievement, which is itself shaped—and in fact *lowered*—by SEP schools, according to Chapter 5. In other words, the quality of AP participation for white and Latinx students is higher in SEP districts *net* of lower eighth grade achievement. This could mean that SEP districts have a neutral or even negative cumulative effect on AP participation if they begin selectively differentiating sufficiently early enough to significantly hamper math achievement by eighth grade. Moreover, it is important to consider possible limitations in the analysis of AP outcomes, including the lack of information on the *number* of AP courses and exams students take and pass, which may mask important differences in overall college preparation.

Across these analyses, I find evidence of positive (AP enrollment for white students, AP passing for white and Latinx students), neutral (AP performance for black students, math achievement for Asian students[78]), and negative (math achievement for white, Latinx, and black

[77] Other research, however, does cast doubt in this regard, by suggesting that marginally admitted SEP school students do not demonstrate significant improvement in test scores (Abdulkadiroğlu et al. 2014; Allensworth et al. 2016; Dobbie and Fryer Jr. 2011).

[78] And, not shown, null effects for reading achievement.

students) effects of SEP schools on districtwide achievement outcomes. Given these wide-ranging effects, and particularly given their variation by race, it is important to consider whether these average effects mask important inequalities.

Inequality of Achievement in SEP Districts

In the foregoing analyses, I examine several indicators of educational inequality. These include (1) standard deviation of achievement, (2) racial achievement gaps, (3) (un)evenness in AP course distribution between schools, and (4) racial gaps AP course access, enrollment, and exam passing. In other words, I consider both overall distributional inequalities (i.e., 1 and 3) as well as racial/ethnic gaps (2 and 4). Importantly, whereas race gaps in access to college preparation are necessarily a sign of discriminatory hindrance to student potential, this is not *necessarily* true of overall distributional inequalities, which could arise simply by lifting the ceiling of opportunity for the most gifted students. It is therefore important to consider both kinds of inequality.

I find mixed evidence of overall distributional inequalities. DiD analyses produce at best marginal evidence of an increase in the districtwide standard deviation of math achievement between third and eighth grade. However, it is important to bear in mind that SEDA's estimates of standard deviation of achievement may be relatively insensitive to extremely high or low performance, due to their reliance on coarse proficiency categories. In terms of resource distribution, I find consistent evidence that the average student's school (across racial groups) has a significantly smaller relative share of per-pupil AP resources in SEP districts than in non-SEP districts, controlling for average eighth grade achievement. Although this implies greater unevenness in school-level resources in SEP district, it likely reflects the presence of one or more schools with an *exceptionally* extensive list of AP course offerings, and does not necessarily imply that the average student in SEP districts has access to fewer total AP courses than the average student in non-SEP districts. To put these districtwide distributional inequalities in further context, I also evaluated racial disparities specifically.

At both the middle school and high school level, there is evidence that racial achievement gaps are exacerbated in SEP districts relative to non-SEP districts, although the minority group most affected differs. In middle school, evidence of a widening math achievement gap in SEP districts exists only between white and Latinx students. Importantly, both white and Latinx students experience slowed growth in middle school math achievement, but this penalty appears somewhat more severe for Latinx students on average (despite similar within-group effect sizes), producing a larger gap. Conversely, although *no* individual racial/ethnic group in SEP districts experiences a disadvantage in AP participation relative to *same-race* peers in non-SEP districts, I find evidence of significantly wider gaps (1) in AP enrollment between white and URM students and (2) in rates of exam passing between white and Black students. Black students in SEP districts also experience a disadvantage in AP course *exposure*, both relative to white students in SEP districts (i.e., increasing white-Black gaps) and relative to same-race peers in non-SEP districts.

The Big Picture

The analyses in this dissertation offer the first evaluation of districtwide outcomes associated with SEP schools. Based on a wide range of outcomes, I show that this model of student differentiation may have certain merits, but it is far from perfect. Specifically, there is strong evidence that selectively differentiating students between schools creates negative externalities—if not in overall quality, then at least in racial gaps in achievement and college preparation at the district level.

Local debates, highlighted in the introduction, suggest that racial equity in access should be the primary policy concern regarding SEP schools. Findings from this study confirm both that race shapes access to SEP schools and that SEP districts shape school segregation. For policymakers who hope to improve rather than remove SEP schools, correcting these inequalities is an important place to start. However, even solving this intractable problem would not preclude the possibility of negative externalities, such as inhibited growth or unequal access to college preparation in traditional public schools throughout SEP districts.

On the other hand, if the primary goal of SEP districts is to minimize the resources required to produce a given level of participation and success in advanced curricula, then offering SEP schools may be a reasonable strategy. However, it is important not to lose sight of the fact that this efficiency may come at the cost of racial equity. What is more, this sacrifice in equity does not appear to increase efficiency beyond what can be accomplished with traditional magnet schools.

Most importantly, this study establishes that the effects of these schools should not be sought solely among choice participants. This, in turn, should motivate continued research on the conditions under which SEP schools are founded, the place they occupy in district-level decision-making processes, and the mechanisms through which they exert their effects.

REFERENCES

Abdulkadiroğlu, Atila, Joshua D. Angrist, and Parag A. Pathak. 2014. "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools." *Econometrica* 82(1):137–96.

Ali, Syed and Margaret M. Chin. 2018. "What ' s Going On With New York ' s Elite Public High Schools ?" *The Atlantic*, June, 1–10.

Allensworth, Elaine M., Pail T. Moore, Lauren Sartain, and Marisa de la Torre. 2016. "The Educational Benefits of Attending Higher Performing Schools: Evidence From Chicago High Schools." *Educational Evaluation and Policy Analysis* 1–23.

Allport, Gordon W. 1954. *The Nature of Prejudice*. Reading, MA: Addison Wesley.

Angrist, Joshua D., Parag A. Pathak, and Román Andrés Zárate. 2019. *Choice and Consequence: Assessing Mismatch at Chicago Exam Schools*. 26137. Cambridge, MA.

Angrist, Joshua D. and Miikka Rokkanen. 2015. "Wanna Get Away? Regression Discontinuity Estimation of Exam School Effects Away From the Cutoff." *Journal of the American Statistical Association* 110(512):1331–44.

Anon. n.d. "Statewide Guarantee." *University of California*. Retrieved March 6, 2020 (https://admission.universityofcalifornia.edu/admission-requirements/freshman-requirements/california-residents/statewide-guarantee/).

Archbald, Doug. 1996. "SES and Demographic Predictors of Magnet School Enrollment." *Journal of Research and Development in Education* 29(3):152–62.

Arsen, David and Yongmei Ni. 2012. "The Effects of Charter School Competition on School District Resource Allocation." *Educational Administration Quarterly* 48(1):3–38.

Attewell, Paul. 2001. "The Winner-Take-All High School : Organizational Adaptations to Educational Stratification." *Sociology of Education* 74(4):267–95.

Ayalon, Hanna and Adam Gamoran. 2000. "Stratification in Academic Secondary Programs and Educational Inequality: Comparisons of Israel and the United States." *Comparative Education Review* 44(1):54–80.

Booker, Kevin, Scott M. Gilpatric, Timothy Gronberg, and Dennis Jansen. 2008. "The Effect of Charter Schools on Traditional Public School Students in Texas: Are Children Who Stay behind Left Behind?" *Journal of Urban Economics* 64(1):123–45.

Braddock II, Jomills Henry and Amaryllis Del Carmen Gonzalez. 2010. "Social Isolation and Social Cohesion : The Effects of K – 12 Neighborhood and School Segregation on Intergroup Orientations." *Teachers College Record* 112(6):1631–53.

Bryk, Anthony S. and Stephan W. Raudenbush. 1988. "Toward a More Appropriate Conceptualization of Research on School Effects: A Three-Level Hierarchical Linear Model." *American Journal of Education* 97(1):65–108.

Burkam, David T., Douglas D. Ready, Valerie E. Lee, and Laura F. LoGerfo. 2004. "Social-Class Differences in Summer Learning between Kindergarten and First Grade: Model Specification and Estimation." *Sociology of Education* 77(1):1–31.

Clark, Damon. 2007. "Selective Schools and Academic Achievement." *Institute for the Study of Labor (IZA), Bonn* IZA Discus.

Clotfelter, Charles. 2004. *After Brown: The Rise and Retreat of School Desegregation*. Princeton, NJ: Princeton University Press.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederick D. Weinfeld, and Robert L. York. 1966. *Equality of Educational*

Opportunity. Washington, DC.
Cook, Jason B. 2018. “The Effect of Charter Competition on Unionized District Revenues and Resource Allocation.” *Journal of Public Economics* 158:48–62.
Cooper, Harris, Barbara Nye, Kelly Charlton, James Lindsay, and Scott Greathouse. 1996. “The Effects of Summer Vacation on Achievement Test Scores: A Narrative and Meta-Analytic Review.” *Review of Educational Research* 66(3):227–68.
Corcoran, Sean Patrick and E. Christine Baker-Smith. 2018. “Pathways to an Elite Education: Application, Admission, and Matriculation to New York City’s Specialized High Schools.” *Education Finance and Policy* 13(2):256–79.
Corcoran, Sean Patrick, Jennifer L. Jennings, Sarah R. Cohodes, and Carolyn Sattin-Bajaj. 2018. *Leveling the Playing Field for High School Choice: Results from a Field Experiment of Informational Interventions*.
Current Population Survey. 2018. “Table 1. Educational Attainment of the Population 18 Years and Over, by Age, Sex, Race, and Hispanic Origin: 2018.”
Davis, James A. 1966. “The Campus as a Frog Pond : An Application of the Theory of Relative Deprivation to Career Decisions of College Men.” *American Journal of Sociology* 72(1):17–31.
Dobbie, Will and Roland G. Fryer Jr. 2011. *Exam High Schools and Academic Achievement: Evidence from New York City*. No. 17286.
Dobbie, Will and Roland G. Fryer. 2014. “The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools.” *American Economic Journal: Applied Economics* 6(3):58–75.
Ebbert, Stephanie. 2016. “Testing Program to Expand, Increase Latin School Access.” *Boston Globe*, April 27.
Education Commission of the States. n.d. “Advanced Placement: All High Schools/Districts Required to Offer AP.” *50-State Comparison*. Retrieved March 6, 2020a (http://ecs.force.com/mbdata/MBQuestRT?Rep=AP0116).
Education Commission of the States. n.d. “Advanced Placement: State Financial Support for AP Course Offerings/AP Success.” *50-State Comparison*. Retrieved March 6, 2020b (http://ecs.force.com/mbdata/MBQuestRT?Rep=AP0216).
Espenshade, Thomas J., Lauren E. Hale, and Chang Y. Chung. 2005. “The Frog Pond Revisited: High School Academic Context, Class Rank, and Elite College Admission.” *Sociology of Education* 78(4):269–93.
Fahle, Erin, Benjamin Shear, Demetra Kalogrides, Sean Reardon, Richard DiSalvo, and Andrew Ho. 2018. *Stanford Education Data Archive (Version 2.1): Technical Documentation*.
Felmlee, Diane H. and Lowell L. Hargens. 1988. “Estimation and Hypothesis Testing for Seemingly Unrelated Regressions: A Sociological Application.” *Social Science Research* 17(4):384–99.
Finn Jr, Chester E. 2012. “Young, Gifted and Neglected.” *New York Times*, September 18, The Opinion Pages.
Finn Jr, Chester E. and Jessica A. Hockett. 2012. *Exam Schools: Inside America’s Most Selective Public High Schools.* Princeton, NJ: Princeton University Press.
Formisano, Ronald P. 2004. *Boston against Busing: Race, Class, and Ethnicity in the 1960s and 1970s*. Univ of North Carolina Press.
Frank, Robert H. and Philip J. Cook. 1995. *The Winner-Take-All Society: Why the Few at the Top Get So Much More Than the Rest of Us*. New York, NY: Random House.

Fuller, Bruce and Richard Elmore, eds. 1996. *Who Chooses? Who Loses? Culture, Institutions and the Unequal Effects of School Choice*. New York, NY: Teachers College Press.
Gamoran, Adam. 2010. “Tracking and Inequality: New Directions for Research and Practice.” Pp. 213–28 in *The Routledge International Handbook of the Sociology of Education*, edited by M. W. Apple, S. J. Ball, and L. A. Gandin. London: Routledge.
Gay, Malcolm. 2019. “The Valedictorians Project: Part 3. The Race to Get into Boston’s Exam Schools.” *The Boston Globe*, January 17.
Grigg, Jeffrey. 2012. “School Enrollment Changes and Student Achievement Growth: A Case Study in Educational Disruption and Continuity.” *Sociology of Education* 85(4):388–404.
Gubbins, E. Jean, Carolyn M. Callahan, and Joseph S. Renzulli. 2014. “Contributions to the Impact of the Javits Act by the National Research Center on the Gifted and Talented.” *Journal of Advanced Academics* 25(4):422–44.
Hanushek, Eric A. and Ludger Wößmann. 2006. “Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries.” *The Economic Journal* 116(510):C63–76.
Horn, Daniel. 2009. “Age of Selection Counts: A Cross-Country Analysis of Educational Institutions.” *Educational Research and Evaluation* 15(4):343–66.
Hoxby, Caroline M. and Gretchen Weingarth. 2005. *Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects*. 7867.
Iatarola, Patrice, Dylan Conger, and Mark C. Long. 2011. “Determinants of High Schools’ Advanced Course Offerings.” *Educational Evaluation and Policy Analysis* 33(3):340–59.
Irizarry, Yasmiyn. 2017. “Commentary: Healing Won’t Begin until LASA Represents All of Austin ISD.” *Statesman*, February 24.
Irwin, Véronique. 2020. “Selective Enrollment Public Schools and District-Level Achievement Outcomes from 3rd to 8th Grade”. *Sociological Science*. 7: 100-127.
Jarwan, Fathi A. and John F. Feldhusen. 1993. *Residential Schools of Mathematics and Science for Academically Talented Youth: An Analysis of Admissions Programs*. Storrs, CT.
Jinnai, Yusuke. 2014. “Direct and Indirect Impact of Charter Schools’ Entry on Traditional Public Schools: New Evidence from North Carolina.” *Economics Letters* 124(3):452–56.
Kahlenberg, Richard. 2014. “Elite, Separate, Unequal: New York City’s Top Public Schools Need Diversity.” *New York Times*, June 23, A21.
Kao, Grace and Jennifer S. .. Thompson. 2003. “Racial and Ethnic Stratification in Educational Achievement and Attainment.” *Annual Review of Sociology* 29:417–42.
Kerckhoff, Alan C. 2001. “Education and Social Stratification Processes in Comparative Perspective.” *Sociology of Education* 74(Extra Issue: Current of Thought: Sociology of Education at the Dawn of the 21st Century):3–18.
Klopfenstein, Kristin. 2004. “Advanced Placement: Do Minorities Have Equal Opportunity?” *Economics of Education Review* 23(2):115–31.
Klopfenstein, Kristin and M. Kathleen Thomas. 2009. “The Link between Advanced Placement Experience and Early College Success.” *Southern Economic Journal* 75(3):873–91.
Klugman, Joshua. 2013. “The Advanced Placement Arms Race and the Reproduction of Educational Inequality.” *Teachers College Record* 115(5):1–34.
Kolluri, Suneal. 2018. “Advanced Placement: The Dual Challenge of Equal Access and Effectiveness.” *Review of Educational Research* 88(5):671–711.
Lauen, Douglas Lee. 2007. “Contextual Explanations of School Choice.” *Sociology of Education* 80(July):179–209.

Lewis, Amanda E. and John B. Diamond. 2015. *Despite the Best Intentions: How Racial Inequality Thrives in Good Schools*. Oxford University Press.
Lucas, Samuel R. 2001. "Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects." *American Journal of Sociology* 10(6):1642–90.
Lucas, Samuel R. and Mark Berends. 2002. "Sociodemographic Diversity, Correlated Achievement, and De Facto Tracking." *Sociology of Education* 75(4):328–48.
Lucas, Samuel R. and Mark Berends. 2007. "Race and Track Location in U.S. Public Schools." *Research in Social Stratification and Mobility* 25(3):169–87.
Lucas, Samuel R. and Adam Gamoran. 2004. "Tracking and the Achievement Gap." Pp. 171–98 in *Bridgin the Achievement Gap*, edited by J. E. Chubb and T. Loveless. Washington, DC: Brookings Institution Press.
Marsh, Herbert W. 1991. "The Failure of Academically Selective High Schools To Deliver Academic Benefits Commensurate with Their Students' Ability Levels." *American Educational Research Journal* 28(2):445–80.
Marsh, Herbert W. and Kit Tai Hau. 2003. "Big-Fish-Little-Pond Effect on Academic Self-Concept: A Cross-Cultural (26-Country) Test of the Negative Effects of Academically Selective Schools." *American Psychologist* 58(5):364–76.
Massey, Douglas and Nancy Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces* 67(2):281–314.
Matthews, Jay. 2017. "America's Most Challenging High School." *The Washington Post*, March 5.
Mickelson, Roslyn Arlin. 2001. "Subverting Swann: First- and Second-Generation Segregation in the Charlotte-Mecklenburg Schools." *American Educational Research Journal* 38(2):215–52.
Mickelson, Roslyn Arlin and Mokubung Nkomo. 2012. "Integrated Schooling, Life Course Outcomes, and Social Cohesion in Multiethnic Democratic Societies." *Review of Research in Education* 36(March):197–238.
Murnane, Richard. 1975. *The Impact of School REsources on Learning of Inner City Children*. Cambridge, MA: Ballinger.
National Association for Gifted Children & the Council of State Directors of Programs for the Gifted. 2015. *State of the States in Gifted Education: Policy and Practice Data 2014–2015*. Washington, DC.
National Center for Education Statistics. 2018. "Table 206.20. Percentage Distribution of Students Ages 5 through 17 Attending Kindergarten through 12th Grade, by School Type or Participation in Homeschooling and Selected Child, Parent, and Household Characteristics: Selected Years, 1999 through 2016." *Digest of Education Statistics*. Retrieved (https://nces.ed.gov/programs/digest/d17/tables/dt17_206.20.asp).
National Commission on Excellence in Education. 1983. *A Nation at Risk: The Imperative for Educational Reform*. Washington, DC.
Ni, Yongmei. 2009. "The Impact of Charter Schools on the Efficiency of Traditional Public Schools: Evidence from Michigan." *Economics of Education Review* 28(5):571–84.
Nye, Barbara, Spyros Konstantopoulos, and Larry V. Hedges. 2004. "How Large Are Teacher Eff." *Educational Evaluation and Policy Analysis* 26(3):237–57.
Orfield, Gary and Jennifer B. Ayscue, eds. 2018. *Discrimination in Elite Public Schools: Investigating Buffalo*. New York, NY: Teachers College Press.
Owens, A., S. F. Reardon, and C. Jencks. 2016. "Income Segregation Between Schools and

School Districts." *American Educational Research Journal* 53(4):1159–97.
Owens, Ann. 2017. "Racial Residential Segregation of School-Age Children and Adults: The Role of Schooling as a Segregating Force." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 3(2):63–80.
Owens, Ann. 2018. "Income Segregation between School Districts and Inequality in Students' Achievement." *Sociology of Education* 91(1):1–27.
Parcel, Toby L. and Mikaela J. Dufur. 2001. "Capital at Home and at School: Effects on Student Achievement." *Social Forces* 79(3):881–911.
Pettigrew, Thomas F. and Linda R. Tropp. 2006. "Interpersonal Relations and Group Processes: A Meta-Analytic Test of Intergroup Contact Theory." *Journal of Personality and Social Psychology* 90(5):751–83.
Pfeffer, Fabian T. 2008. "Persistent Inequality in Educational Attainment and Its Institutional Context." *European Sociological Review* 24(5):543–65.
Pfeffer, Fabian T. 2012. *Equality and Quality in Education Population Studies Center Research Report 774*.
Quiroz, Pamela Anne and Vernon Lindsay. 2015. "Selective Enrollment, Race, and Shifting the Geography of Educational Opportunity: Where '"Diversity"' and Opportunity Compete with Tax Increment Financing." *Humanity & Society* 39(4):1–18.
Reardon, Sean. F., A. D. Ho, B. R. Shear, E. M. Fahle, D. Kalogrides, and R. DiSalvo. 2018. "Stanford Education Data Archive (Version 2.1)."
Reardon, Sean F., Elena Tej Grewal, Demetra Kalogridges, and Erica Greenberg. 2012. "Brown Fades: The End of Court-ORdered School Desegregation and the Resegregation of American Public Schools." *Journal of Policy Analysis and Management* 31(4):876–904.
Reardon, Sean F., Demetra Kalogrides, and Andrew Ho. 2017. "Linking U. S. School District Test Score Distributions to a Common Scale." *CEPA Working Paper No. 16-09* (16):1–50.
Reardon, Sean F. and Ann Owens. 2014. "60 Years After Brown : Trends and Consequences of School Segregation." *Annual Review of Sociology* 40(199:218).
Reardon, Sean F. and John Yun. 2001. "Suburban Racial Change and Suburban School Segregation, 1987–1995." *Sociology of Education2* 74:79–101.
Reardon, Sean F., John T. Yun, and Tamela Mcnulty Eitle. 2000. "The Changing Structure of School Segregation : Measurement and Evidence of Multiracial Metropolitan-Area School Segregation 1989-1995." *Demography* 37(3):351–64.
Renzulli, Joseph S., Carolyn M. Callahan, and E. Jean Gubbins. 2014. "Laying the Base for the Future: One Cornerstone of the Javits Act." *Journal of Advanced Academics* 25(4):338–48.
Rey, Jay. 2017. "African-American Enrollment Dips at City Honors, despite Efforts to Correct Disparity." *Buffalo News*, September 26.
Rumberger, Russell W. and Gregory J. Palardy. 2005. "Does Segregation Still Matter? The Impact of Student Composition on Academic Achievement in High School." *Teachers College Record* 107(9):1999–2045.
Sacerdote, Bruce. 2011. "Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" Pp. 249–77 in *Handbook of the Economics of Education*. Vol. 3. Elsevier B.V.
Saporito, Salvatore and Deenesh Sohoni. 2006. "Coloring Outside the Lines: Racial Segregation in Public Schools and Their Attendance Boundaries." *Sociology of Education* 79(2):81–105.
Sattin-Bajaj, Carolyn and Allison Roda. 2018. *Opportunity Hoarding in School Choice Contexts: The Role of Policy Design in Promoting Middle-Class Parents' Exclusionary Behaviors*.

Schneider, Jack. 2009. “Privilege, Equity, and the Advanced Placement Program: Tug of War.” *Journal of Curriculum Studies* 41(6):813–31.
Schonfeld, Zach. 2014. “Newsweek ’ s High School Rankings for 2014 : Two Lists Are Better than One.” *Newsweek*, September 8.
Scott Krei, Melinda. 1998. “Intensifying the Barriers: The Problem of Inequitable Teacher Allocation in Low-Income Schools.” *Urban Education* 31(1):71–84.
Shapiro, Eliza. 2019. “Only 7 Black Students Got Into Stuyvesant, NY’s Most Selective High School, Out of 895 Spots.” *New York Times*, March 18.
Sørensen, Aage Bøttger. 1970. “Organizational Differentiation Of Students And Educational Opportunity.” *Sociology of Education* 43(4):355–76.
Stearns, Elizabeth. 2010. “Long-Term Correlates of High School Racial Composition : Perpetuation Theory Reexamined.” *Teachers College Record* 112(6):1654–78.
Stevens, Mitchell L. 2007. “Creating a Class: College Admissions and the Education of Elites.” Cambridge, MA: Harvard University Press.
The Editorial Board. 2018. “It’s Time to Integrate New York’s Best Schools.” *New York Times*, June 24.
Tucker, Jill. 2018. “Lowell High School’s Competitive Selection Process May Violate State Law.” *San Francisco Chronicle*, September 22.
U.S. Department of Education. 2019. “Funding Status: Jacob K. Javits Gifted and Talented Students Education Program.” Retrieved February 20, 2020 (https://www2.ed.gov/programs/javits/funding.html).
Wells, Amy Stuart. 1993. *Time To Choose: America at the Crossroads of School Choice Policy (New Frontiers of Education)*. New Jersey: Putnam Publishing Group.
van de Werfhorst, Herman G. and Jonathan J. B. Mijs. 2010. “Achievement Inequality and the Institutional Structure of Educational Systems: A Comparative Perspective.” *Annual Review of Sociology* 36:407–28.
West, Anne and Rita Nikolai. 2013. “Welfare Regimes and Education Regimes: Equality of Opportunity and Expenditure in the EU (and US).” *Journal of Social Policy* 42(3):469–93.
Yin, Alice. 2017. “Asian Test-Prep Centers Offer Parents Exactly What They Want: ‘Results.’” *New York Times Magazine*, October 25.
Yu, Corrine M. and William L. Taylor. 1997. *Difficult Choices: Do Magnet Schools Serve Children in Need? Report of the Citizens’ Commission on Civil Rights*. Washington, DC.
Zimmer, Ron and Richard Buddin. 2009. “Is Charter School Competition in California Improving the Performance of Traditional Public Schools?” *Public Administration Review* 69(5):831–45.

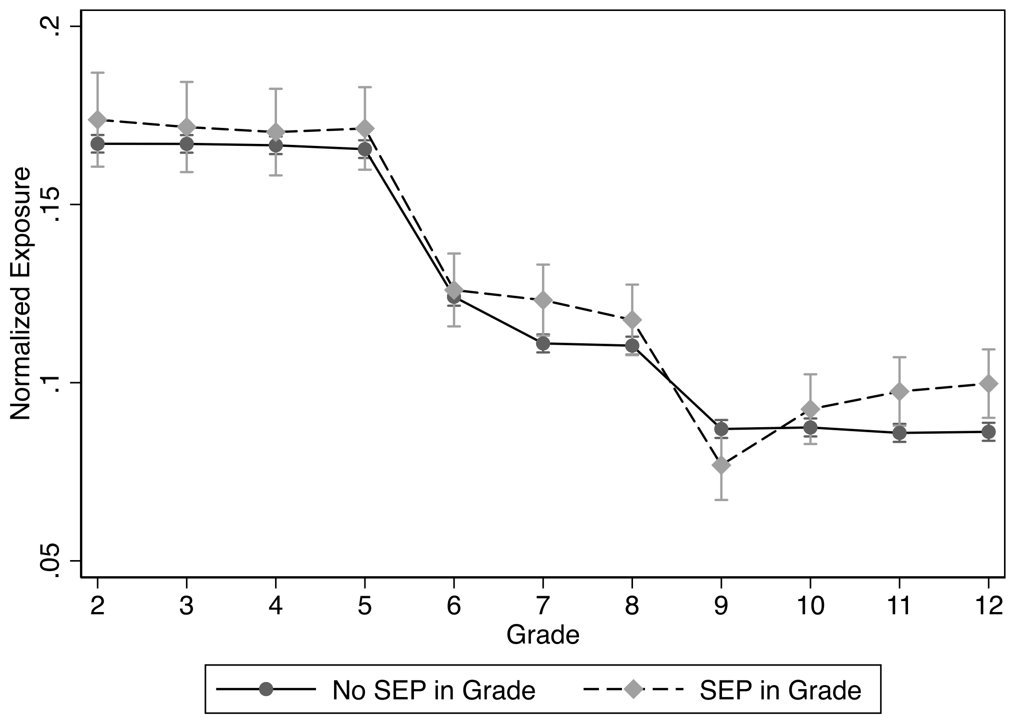
APPENDIX A: Supplement to Chapter IV

Normalized Isolation of White and URM Students

Like most measures of exposure, ${}_{w}P_{urm}$ is sensitive to the relative population share of each group. Below I present a replication of Figure 4.2 Panel A using the Normalized Interaction Index, which accounts for the overall proportion of white students (P), as shown in Equation A1. Findings using this measure neatly parallel findings based on the Index of Dissimilarity, which captures a different dimension of segregation, but also accounts for the overall population composition with P.

$$N = [({}_{w}P_{urm} - P)/(1 - P)] \tag{A1}$$

Figure A1.



Relative Representation from EDGE

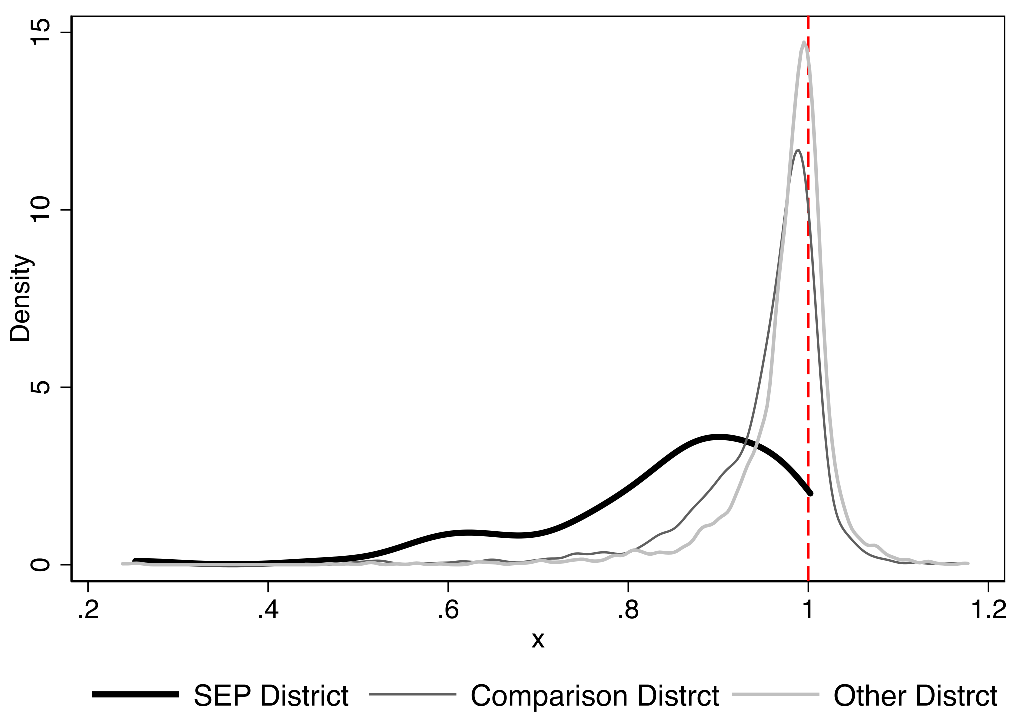
I generate an additional measure of racial disproportionality in schools, which I refer to as a "relative representation ratio". Unlike traditional measures of segregation, this index does not measure dispersion or dissimilarity between subunits within a larger entity, but is rather a comparison of the composition of two overlapping entities—(1) the school district and (2) the residential community within the geographic boundaries of the district. The Education Demographic and Geographic Estimates (EDGE) from the Census provide population estimates based on the geographic boundaries of public school districts, including (1) all residents of the district, (2) all school-aged residents of the district, and (3) all residents (children) who are actually enrolled in the district. I define the relative representation of a given demographic group as the ratio of district enrollments to school-aged district residents, as shown in Equation 4. Again, because of the advantage white students experience in school choice—and because the over/under-representation of white and URM students are mechanically inversely related—I focus on the choice behaviors of white students.

$$R = \left[\frac{P_{enr}}{P_{res}}\right] \qquad (A2)$$

Where P_{enr} and P_{res} are the proportion of white students enrolled and residing in the school district, respectively.

So defined, a relative representation ratio of 1 means that white students are proportionately represented in public schools, whereas a relative representation ratio of 1.2 or 0.8 would mean that white students are over- or under-represented by 20%, respectively. I estimate a proportional imbalance, rather than a difference, because differencing would commit the error of treating a 2 percentage-point difference in enrollment composition in SEP versus non-SEP districts as the same "effect size", regardless of whether the demographic composition was 40% or 80% white. The relative representation ratio treats the former as a larger change (5% compared to 2.5%). Whereas the traditional segregation measures defined above range from 0 to 1, the relative representation ratio can technically range from 0 to infinity. However, in practice white students are underrepresented in the vast majority of public school districts (R<1).

Figure A2. District-Level Relative Representation of White Students



APPENDIX B: Supplement to Chapter V

Table B1. Replication of Tables 5.3 and 5.4 using Grade Cohort Standardized (GCS) Scale

SEP Scope Definition	binary	binary	grade count	enrollment %
District controls	x	x	x	x
Future-SEP		x	x	x
Panel 1: Pooled (N=1570)	(B1)	(B2)	(B3)	(B4)
SEP District DiD	-0.428***	-0.430***	-0.120***	-0.077***
	(0.096)	(0.096)	(0.031)	(0.018)
SEP District (3rd grade difference)	0.174*	0.169*	0.053*	0.036*
	(0.069)	(0.069)	(0.023)	(0.014)
8th Grade (ref, 3rd Grade)	4.622***	4.624***	4.621***	4.619***
	(0.027)	(0.027)	(0.027)	(0.027)
Future-SEP DiD		-0.069	-0.066	-0.066
		(0.080)	(0.080)	(0.080)
Constant	2.258***	2.259***	-0.283***	-0.285***
	(0.087)	(0.048)	(0.028)	(0.029)
R-squared	0.956	0.782	0.788	0.788
Panel 2: White (N=1488)	(B5)	(B6)	(B7)	(B8)
SEP District DiD	-0.345**	-0.343**	-0.106*	-0.062*
	(0.120)	(0.120)	(0.045)	(0.024)
Panel 3: Asian (N=1078)	(B9)	(B10)	(B11)	(B12)
SEP District DiD	-0.187	-0.186	-0.038	-0.018
	(0.146)	(0.147)	(0.046)	(0.027)
Panel 4: Black (N=1250)	(B13)	(B14)	(B15)	(B16)
SEP District DiD	-0.351**	-0.355**	-0.090**	-0.045**
	(0.112)	(0.112)	(0.032)	(0.017)
Panel 5: Latinx (N=1478)	(B17)	(B18)	(B19)	(B20)
SEP District DiD	-0.391***	-0.392***	-0.114***	-0.065***
	(0.080)	(0.080)	(0.027)	(0.016)

Huber-White heteroscedasticity-robust clustered standard errors in parentheses
+ p<.1, * p<.05, ** p<.01, *** p<.001

Exploratory Math Achievement Mediation

Table B2. Testing Potential Mediators

	Pooled	White	Asian	Black	Latinx
District Controls	x	x	x	x	x
Future SEP	x	x	x	x	x
Panel 1: Replication of Results	(3)	(3-W)	(3-A)	(3-B)	(3-L)
SEP District DiD	-0.085**	-0.091**	-0.033	-0.087*	-0.110***
	(0.028)	(0.034)	(0.041)	(0.036)	(0.026)
Panel 2: 7th Grade Algebra	(C1)	(C2)	(C3)	(C4)	(C5)
SEP District DiD	-0.082**	-0.089**	-0.027	-0.085*	-0.106***
	(0.028)	(0.034)	(0.042)	(0.036)	(0.027)
District-wide Algebra Exposure [b]	-0.007	-0.007	0.051	-0.031	-0.048*
(effect in grade 3)	(0.017)	(0.019)	(0.039)	(0.023)	(0.020)
District-wide Algebra Exposure	0.027	0.047*	0.116**	-0.002	-0.010
(effect in grade 8)	(0.019)	(0.021)	(0.044)	(0.025)	(0.020)
Panel 3: Free-Lunch Exposure	(C6)	(C7)	(C8)	(C9)	(C10)
SEP District DiD	-0.074**	-0.086**	-0.018	-0.082*	-0.090***
	(0.028)	(0.033)	(0.043)	(0.038)	(0.026)
District-wide Algebra Exposure	0.006	-0.019	0.062	-0.017	-0.035+
(effect in grade 3)	(0.017)	(0.018)	(0.039)	(0.022)	(0.019)
District-wide Algebra Exposure	0.027	0.037+	0.119**	-0.005	-0.011
(effect in grade 8)	(0.018)	(0.020)	(0.044)	(0.024)	(0.020)
Exposure to FL peers [b]	-0.642***	-1.106***	-0.496	-1.042***	-0.966***
	(0.151)	(0.156)	(0.364)	(0.130)	(0.133)
Observations	1570	1488	1078	1250	1478

Huber-White heteroscedasticity-robust clustered standard errors in parentheses
+ p<.1, * p<.05, ** p<.01, *** p<.001
Notes:
[a] This is an indicator of distribution of and access to advanced coursework, which is likely to be a direct target of SEP school policy interventions. Math placement is the only advanced coursework for which data is available prior to high school from the CRDC. It is also relevant to focus on advanced coursework in math, because these models specifically test the association between SEP schools and district-level math achievement
[b] I also test exposure to Free Lunch eligible peers as an indicator of between-school segregation by SES, which research shows is associated with achievement, and is more consequential for outcomes than racial segregation (Owens, Reardon, and Jencks 2016; Rumberger and Palardy 2005).

APPENDIX C: Supplement to Chapter VI

Figure C1. Noteworthiness of Resources in SEP Schools Relative to the District

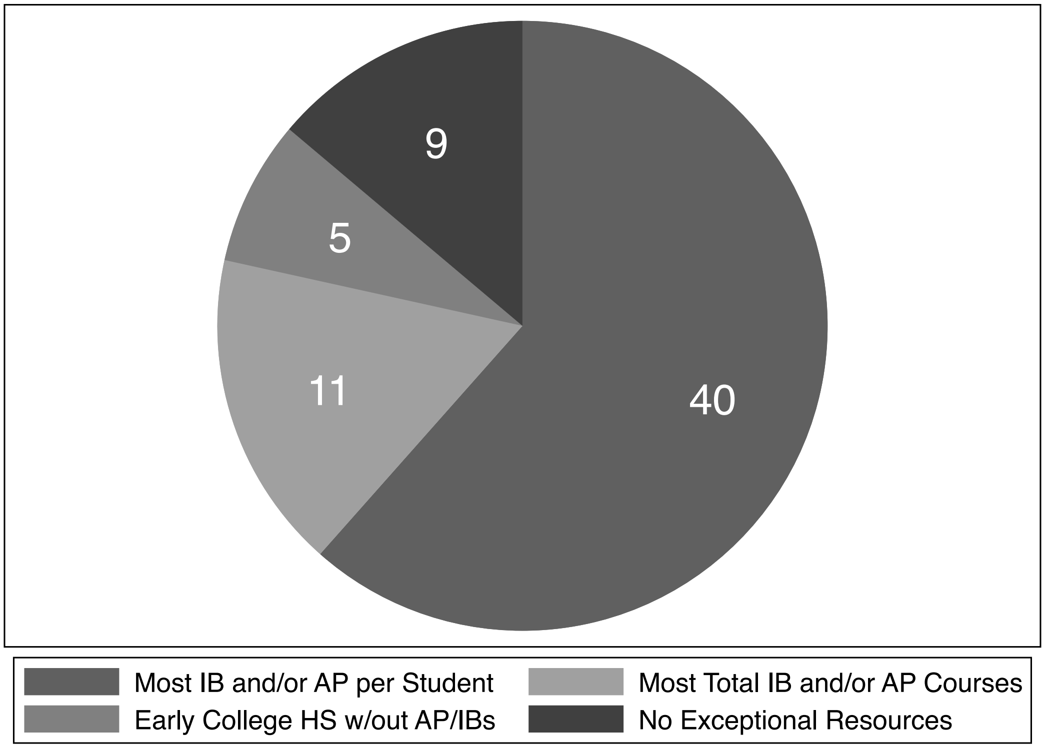


Table C1. SUR of AP Course Allocation for District Average Student, by Race

	AP Per Student Evenness			AP Course Exposure		
	(1)	(2)	(3)	(4)	(5)	(6)
Race-Specific Achievement [a]		x	x		x	x
N	326	325	325	326	325	325
White						
SEP District	-0.176***	-0.177***	-0.181***	-0.220	-0.281	-0.369
	(0.038)	(0.039)	(0.038)	(0.727)	(0.702)	(0.701)
(Whole-)Magnet District	-0.059*	-0.055+	-0.091**	0.307	0.368	-0.290
	(0.030)	(0.030)	(0.035)	(0.564)	(0.541)	(0.626)
Partial-Magnet District			0.006			1.488+
			(0.042)			(0.767)
Black						
SEP District	-0.188***	-0.191***	-0.194***	-1.031	-1.328*	-1.409*
	(0.035)	(0.035)	(0.035)	(0.696)	(0.678)	(0.674)
(Whole-)Magnet District	-0.058*	-0.056*	-0.079*	0.449	0.449	-0.147
	(0.027)	(0.028)	(0.032)	(0.542)	(0.527)	(0.609)
Partial-Magnet District			-0.017			1.463+
			(0.039)			(0.747)
Latinx						
SEP District	-0.188***	-0.192***	-0.196***	-0.670	-0.996	-1.084
	(0.037)	(0.037)	(0.037)	(0.712)	(0.693)	(0.690)
(Whole-)Magnet District	-0.049+	-0.046	-0.074*	0.646	0.684	0.039
	(0.029)	(0.029)	(0.033)	(0.554)	(0.539)	(0.623)
Partial-Magnet District			0.000			1.779*
			(0.041)			(0.762)

Table C1, cont'd

Wald Tests						
SEP(Black)=SEP(White)	0.524	0.433	0.479	0.042	0.012	0.012
SEP(Latinx)=SEP(White)	0.386	0.259	0.283	0.226	0.063	0.061
SEP(White)=Magnet(White)	0.001	0.001	0.023	0.450	0.335	0.912
SEP(Black)=Magnet(Black)	0.000	0.000	0.002	0.027	0.006	0.071
SEP(Latinx)=Magnet(Latinx)	0.000	0.000	0.001	0.054	0.012	0.118

Standard errors in parentheses; + $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

[a] See page 68 for full list of other controls included in all models

Table C2. SUR of District-Level AP “Efficiency”, by Race

	AP Enrollment			Exam Passing per Enrollment		
	(7)	(8)	(9)	(10)	(11)	(12)
Race-Specific Achievement [a]		x	x		x	x
N	326	325	325	320	319	319
White						
SEP District	0.030*	0.031*	0.030*	0.072**	0.081***	0.083***
	(0.014)	(0.014)	(0.014)	(0.023)	(0.022)	(0.022)
(Whole-)Magnet District	0.012	0.013	0.003	0.033+	0.040*	0.057**
	(0.011)	(0.011)	(0.012)	(0.017)	(0.017)	(0.019)
Partial-Magnet District			0.029+			0.009
			(0.015)			(0.024)
Black						
SEP District	0.011	0.007	0.006	0.017	0.016	0.018
	(0.009)	(0.008)	(0.008)	(0.019)	(0.019)	(0.019)
(Whole-)Magnet District	0.003	0.002	-0.005	-0.011	-0.011	0.004
	(0.007)	(0.006)	(0.007)	(0.015)	(0.015)	(0.017)
Partial-Magnet District			0.016+			-0.037+
			(0.009)			(0.021)
Latinx						
SEP District	0.009	0.002	0.001	0.048*	0.053*	0.055**
	(0.010)	(0.009)	(0.009)	(0.021)	(0.021)	(0.021)
(Whole-)Magnet District	0.006	0.003	-0.006	0.030+	0.032*	0.054**
	(0.008)	(0.007)	(0.008)	(0.016)	(0.016)	(0.018)
Partial-Magnet District			0.019+			-0.005
			(0.010)			(0.022)

Table C2, cont'd

Wald Tests						
SEP(Black)=SEP(White)	0.081	0.022	0.024	0.020	0.007	0.007
SEP(Latinx)=SEP(White)	0.048	0.005	0.005	0.287	0.208	0.220
SEP(White)=Magnet(White)	0.201	0.160	0.059	0.074	0.052	0.270
SEP(Black)=Magnet(Black)	0.313	0.552	0.180	0.125	0.134	0.482
SEP(Latinx)=Magnet(Latinx)	0.760	0.923	0.482	0.352	0.294	0.941

Standard errors in parentheses; + $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

[a] See page 68 for full list of other controls included in all models