The Long-Run Impacts of Reducing Racial Gaps in Special Education

Briana Ballis* and Katelyn Heath†

January 21, 2026

Abstract

Black students are about 1.5 times more likely to be receiving special education (SpEd) services relative to White students. While there is concern that this implies some Black students are inappropriately placed in SpEd, there is little evidence for whether this helps or harms Black students. Using administrative data from Texas, we find that policy capping Black over-representation in SpEd led to reductions in SpEd placement by early high school and small gains in high school completion and college attainment for Black students in special and general education. These results are driven by students with specific learning disabilities (SLD), males, and those in urban and midsize districts. Our findings from Texas suggest that policy aimed at reducing SpEd misclassification among Black students may serve to reduce gaps in later-life success across race.

Keywords: Special education, disability, disproportionality, racial gaps, educational attainment

JEL Codes: I24, I26, J14, J15

*Department of Economics, UC Merced. 5200 N. Lake Rd. Merced, CA 95343 Email: bballis@ucmerced.edu

†**Corresponding author.** Hobby School of Public Affairs, University of Houston. 4104 Martin Luther King Blvd. Houston, Texas 77004 Email: kmheath@uh.edu.

We would like to thank Maria Fitzpatrick, Michael Lovenheim, Doug Miller, Jordan Matsudaira, Rodney Andrews, Scott Carrell, Paco Martorell, Marianne Page, and Marianne Bitler for their advice and guidance. Briana Ballis acknowledges financial support from the Disability Research Consortium Fellowship from SSA via Mathematica Policy, the National Science Foundation, and the UC Davis Economics department. Katelyn Heath acknowledges financial support from the National Academy of Education and the National Academy of Education/Spencer Dissertation Fellowship Program, the Bronfenbrenner Center for Translational Research, Christopher Wildeman, Cornell University Graduate School, Cornell Policy Analysis and Management Department, and Cornell Economics Department Labor Grant in Economics. The conclusions of this research do not necessarily reflect the opinions or official position of the Texas Education Agency, the Texas Higher Education Coordinating Board, Texas Workforce Commission, or the State of Texas. This paper uses confidential data from the Texas Schools Project (TSP) through the University of Texas at Dallas. The data can be obtained by sending an inquiry to the director of the TSP, currently Trey Miller (tmiller@utdallas.edu). Please follow this link for additional information: `https://tsp.utdallas.edu/erc/initiating/`.

1 Introduction

Racial disparities during adulthood in the U.S. are well documented along several dimensions, including health (Leive & Christopher, 2022; Cutler & Vogl, 2012; Bound et al., 1995), educational attainment (Reardon, 2016; Rothstein & Wozny, 2013; Card & Rothstein, 2007; Cameron & Heckman, 2001; Jencks & Phillips, 1998), and employment (Chetty et al., 2019; Bayer & Charles, 2018). A growing literature demonstrates these disparities later in life can, in part, be linked to educational experiences during childhood. For instance, higher school spending and preschool programs have been shown to reduce racial gaps later in life (Rothstein & Schanzenbach, 2022; Heckman & Karapakula, 2019). An important, yet understudied, possible contributor to racial disparities in adulthood is special education (SpEd), one of the largest K-12 programs, which provides specialized services to students with disabilities.

The share of students in SpEd has more than doubled since 1975, with over 14 percent currently receiving services. Black students participate at even higher rates, and are about one and a half times more likely to receive SpEd services relative to White students (Gordon, 2017; Hosp & Reschly, 2003; Donovan & Cross, 2002; Oswald et al., 1999). However, there is no consensus in the literature on why this over-representation exists. Race may be correlated with other factors (such as socioeconomic status) that contribute to a greater underlying need for SpEd services.[1] Alternatively, racial biases in the SpEd referral or evaluation process may lead to misidentification.[2] While extensive research has documented the disproportionate representation of Black students in SpEd, there is little evidence on whether it helps or harms students. And, even less is known about its possible spillover effects on peers.

In this paper, we investigate the direct and spillover effects of policy that limited over-representation of Black and Hispanic students in SpEd in Texas. Although SpEd aims to benefit students through personalized instruction and accommodations (such as one-on-one or small group instruction, a classroom aide, or standardized testing modifications) there

[1] Interestingly, after conditioning on important confounds such as prior academic achievement and socioeconomic status, literature has shown that minority students are *less* likely to be receiving SpEd services relative to their observationally-equivalent White peers (Elder et al., 2021; P. Morgan, Farkas, Hillemeier, & Maczuga, 2017; Shifrer et al., 2011; Hibel et al., 2010). However, although these studies provide useful evidence about conditional representation, the endogeneity of achievement limits our ability to draw conclusions about whether observed disparities represent true over- or under-representation.

[2] As we will discuss in Section 2, for some disability types, determining whether a student qualifies for SpEd is a subjective process and teachers may interpret the same behavioral or academic challenges differently across race (Hosp & Reschly, 2003; Prieto & Zucker, 1981).

are also costs associated with participation. For example, SpEd can impose stigma, lower expectations from teachers and/or parents, reduce students' self-perceptions and aspirations, reduce students' exposure to the general curriculum, or result in negative spillovers from other SpEd students with relatively more challenging classroom behaviors (Shifrer, 2013; Harrison et al., 2013; Blackorby & Cameto, 2004). As will be discussed in more detail with a theoretical framework (Section 2.3), the impacts of limiting over-representation in SpEd are a-priori ambiguous. If students are over-represented for reasons that are unrelated to their actual need for services, such as racial biases, reducing SpEd access could improve outcomes. However, if SpEd is reduced among a population that genuinely requires services at higher rates, they are likely to be negatively impacted. Despite these important implications, to our knowledge, no prior study has been able to causally identify how reducing over-representation influences long-run trajectories of those directly affected and their peers.

Our research design utilizes a 2004 policy change that introduced district-level caps on the over-representation of Black and Hispanic students in SpEd, hereto referred to as disproportionality caps (consistent with the language in the original policy). Specifically, if the percent of Black (or Hispanic) students in SpEd was 1 p.p. higher than the percent of Black (or Hispanic) students in a district overall, the district was deemed out of compliance and had to reduce access to SpEd for Black (or Hispanic) students. In 2004, Black students were significantly over-represented in SpEd, with Black SpEd enrollment 4.2 p.p. higher than the percent of Black students in a district overall, on average. In contrast, Hispanic students were under-represented, with Hispanic SpEd enrollment 2.9 p.p. *lower* than the percent of Hispanic students in a district overall, on average.[3] Since the majority of districts were already in compliance with the Hispanic disproportionality cap, the policy did not have a strong impact on Hispanic SpEd students and we focus primarily on Black students moving forward.

Using administrative data from Texas and a dose-response difference-in-differences estimation strategy, we leverage variation across districts in pre-policy disproportionality rates and across cohorts in their exposure to the policy to identify effects on SpEd removal and long-run outcomes. The disproportionality policy also coincided with the implementation of

[3]It is interesting that Hispanic students are under-represented in SpEd at baseline. Similar to the possible reasons behind the over-representation of Black students in SpEd, this could be due to Hispanic students having a smaller underlying need for services, and/or a result of bias that leads to fewer Hispanic students being placed in SpEd. Ultimately, we are not able to investigate the effects of under-representation in this context, given that the policy was focused on reducing over-representation.

a district-level cap on SpEd enrollment at 8.5%. In previous work, we found that reducing access to SpEd through this cap harmed long-run outcomes for SpEd students (Ballis & Heath, 2021). We elaborate on this aspect of the policy in Sections 2.2 and 4, and provide an in-depth comparison of these two papers in Appendix B. Importantly, our results are robust whether we include or omit controls for the SpEd enrollment cap, allowing us to isolate the effects of the disproportionality caps.

We focus primarily on three groups of students. First, we focus on a group of *marginal* SpEd students, defined as those in SpEd as of 5th grade prior to policy implementation with Specific Learning Disabilities (SLD). Using pre-policy SpEd status avoids changes in the ability distribution of students in SpEd post-policy. Determinations as of 5th grade provides a stable sample, since SpEd enrollment rises rapidly in earlier grades but levels off by this point.[4] We refer to this group as marginal since SLD eligibility relies on relatively more subjective criteria, making them more likely to be removed from SpEd after the policy. Second, we focus on a group of *inframarginal* SpEd students, defined as non-SLD SpEd students. These students are more likely to remain in SpEd post-policy, but may still be impacted in the long-run through smaller SpEd caseloads, peer spillovers, changes in instructional attention, or racial bias. Third, we focus on 5th grade general education (GE) students who may have been directly affected by the policy by being less likely to be placed in SpEd in later grades, or indirectly affected through similar channels as those who remain in SpEd.

For our *marginal* SLD students, we find a 1.8 p.p. (2.2%) decrease in the likelihood of remaining in SpEd in 9th grade, but *increases* in high school graduation by 1.4 p.p. (2.2%) and college-going by 1.4 p.p. (4.6%). These estimates reflect effect sizes for students at the average district (that was 3.2 p.p. above the 1% disproportionality cap) who were exposed to the policy in every year between 5th and 9th grade (or 12th grade for the long-run outcomes). In contrast, our *inframarginal* non-SLD SpEd students did not experience significant changes in SpEd removal or longer-run educational attainment. Although the coefficients on the long-run outcomes are positive, the standard errors are large.

For Black GE students, we find small but significant declines in the likelihood of SpEd placement at 9th grade (0.32 p.p. or 6.9%), and positive effects on high school completion (0.51 p.p. or 0.7%) and college enrollment (1.7 p.p. or 3.1%). While not being placed

[4]Results are robust to defining SpEd status in 3rd, 4th, or 6th grade (Appendix Table A.2).

in SpEd in later grades contributed to improving outcomes, our results suggest that indirect mechanisms (such as broader improvements in school climate and possible changes in teaching practices) played a crucial role. The magnitude of the effect on SpEd placement among Black GE students is about one quarter of the size of the effect on SpEd students. And yet, the effect on college completion for GE students is slightly larger. Moreover, college enrollment improves significantly across all baseline levels of achievement for Black GE students, even among high achievers who were not likely to be placed in SpEd in later grades.

While our data prevent us from precisely determining mechanisms behind the positive impacts of SpEd removal for marginal Black SLD students, we find several pieces of evidence consistent with our results being driven by students who were misidentified for SpEd. First, impacts for SLD students are larger in districts that we predict are over-identifying Black SpEd students, that is, where Black students are more likely to be in SpEd even after conditioning on confounding characteristics (such as socioeconomic and early achievement measures). Second, impacts are concentrated in districts with characteristics that are linked to higher misclassification such as lower shares of Black students and fewer Black teachers (Stiefel et al., 2024; Hart & Linday, 2024; Elder et al., 2021; Fish, 2019; Gershenson et al., 2016; Dee, 2005). Finally, the policy removed students with relatively milder disabilities (i.e., spent more time in GE and less likely to receive testing exemptions) than those removed pre-policy. This is consistent with districts removing students who were misidentified for SpEd, since in the absence of misidentification we would expect districts to remove students with increasingly more severe disabilities post-policy.

While these patterns point to misclassification as a key mechanism, indirect effects also play an important role in the policy's impacts on both GE and SpEd students. Even for SLD SpEd students, effects on long-run outcomes exceed what SpEd status changes alone could plausibly explain. Three factors likely drive indirect effects. First, since students who lost SpEd services spent most of their day in GE classrooms, they shared learning environments with GE and mainstreamed SpEd peers.[5] Misclassified students may have created classroom disruptions that negatively affected their peers. Research shows that disruption even from just one peer can have strong negative impacts on classroom learning (Carrell, Hoekstra, & Kuka,

[5]Our data contain a categorical variable indicating whether SpEd students spend the majority of their day in GE classrooms. However, our data do not contain classroom-level information, and thus we cannot identify which students are in class with whom.

2018). Their removal from SpEd could have improved learning environments for all students by allowing teachers to focus more on instruction and by improving peer-to-peer interactions. Second, while we do not observe changes in per-pupil SpEd spending, removing services from misidentified students could have freed up resources, improving instructional quality and individualized attention for those who remained in SpEd (Miles & Darling-Hammond, 1998). Finally, the racial specificity of our results suggests that racial bias reduction is a plausible mechanism– benefits of the policy are concentrated almost exclusively among Black students. If initial SpEd placements for some Black students were racially motivated (as evidenced by our findings that SLD students benefit from SpEd removal), being told to reduce these placements may have prompted teachers to re-evaluate their perceptions and biases toward Black students, leading to broader improvements in how these students are treated in schools (Santiago-Rosario et al., 2021; H. Morgan, 2020; Gershenson & Papageroge, 2018).

Our study offers three primary contributions to the literature. First, we provide novel estimates of the long-term impacts of a policy that aims to reduce racial disparities in SpEd classification rates. The U.S. Department of Education requires that school districts monitor and report on the extent to which there is disproportionate representation of students in SpEd across race driven by inappropriate identification for SpEd (Office of Special Education and Rehabilitative Services, 2009). Over the past two decades, there has been on-going debate about whether the over-representation of Black students in SpEd is driven by a greater need for SpEd services or by racial bias (Stiefel et al., 2024; Elder et al., 2021; Fish, 2019; P. Morgan, Farkas, Hillemeier, & Maczuga, 2017; P. Morgan, Farkas, Cook, et al., 2017; Gordon, 2017; P. Morgan et al., 2016; Shifrer et al., 2011; Hibel et al., 2010; Hosp & Reschly, 2003; Donovan & Cross, 2002; Oswald et al., 1999). While many previous studies have documented racial disparities in SpEd, they have not been able to evaluate the long-term impacts of these disparities. Given the large population in Texas, as well as the detailed administrative data we utilize, we are able to make a novel contribution to this literature by being the first, to our knowledge, to trace out the long-run impacts of a policy intervention that reduced Black students' disproportionate placement in SpEd.

Second, our paper contributes to the small but growing literature on the effectiveness of SpEd. On the whole, previous literature finds that SpEd improves student's short and long-run outcomes (Ballis & Heath, 2021; Sallin, 2021; Schwartz, Hopkins, & Stiefel, 2021;

Cohen, 2007; Hanushek, Kain, & Rivkin, 2002). Notably, these papers examine changes in overall access to SpEd, which are not race-specific. We leverage a distinct source of variation that specifically targeted Black student representation in SpEd, allowing us to be the first paper in this literature to identify whether limiting the over-representation of Black students in SpEd helps or harms students' outcomes. Contrary to most of the prior literature, we find that reducing access to SpEd for Black students led to improvements in long-run outcomes. Our results suggest that this improvement operates through multiple channels: direct effects for students who were likely misidentified for SpEd, spillovers due to removing these misidentified students, and broader improvements in racial bias. Ultimately, our findings point to the importance of carefully examining disability evaluation criteria to ensure that students of all races are appropriately evaluated for SpEd placement. Finally, our results highlight the importance of understanding heterogeneity in treatment effects across students when evaluating the benefits of special programs in public schools.

Third, this paper contributes to a large and growing literature evaluating the sources of racial gaps in adulthood. While the quality of early life health care and schools have been linked to racial gaps during adulthood, to our knowledge, there is no existing literature on how SpEd programs and policies may be impacting long-run racial gaps. We know that SpEd can have large and important impacts on long-run outcomes. Yet, how SpEd participation among Black students contributes to disparities across race in adulthood is largely unknown. Our paper helps to fill this gap by investigating the impacts of capping disproportionality on educational attainment in adulthood. We find that students who are likely misidentified for SpEd do better in the long-run. We find positive impacts on their GE and SpEd peers as well. We conclude that reducing misidentification and racial bias in the SpEd process in school can help reduce racial gaps in adulthood. Importantly, this *does not* imply that caps on disproportionality policy are the best way to improve SpEd identification, and instead reiterate the importance of carefully considering the process for identifying students for SpEd to ensure that all students are appropriately placed.

2 Background

2.1 Special Education

In 1975, Congress enacted the Education for All Handicapped Children Act (later renamed the Individuals with Disabilities Education Act (IDEA)). Under IDEA public schools are required to provide a "free and appropriate" education for all students regardless of physical or cognitive disability. This initiated the provision of SpEd services to students with disabilities. To qualify, students must fall within at least one of thirteen disability categories, which include SLD (by far the largest category), followed by speech impairments, intellectual disabilities, and emotional disturbance (see Table 1 for full list).

To be evaluated for SpEd, students are typically referred by a parent or teacher. After the initial referral, students are evaluated to determine what, if any, disability they have and whether this disability adversely affects their educational performance. If a student is deemed eligible, an Individualized Education Plan (IEP) is written for them by a team of professionals, which includes both special and general educators. The IEP states what support and instructional services a student will receive over the course of the school year. For example, typical services may include one-on-one or small group instruction, therapies such as speech, occupational, or physical therapy, counseling for social or emotional challenges, or help with special devices and equipment. IEPs also document accommodations such as specific seating assignments, special equipment (e.g., headphones), and additional time on tests or other assignments (Exceptional Lives, 2025). IEPs are *individualized* and may vary widely so that each student receives a different set of services and accommodations depending on the student's disability and the school they attend.[6]

For many students, the process for determining SpEd eligibility is complex. This is especially pronounced for SLD students. SLD refers to students with learning disorders in reading, math, and/or written expression (Luna, 2024). Unlike more objective disabilities (such as deafness or blindness) with clear medical markers, SLD evaluation relies on profes-

[6]We note that 504 plans are distinct from IEPs. 504 plans are legislated by Section 504 of the Rehabilitation Act of 1973. A 504 plan documents the accommodations a student may receive as a result of a disability, in order to help prevent discrimination and allow equal access to education. 504 plans do not include any instructional components (Rawe, 2024). Our data do not contain any information about 504 plans. Thus, we focus only on IEPs in this paper. Throughout the paper we use the term special education (SpEd) to refer only to those students with IEPs.

sional judgment, which can vary across evaluators. The tests used to evaluate SLD can vary widely, and determining which tests to use and whether a student's academic achievement is sufficiently hindered by their disability to qualify for SpEd can be highly subjective. For example, the Center for Parent Information & Resources (2022) states that "Only by collecting data through a variety of approaches (e.g., observations, interviews, tests, curriculum-based assessment, and so on) and from a variety of sources (parents, teachers, specialists, child) can an adequate picture be obtained of the child's strengths and weaknesses."

2.2 Policy Environment

Amid concerns that minority students were being placed in SpEd at rates which were too high and potentially harmful, the U.S. Department of Education (DOE) began requiring that school districts monitor the disproportionate representation of minority students in SpEd in its re-authorization of IDEA in 1997. The DOE strengthened this requirement in 2004 by requiring that districts allocate 15% of their federal SpEd funding to improving early intervention services for students with disabilities aged six or younger in districts with significant disproportionality. The threshold for what constitutes "significant" disproportionality is left up to states to decide (Office of Special Education and Rehabilitative Services, 2009).

In 2004, the Texas Education Agency introduced the Performance-Based Monitoring Analysis System (PBMAS), which included a cap on disproportionality for Black and Hispanic students.[7] Under PBMAS, any district with a disproportionality rate (defined as the percent of Black or Hispanic students in SpEd minus the overall district percent of Black or Hispanic students) greater than 1 percent was considered out of compliance, and faced state interventions if they did not reduce disproportionality to meet this new target.[8] Based on a district's distance from the relevant threshold and how long they had been out of compliance, sanctions ranged in intensity from districts needing to develop improvement plans to third party on-site monitoring visits.[9]

[7]PBMAS was developed by stakeholders during the summer of 2004 and school districts received their first PBMAS report in December of 2004. Thus, we conservatively treat the 2004-2005 school year as the first year of policy implementation.

[8]Districts were assigned performance levels based on their distance above the 1% disproportionality threshold: less than 1%, between 1% and 2%, between 2% and 5%, and greater than 5%. The greater the disproportionality rate, the more intense the scrutiny and subsequent actions taken to remedy a district's "performance" (Texas Education Agency, 2004).

[9]For districts to comply with the disproportionality caps they could either reduce the SpEd rate for Black

Although federal law requires all states to report on disproportionality, Texas additionally implemented the PBMAS system, which differed substantially from federal requirements in two ways. First, PBMAS uses percentage point differences between SpEd and the overall district rather than the standard risk ratio, which compares the rate of SpEd placement for one racial group to the rate for all others. The PBMAS percentage point threshold is much more restrictive than federal risk ratio standards. For example, a district with 10% Black enrollment could have Black students comprise up to 21.7% of SpEd under federal guidelines (using a risk ratio of 2.5, which is what Texas currently uses for federal reporting (Texas Education Agency, 2025)). In contrast, a district with 10% Black enrollment can have only 11% of Black students in SpEd under the PBMAS disproportionality cap. Second, while federal law requires funding reallocation to early intervention, PBMAS directly incentivizes removing students from SpEd.

Figure 1b illustrates trends in the percent of SpEd students in Texas overall and by race. We see a dramatic decrease in SpEd enrollment after 2004 for each race. Across all years, Black students have higher rates of SpEd compared to White students, whereas Hispanic students have lower rates of SpEd compared to White students. In Figure 2, we show the district-level rates of disproportionality among Black and Hispanic students in Texas across our study period. Of note is the fact that throughout, rates of disproportionality are much higher among Black students compared to Hispanic students. By 2004, the statewide average Hispanic disproportionality rate was already below 0.

It is important to note that PBMAS also monitored the overall SpEd enrollment rate. Any district with over 8.5% of students in SpEd was out of compliance under PBMAS standards. In Ballis and Heath (2021), we explore the impacts of the SpEd enrollment cap in depth. This cap led to significant reductions in SpEd access, which in turn, generated large reductions in educational attainment. As detailed further in Appendix B, we control for the SpEd enrollment cap in this paper to account for any confounding effects, but demonstrate that including this cap does not change the estimated effect of the disproportionality caps. Appendix B also reconciles the differences between the negative impacts of the SpEd enrollment cap identified in Ballis and Heath (2021) and the positive impacts of the Black disproportionality cap here.

or Hispanic students, or increase the SpEd rate for other races. However, we find strong evidence of reductions in Black SpEd enrollment and do not find evidence that the Black or Hispanic disproportionality caps led to increases in SpEd placement for other races.

Ultimately, the negative impacts of the SpEd enrollment cap are driven by Hispanic students who benefited from being in SpEd, whereas the positive impacts of the Black disproportionality cap are driven by Black students who were misidentified for SpEd.

Before 2004, Texas's SpEd system was unremarkable compared to other states in terms of student enrollment. At that time, approximately 12% of Texas students received SpEd services, which was slightly below the national average of 14%, as can be seen in Figure 1a. SpEd enrollment varied significantly across the country. For example, in 2003, Rhode Island had 19.9% of its students in SpEd, while California had only 10.5% (Dhuey & Lipscomb, 2011). By 2016-2017, the average in Texas had fallen to about 8.5%, while the national average remained stable at around 13.5%. This sharp decline highlights the strong downward pressure on SpEd enrollment during this period, which could bias our estimates of the disproportionality caps on student outcomes downward, given the negative effects of the SpEd enrollment cap on SpEd students. However, again we show in Appendix B that controlling for the SpEd cap does not affect our estimates of the disproportionality cap. Further, since we find positive impacts of reducing disproportionality, any potential bias would only attenuate our estimates. Thus, in states without a similar enrollment cap, the positive effects of such policies on Black students could be even more pronounced.

In addition to the outcomes described above, PBMAS also monitors other outcomes related to improving the performance of SpEd students and reducing the amount of time they spent in separate classroom settings, as well as monitoring other groups of students including Bilingual/English as a Second Language and Migrant students.[10] In Ballis and Heath (2021), we show that the majority of districts were already meeting, or nearly meeting, the other thresholds pertaining to SpEd performance prior to policy implementation. In 2005, 99% of districts were meeting or nearly meeting the thresholds limiting disciplinary actions and academic performance, 80% were meeting or nearly meeting the inclusive setting threshold, and 89% were meeting or nearly meeting the unmodified test-taking threshold. Overall, we view it as unlikely that districts responded in significant ways to these other thresholds.

[10]The monitored outcomes for Bilingual/English as a Second Language and Migrant students do not include any thresholds limiting the percent of students in these programs, rather they include outcomes such as passing rates on the standardized exams and high school dropout.

2.3 Theoretical Framework

To motivate our empirical analysis and gain insight into potential mechanisms, we introduce a simple theoretical framework. This framework allows us to examine the decision-making process behind SpEd classification and it's potential impact on educational attainment. We will use this framework to illustrate how the impacts of SpEd can differ across racial groups and the two sources of policy pressure in our context (i.e. to reduce disproportionality vs. overall enrollment in SpEd). Ultimately, our model rests on the presence of racial bias in the SpEd identification process for Black students. We note that while we find several pieces of empirical evidence in support of this, our data does not allow us to implement a direct test of our model. Nonetheless, we believe this framework offers valuable insights into our findings.

Our model assumes that there are just two alternatives: remaining in GE or moving to SpEd. We model the utility of remaining in GE separately for students and schools. First, we consider the utility of remaining in GE for a given student:

$$u_i = a_i - c \tag{1}$$

For simplicity, we assume each student has an underlying "ability" level, a_i, which represents a student's ability to perform in school without SpEd services.[11] There is some fixed threshold, c, such that students with a_i below this threshold obtain greater utility from being in SpEd. Above this threshold, students obtain greater utility from remaining in GE (driven by the fact that there are costs associated with SpEd placement as described in Section 2.1). We can think of c as the net benefit of receiving SpEd services.

Next, we consider the school's utility of a given student receiving GE services. The school's utility for keeping a child in GE is the sum of the utility each student gains from participating in GE ($a_i - c$) plus the cost of providing SpEd services. For simplicity, we assume a fixed cost of providing services, β (in practice cost varies by the severity of a student's disability and the types of services they receive):

$$v_{is} = a_i - c + \beta \tag{2}$$

[11] We are oversimplifying the notion of "ability" with respect to the need for services. We are treating this singular term as representing the spectrum of need for services, whether it be cognitive or physical in nature.

For a school to place a student in SpEd, it would need to be the case that $a_i + \beta < c$. Thus, our model predicts that there is a gap between the student's optimal level of SpEd participation and the school's optimal level, due to the cost to the school of providing services. This implies that students will be placed in SpEd at lower than optimal rates.[12] Therefore, if schools are required to further reduce SpEd placement (such as when Texas implemented the SpEd enrollment cap), removals from SpEd are predicted to worsen students' outcomes, as these are students for whom $a_i < c$.

However, we have yet to consider the issue of race. As previously noted, Black students are placed in SpEd at much higher rates than other races. One explanation for this could be that Black students have a greater underlying need for services relative to other races (i.e., their distribution of a_i could be more skewed to the right). In this case, there would be higher proportions of Black students with a_i below c. This would imply that reducing SpEd would negatively affect Black students' outcomes. However, it is also possible that schools could have an implicit or explicit racial bias that leads to higher rates of Black students in SpEd (Skiba et al., 2008). This would imply that a school's utility of GE placement is:

$$v_{is} = a_i - c + \beta - \gamma_s \tag{3}$$

where γ_s is an indicator equal to 1 if a student is Black. We allow the bias term to be school-specific (denoted by subscript s), as schools are likely to differ in the extent to which they exhibit racial bias in the SpEd identification process. The size of the bias term affects the degree to which students will be misclassified for SpEd. If there is little racial bias (i.e., $\gamma_s \leq \beta$), then removal from SpEd will harm the marginal student. If there is high racial bias (such that $\gamma_s - \beta > a_i - c > 0$) then SpEd removal will benefit the marginal student. Ultimately, we find positive long-run impacts of policy-induced SpEd removal, driven by students who are most likely to be misclassified for SpEd (Section 5.1.2), which is consistent with this theory of racial bias.

[12] This gap in the optimal allocation of SpEd may be reduced if the burden of paying for SpEd fell less directly on the school. While the federal government under IDEA is meant to help offset the cost of providing services, SpEd has been underfunded by the federal government since its inception (Kolbe, Dhuey, & Doutre, 2023). Thus, the majority of the costs fall to the state and local levels.

3 Data

3.1 Data Sources

Data for this paper come from the Texas Schools Project housed at the Education Research Center at the University of Texas at Dallas. These restricted-access administrative data allow us to link individual-level data from public school records from the Texas Education Agency to public post-secondary data from the Texas Higher Education Coordinating Board. We merge these data together to obtain panel data from 1994 to 2017, containing a rich set of individual-level background characteristics and outcomes. Importantly, these data track participation in SpEd, with information on the type of disability and level of classroom inclusion.

In this paper, we do not estimate effects on math and reading exam performance. SpEd students are often exempt from the exams or take modified/accommodated versions. Losing SpEd services is likely to reduce test scores mechanically as a result of no longer having access to modified/accommodated versions. In addition, modified/accommodated versions were not offered until 2001 and are not available in our data until 2008. Therefore, we do not expect the selected scores of only those students who take unmodified versions of the exam to provide an accurate estimate of the effects of the policy on performance in school for SpEd students.

Instead, we focus on long-run outcomes, which include whether an individual graduated from high school and attended post-secondary school in Texas. High school graduation is measured as an indicator for receiving a high school diploma within 2 years of expected graduation, for students observed in our data as of 9th grade. We choose 9th grade in particular to capture students before dropout decisions are made and to minimize counting other reasons for leaving the data in earlier grades as dropping out (such as moving out of state or to private school). Our results are robust to conditioning on 8th grade enrollment instead. For college enrollment, we do not condition on high school graduation and it is censored so that individuals have 6 years after expected high school graduation to enroll in college.

We highlight here that these data only capture public college attendance in the state of Texas.[13] However, outmigration from Texas is very low. As of 2012, Texas had the lowest outmigration of any state, with 82% of people born in Texas living in Texas (Aisch, Gebeloff,

[13]Private universities are not required to share their data with THECB. Those that do share their data account for just 3.5% of our sample, and results are robust to including or excluding these observations.

& Quealy, 2014). College attendance out of state is also very low among students in Texas. For a subset of cohorts that can be linked to the National Student Clearinghouse, in 2008 and 2009 only 3.7% of students attended college out of state (compared to 64.5% who attended in-state) (Mountjoy, 2022) and from 2008 to 2012 only 1.7% of SpEd students enrolled in college out of state within two years of their high school graduation (Ballis & Heath, 2021). Finally, although post-secondary completion and earnings are available in the data, the policy change occurs too close to the end of our data to provide accurate estimates of changes in these outcomes.[14] Thus, we leave for future work estimates of the impact of the policy on changes in college completion and earnings in the labor market.

3.2 Summary Statistics

Table 1 presents descriptive statistics for all students, Black students, and Hispanic students, as well as SpEd students by race for our main analysis sample. As we will justify in Section 4, we focus on students entering 5th grade between 1994 and 2004. Overall, about 14% of students are in SpEd, 14% are Black, and 39% are Hispanic. Black students have a higher SpEd rate at 19% relative to Hispanic students at 14%. Our final analysis sample consists of 72,725 Black students in SpEd at 5th grade and 158,855 Hispanic students in SpEd at 5th grade. For SpEd students, we only have information on disability type and classroom setting. Among all races, SLD is the most common disability type, followed by speech impairments. The vast majority (roughly 85%) of SpEd students spend greater than 50% of their day in the GE classroom.

Appendix Table A.1 illustrates raw differences in the pre-policy characteristics of districts that are above and below the 1% thresholds for the Black and Hispanic disproportionality caps. Districts above the Black disproportionality threshold have more Black students and fewer Hispanic students. They also have lower rates of FRL and Title I students, implying that these districts are less economically disadvantaged. Most other observable characteristics do not vary significantly across districts above and below the threshold, and for those that do

[14] We have access to data through 2017 and are thus only able to follow the youngest cohort of students in our sample through 6 years post expected high school completion. However, for Black and Hispanic students in SpEd as of 5th grade between 1994 and 2000 the average number of years between expected high school completion and associate's degree attainment is 7.3 and for bachelor's degree attainment is 6.2. The 75th percentile for earning an associate's degree is 10 and for a bachelor's degree is 8. We would thus ideally examine the effects of the policy on college completion and earnings at least 10 years after high school completion.

the differences are small in magnitude. A similar pattern emerges across districts above and below the Hispanic disproportionality threshold. There are more Hispanic students and fewer Black students in districts above the Hispanic threshold. We account for differences in baseline characteristics in our empirical strategy by including controls for each of these variables at the individual and cohort-district level. In addition, our results are robust to controlling for district-level time-trends in the baseline levels of the demographic variables (Section 5.1.1).

4 Empirical Strategy

We estimate the causal impact of reducing disproportionality using cross-district and cross-cohort variation in exposure to the disproportionality caps. We employ a dose-response difference-in-differences estimation strategy to determine whether students in districts with higher rates of disproportionality at baseline experience larger changes in outcomes.[15] We estimate effects separately for Black and Hispanic students, and include the Black disproportionality rate in models estimated for Black students and the Hispanic disproportionality rate in models estimated for Hispanic students.

Given the nature of the policy change, we are not able to causally estimate the effect of the policy by simply comparing SpEd student outcomes before and after policy implementation. To reduce their SpEd rate, districts must decide which students will be removed from SpEd and which students will never be placed in SpEd, thus impacting the underlying ability distribution of students who remain in SpEd. Instead, we estimate the effect of limiting access to SpEd for students already identified before the policy. To do so, we select students who were in SpEd as of 5th grade before the policy. This is a reasonable choice since most SpEd enrollment decisions take place prior to 5th grade. Furthermore, Appendix Table A.2 illustrates that results remain similar when we use students in SpEd as of 3rd, 4th, or 6th grade prior to policy implementation instead.[16] Similarly, to estimate effects on GE students we focus on

[15]Recent literature has made important advancements in our understanding and use of the DiD model. Unfortunately, most of this recent literature is not applicable to our setting. For example, Callaway and Sant'Anna (2021) is applicable to settings where treatment turns on (and off) at different times. In our case, treatment turns on at the same time for all units. Sant'Anna and Zhao (2020) propose a doubly-robust DiD estimator, however, this work has not yet been extended to the continuous treatment case. Additionally, we do not use time-varying controls in our DiD specification and we show in Section 5.1.1 that our results are robust to the inclusion of background characteristic-specific trends.

[16]Appendix Figure A.1 illustrates the percent of all students entering SpEd by grade. The fraction of new entries levels off around 4th grade and drops each year after that. Appendix Figure A.2 shows that less than 6% of students exit SpEd in each grade prior to 5th grade. Thus, estimating effects as of 5th grade captures

students in GE as of 5th grade before the policy.

Specifically, we estimate the following difference-in-differences specification on either the sample of 5th grade SpEd or GE students:

$$Y_{idc} = \beta_0 + \beta_1 Disp_{2004,d} * FracExpo_c + \beta_2 SpEd_{2004,d} * FracExpo_c + \beta_3 X_{idc} + \eta_d + \theta_c + \varepsilon_{idc} \quad (4)$$

where Y_{idc} is an outcome of interest for individual i, enrolled in school district d, in cohort c. We estimate the impact of the disproportionality caps on the likelihood of participating in SpEd by expected 9th grade and on the long-run outcomes of high school completion and post-secondary enrollment. We measure SpEd status as of 9th grade since this is prior to when most dropout decisions are made. Additionally, this is measured as *expected* 9th grade, that is, 4 years after 5th grade to avoid endogenous changes in grade repeating.

The term $Disp_{2004,d}$ is the *relative* Black (or Hispanic) disproportionality rate at the district-level in 2004. That is, $Disp_{2004,d}$ equals the 2004 Black (or Hispanic) disproportionality rate minus 1% if above the 1% target, and is set to 0 if below 1%. Rather than interact this with an indicator for the post-policy period, we interact it with $FracExpo_c$, which measures the fraction of years an individual is in school under the policy. For the outcome of SpEd status at 9th grade, $FracExpo_c$ takes on the values of 1/4, 1/2, 3/4, and 1, depending on how many years the policy was in effect between 5th and (expected) 9th grade for each student. For high school completion and college enrollment, $FracExpo_c$ equals 1/7, 2/7, ..., 7/7, depending on how many years the policy was in effect between 5th and (expected) 12th grade. This captures the fact that students exposed to the policy for longer are more likely to experience larger impacts. The main coefficient of interest is β_1, which gives the effect of reducing disproportionality among Black (or Hispanic) students who were exposed to the policy in *every year* between 5th grade and when the outcome is measured.

The term X_{idc} represents a vector of individual and district-cohort level controls including gender, free and reduced-price lunch (FRL) status, English as a Second Language (ESL) status, gifted status, and Title I status measured as of 5th grade. When estimating results for the SpEd sample, we include controls for baseline disability type and an indicator for whether the

the majority of students identified between kindergarten and 5th grade. Appendix Table A.2 shows that the effect on 9th grade SpEd status becomes larger as we look across grades, likely due to the fact that relative to students in SpEd at 6th grade, students in SpEd at 3rd grade have more severe disability types.

student spent greater than 50% of the day in a GE classroom at baseline. When estimating results for the GE sample, we control for 5th grade math and reading standardized exam scores. In all models, we also include district fixed effects, η_d, and cohort fixed effects, θ_c. Standard errors are clustered at the district level, since this is the level at which treatment varies. As mentioned previously, the policy simultaneously introduced a cap on overall SpEd enrollment at 8.5%. To account for this pressure, we control for the interaction of $SpEd_{2004,d}$, the relative percent of students in SpEd in 2004 above the 8.5% target (and 0 otherwise) in each district and $FracExpo_c$. We expand on the discussion and justification of including this control in Appendix B. Importantly, Appendix Table B.1 illustrates that the effect of the disproportionality caps remain quantitatively and qualitatively similar when we do not control for the SpEd enrollment cap.

The main identifying assumption for our models is: conditional on the fixed effects and observable characteristics, trends in outcomes among districts with low disproportionality rates provide an accurate counterfactual for trends among districts with high disproportionality rates. We test this assumption directly with an event study analysis:

$$Y_{idc} = \beta + \sum_{t=1998, t\neq 2004}^{2008} \{(\gamma_t Disp_{2004,d} * t) + (\alpha_t SpEd_{2004,d} * t)\} + \delta X_{idc} + \eta_d + \theta_c + \varepsilon_{idc} \quad (5)$$

where 9th grade cohort indicator variables are interacted with $Disp_{2004,d}$ and $SpEd_{2004,d}$. We exclude the 2004 9th grade cohort for the outcome of SpEd status at 9th grade, since this is the last cohort to be unexposed to the policy during 9th grade. For the long-run outcomes, we exclude the 2001 9th grade cohort, since this is the last cohort unexposed to the policy between 5th and 12th grade. The results of this analysis are presented in Section 5. On the whole, we do not find evidence of pre-treatment trends.[17] Finally, for our specifications to be identified it must also be the case that there are no contemporaneous shocks correlated with treatment and outcomes. We address this assumption in Section 5.1.1 and conclude that there were no contemporaneous shocks likely to influence our results.

[17]It is also unlikely that school districts anticipated these policies and modified their SpEd practices leading up to the policy implementation. Not only were the disproportionality caps not widely known to the public, neither was the SpEd enrollment cap, which was revealed 12 years after its implementation in a 2016 investigative Houston Chronicle article (Rosenthal, 2016).

5 Results

5.1 Black Disproportionality on Special Education Students

We begin with the effect of the Black disproportionality cap on Black SpEd students. Starting with event study estimates, Figure 3 plots the coefficients from 9th grade cohort indicator variables interacted with the 2004 Black disproportionality rate. The figures on the left present results for the full sample, and those on the right present results for SLD students. In the pre-policy period, for 9th grade SpEd placement, figures 3a and 3b show that cohorts entering 9th grade before 2004 exhibited parallel trends across high and low disproportionality districts. For long-run outcomes, Figures 3c through 3f show that cohorts entering 9th grade before 2001 (who reached 12th grade in 2004) also exhibit parallel trends across high and low disproportionality districts.

In the post-policy period, figures 3a and 3b show a downward trend in the likelihood of continuing in SpEd in 9th grade. The negative impact of the policy on SpEd participation is stronger for students exposed to the policy for longer. SpEd removal can be a lengthy process, requiring meetings and re-evaluation to determine that the student no longer qualifies for services. Consequently, the 2005 9th grade cohort (one year under the policy) shows little change, whereas the 2008 9th grade cohort (4 years under the policy) experiences large declines in SpEd participation at 9th grade.[18] The negative impact is especially pronounced for SLD students, consistent with these students being relatively more marginal.

In the long-run, Figures 3c and 3d show a small upward trend in high school completion post-policy. Again, the impacts of the policy become stronger as students are exposed for longer. In particular, the policy did not have a strong impact on the 2002 9th grade cohort, as these students were only exposed to the policy in 12th grade. Both direct (i.e. SpEd removal) and indirect channels (i.e. peers' SpEd removal and classroom environment changes) are likely to be small for the earlier cohorts. The 2005 9th grade cohort (exposed to the policy in all four years of high school) marks the first clear increase in high school completion. Effects for subsequent cohorts are also high, as their exposure extends back into middle school. Figures 3e and 3f, also show a positive trend for college enrollment, with the largest and most significant

[18] We also note that the policy did not require immediate compliance with the 1% threshold. Rather, it required incremental progress toward compliance.

impacts experienced by those exposed to the policy for longer.

Overall, the fact that we find a stronger negative pattern for 9th grade SpEd placement for SLD students (Figure 3b) compared to the full sample (Figure 3a) suggests that non-SLD were *not* more likely to lose SpEd. The full sample estimates are attenuated relative to the SLD sample, implying close to zero impacts for the non-SLD group. In contrast, the event studies for high school completion and college enrollment are quite similar for the full sample and SLD students. This similarity suggests that non-SLD students likely experienced similar long-term gains as SLD students. Although we will find that results for the non-SLD subgroup are imprecisely estimated (Table 2), this comparison provides suggestive evidence that the policy generated long-run benefits for non-SLD students that are not driven by changes in SpEd placement. As we will discuss in Section 5.1.2, these patterns align with mechanisms operating through indirect effects, such as changes in classroom environments.

Turning to our main table of results, columns (1)-(3) of Table 2 present estimates of the impact of the Black disproportionality cap for all Black SpEd students. We start with a model that only includes district and cohort fixed effects, and consecutively add individual and then district-cohort level controls. The significance of our estimate for SpEd participation at 9th grade increases when we control for classroom setting in 5th grade. Intuitively, this makes sense, as whether a student spends the majority of their day in a GE or SpEd classroom likely serves as a proxy for the severity of the disability.[19] Additional controls have minimal effect on the magnitude and significance of our estimates.

In the fully specified model in column (3), we find that Black SpEd students at the average district (that was 3.2 p.p. above the 1% threshold) who were exposed to the policy in every year between 5th and 9th grade experienced a 1.2 p.p. (0.376*3.2) decrease in the likelihood of continuing in SpEd at 9th grade. This represents a 1.6% (1.2/77.4) decline in 9th grade SpEd participation. In the long-run, the Black disproportionality cap improved Black SpEd student's outcomes. For those at the average district who were exposed to the policy in every year between 5th and 12th grade, the likelihood of completing high school increased by 1.3 p.p. (2.1%) and college enrollment increased by 1.5 p.p. (4.9%).[20] This increase in college

[19]The estimate is stable as we consecutively add individual and district-cohort level controls when estimating effects separately for SLD and non-SLD students, in columns (4) and (5), respectively.

[20]To account for multiple inference, we examine the impact of the Black disproportionality cap on a summary index of long-run outcomes, which is computed as the equally weighted average of the z-scores of high school

enrollment is driven by increases in 2-year college (rather than 4-year college).[21]

We additionally estimate effects for SLD and non-SLD students in columns (4)-(5) of Table 2. As expected, the effects of the Black disproportionality cap are driven by SLD students. For Black SLD students, we estimate a 1.8 p.p. (2.2%) decline in SpEd enrollment, a 1.4 p.p. (2.2%) increase in high school completion, and a 1.4 p.p. (4.6%) increase in college enrollment. In contrast, we do not find statistically significant effects on non-SLD Black SpEd students. The magnitude of the point estimate on SpEd removal is very close to zero, as expected. While the coefficients for educational attainment are positive and the magnitudes are reasonably sized, they are imprecisely estimated with large standard errors.[22]

Heterogeneity- We begin by examining heterogeneity across gender and income. Table 3 demonstrates that while the disproportionality cap led to declines in SpEd participation for both males and females, the positive long-run effects are driven by males (column 1). By income, we find statistically significant declines in SpEd participation for FRL students. The effect on non-FRL students is also negative but not precisely estimated, perhaps driven by the smaller sample size.[23] In the long-run, we find statistically significant increases in the likelihood of high school completion for FRL students, and increases in college enrollment for both FRL and non-FRL students. The fact that we find large positive impacts on college enrollment among higher income students (who do not experience significant changes in SpEd placement) provides compelling evidence that spillovers such as less exposure to disruptive peers, changes in teaching practices, and changes in school culture played an important role. We discuss mechanisms further in Section 5.1.2.

completion and college enrollment (Kling, Liebman, & Katz, 2007). This summary index mitigates the issue of multiple hypothesis testing by reducing the number of individual tests, simplifying the inference process, improving statistical power, and providing a clearer and more interpretable overall measure of the effect. Results using this summary measure (available upon request) indicate a statistically significant improvement in the long-run outcomes of Black SpEd students.

[21]These results are available upon request. Note also that 4 year college is a much rarer occurrence than 2 year college enrollment. Table 1 illustrates that 28% of Black SpEd students enroll in 2 year college overall, whereas just 4% enroll in 4 year college.

[22]In Appendix Table A.3 we estimate effects for each disability type separately. The other disability categories are quite small and all have large standard errors, making it challenging to draw firm conclusions. The only other disability type that experiences long-run impacts is Other Health Impairment (OHI). For these students, we find a 1.9 p.p. (2.2%) decrease in SpEd enrollment (significant at the 10% level) and a 3.2 p.p. (8.6%) increase in college enrollment.

[23]Further, if SpEd removal is viewed as non-advantageous by parents, non-FRL students are more likely to have parents who have the resources (e.g., time and institutional knowledge) to advocate that their children remain in SpEd.

We also investigate heterogeneity across urban/rural districts and district size. In Appendix Table A.4, we find that effects are driven by urban districts (column 1), with very little impact on rural districts. In columns (3) through (6), we split districts into quartiles of district size. We find that impacts are driven by midsize districts - those in the 2nd and 3rd quartiles, with district size ranging from 5,282 to 51,838 students.

Intermediate Outcomes- Next, we investigate several intermediate outcomes with the goal of better understanding what may have led to the improvements in educational attainment. First, we investigate the effect of the policy on the likelihood of switching districts between 5th and 9th grade. Appendix Table A.5 shows a statistically significant increase in the likelihood of switching districts for Black SpEd students. We cannot precisely determine whether students switched districts before or after losing SpEd, since each are observed only once per year. However, the grade of SpEd removal and switching roughly coincide.[24] Previous literature finds that moving districts disrupts learning (Welsh, 2017; Gasper et al., 2012; Hanushek et al., 2004). Thus, we would expect increased switching to reduce long-run outcomes, which could attenuate our estimates. As shown in Appendix Table A.6 column (2), when we re-estimate our model including a control for district switching, we find that the coefficient on switching predicts a reduction in high school completion and college enrollment.[25] Reassuringly, the effect of the disproportionality cap remains quantitatively and qualitatively very similar. This provides compelling evidence that the positive long-run effects are a result of the disproportionality policy, rather than the result of switching districts.

We also estimate effects on absences, suspensions, expulsions, grade repeating, and 8th grade test participation in Appendix Table A.7. Across these outcomes we find no statistically significant changes, except for an increase in the likelihood of *taking* the 8th grade math and reading exams. This is consistent with removal from SpEd eliminating test-taking exemptions. Importantly, an increase in test-taking biases us against finding a positive impact of SpEd removal on high school completion. Once removed from SpEd, students no longer qualify for exemptions from the standardized exit exams required for high school graduation, thereby

[24] We estimate separately the effect of the policy on losing SpEd and on switching districts for each grade between 5th and 9th grade. We find the effect is statistically significant for the first time 3 years after 5th grade for both the likelihood of losing SpEd and switching districts.

[25] Although switching is an intermediate outcome, this type of specification is similar in spirit to the remediation analyses performed in Baron, Hyman, and Vasquez (2022).

raising the standards to graduate.

5.1.1 Robustness

While our event studies provide evidence in support of the parallel trends assumption, we implement a series of additional checks to test the plausibility of this assumption. First, we rule out the possibility that districts facing greater pressure under the policy were on differential trends driven by differences in observable baseline characteristics. To do so, we add to our specification one-at-a-time trends in district-level demographic characteristics based on the demographics that were statistically significantly different across districts above and below the black disproportionality threshold. In Appendix Table A.8 we find that our results are robust to including these trends.

Next, we investigate whether the policy led to differential rates of attrition. If students more exposed to the policy pressure to reduce disproportionality moved out of Texas public schools (perhaps upon being denied SpEd in their current district), this could have changed the underlying composition of students in districts with high rates of disproportionality such that parallel trends were less likely to continue. In Appendix Table A.9, we directly test whether 5th grade Black SpEd students are systematically moving out of public schools by estimating the effect of the policy on the likelihood of leaving the data between 6th and 9th grade. We find marginally significant *increases* in the likelihood of being enrolled for FRL and non-FRL students. The fact that we do not find a decrease in enrollment by 9th grade, especially for non-FRL students provides evidence that the disproportionality cap did not lead parents to seek SpEd services elsewhere (such as out of state, in home school, or in private school). Instead, we conclude that this reflects the fact that Black SpEd students are more likely to stay in school, and in turn more likely to graduate from high school and enroll in college as a result of the Black disproportionality cap. Furthermore, increases in students on the margin of dropout in our sample would only attenuate the positive effects of the Black disproportionality cap on Black SpEd students' long-run outcomes.

Another important assumption of our model is that there were no other policy changes around the same time that are confounding our estimates. The only educational policy, to our knowledge, implemented around the same time as the PBMAS was the federal accountability system, No Child Left Behind (NCLB), implemented by former President George W. Bush

in 2003. Texas already had an accountability system in place that had been implemented under President Bush when he was governor of Texas. Since many features of NCLB mirrored those of the existing accountability system that had been in place in Texas since 1993, we do not expect that NCLB played a large role in Texas. The main difference between Texas' accountability system and NCLB is that NCLB monitored the performance of SpEd students as their own subgroup on the standardized exams. However, the achievement standards that were set under NCLB were very low, as the vast majority of districts (97%) were already meeting the performance ratings set by NCLB, which were identical to those under PBMAS (Ballis & Heath, 2021). In addition, Prenovitz (2017) finds that NCLB led to incentives to place relatively higher performing students into SpEd in North Carolina to boost the performance ratings of the SpEd subgroup, which is an incentive working in the opposite direction of the disproportionality cap aimed at reducing access to SpEd in our setting.

5.1.2 Mechanisms

To this point, we have found that the Black disproportionality cap had positive long-run impacts on Black SpEd students, which were driven by SLD students, males, those in midsize districts, and those in urban districts. While SpEd removal plays a critical role in the long-run impacts for SLD students, the magnitude of our long-run estimates relative to the magnitude of SpEd removal suggests that indirect effects also play an important role. In particular, if SpEd removal was the only channel affecting educational attainment, the implied IV for SLD students would be quite substantial, implying an increase in high school completion by 78% and college enrollment by 81%.[26] Further, for several subgroups, such as non-FRL and male students (Table 3), the impact on college enrollment is larger than the impact on SpEd participation. Although we cannot fully separate the direct impacts of SpEd removal from indirect effects such as classroom spillovers and reductions in racial bias, we explore evidence for each in turn.

Direct Effects - As described in the Theoretical Framework (Section 2.3), if students are misclassified for SpEd, their removal should lead to long-run gains by correcting inappropriate

[26] Although the estimates in this context are driven by a combination of the direct effects of SpEd removal as well as spillover effects, it is conceivable that SpEd removal would have large impacts. Ballis and Heath (2021) found that SpEd removal decreased high school completion and college enrollment by 51.9% and 37.9%, respectively.

placements. Consistent with this framework, we find several pieces of evidence showing that subgroups most likely to be misclassified for SpEd benefited the most in the long run. This provides suggestive evidence that the positive impacts for SLD students removed from SpEd as a result of the policy are driven by those who were misidentified for SpEd.

First, we investigate how effects of the policy vary based on whether districts conditionally over- or under-represent Black students in SpEd. As discussed previously, Black students in Texas (and in the U.S. in general) are placed in SpEd at higher rates than White students. If after conditioning on background characteristics, such as income and academic achievement, Black students are still over-represented relative to White students this could indicate that these Black students are misidentified for SpEd. If conditional over-representation reflects *misidentification* (that is, placements where the costs of SpEd exceed the benefits), SpEd removal could improve long-run outcomes. Conversely, if districts with a conditional under-representation of Black students do not have misidentified students (and thus have placements where benefits exceed costs), SpEd removal could harm long-run outcomes.

Following Elder et al. (2021), we develop a model to predict whether a district conditionally over- or under-represents Black students in SpEd. We then investigate whether the effects of the policy differ across these two types of districts. We use a Blinder-Oaxaca decomposition to categorize districts into those having a conditional over- or under-representation of Black students in SpEd. First, we use a logit model to predict the likelihood of SpEd placement for White students, based on pre-treatment characteristics.[27] Next, we apply the coefficients from this model to Black students, to predict the likelihood of SpEd placement for Black students as if they were White. Then, we subtract the prediction from an indicator for whether a student is actually in SpEd. This gives us a measure of whether the student is predicted to be over- or under-represented in SpEd relative to an observationally-equivalent White student. Finally, we aggregate these differences to the district-level, to obtain a prediction for whether each district has an over- or under-representation of Black students in SpEd on average.

[27]SpEd status is predicted as of 5th grade, using baseline covariates measured as of 3rd grade, including age as of September 1st, gender, FRL, ESL, Bilingual, Title I, At Risk, Gifted, LEP, Migrant status, math and reading scores, and disciplinary actions (available upon request). One important caveat of this analysis is that predicting SpEd participation is very difficult with most datasets available to researchers. In particular, the R-squared from our logit model predicting SpEd participation for White students is 0.221. While prior papers have utilized birth certificate records (Elder et al., 2021) to improve the predictive power of these models, there is still a lot of variation in SpEd placement that cannot be explained even with administrative datasets on test-scores, discipline, and measures of infant health.

Our estimates for the impact of the policy on Black SLD students separately by districts conditionally over- or under-representing Black SpEd students are presented in the first two columns of Table 4. In line with our prior, the negative impact of the disproportionality cap on 9th grade SpEd placement is driven by districts with a conditional over-representation of Black students in SpEd. Additionally, we find statistically significant increases in the likelihood of high school completion and college enrollment for Black SLD students in districts predicted to over-identify Black students for SpEd. For districts predicted to under-identify Black students for SpEd, we do not find a statistically significant impact on SpEd status or educational attainment. These results provide suggestive evidence that the positive impacts of SpEd removal are driven by those who were initially misidentified for SpEd.

Since our Blinder-Oaxaca decomposition has limited predictive power for SpEd placement, we complement it by examining impacts for Black SLD students across districts with low vs. high shares of Black students. Elder et al. (2021), Fish (2019), and Stiefel et al. (2024) find that minority students tend to be conditionally over-represented in SpEd in districts where they are most dissimilar to their peers (i.e., districts with low minority shares) and tend to be conditionally under-represented in districts where they are not dissimilar to their peers (i.e., districts with large minority shares).[28] Therefore, we compare the effects for Black SLD students in districts with less than 20% Black students vs. districts with greater than 80% Black students.[29] We use this district-level split as a proxy for contexts where Black students are more likely to be misidentified for SpEd (low Black share districts) versus appropriately identified (high Black share districts). Consistent with our theory, the negative impact on 9th grade SpEd placement and positive impacts on educational attainment are driven by districts with low Black shares (column 3 of Table 4). In fact, we find negative impacts on educational attainment (although with large standard errors) in high Black share districts (column 4). This provides additional evidence that the benefits of SpEd removal are strongest in contexts where Black SLD students were more likely to be misclassified.

[28]The authors of these papers argue that if SpEd identification depends on relative peer comparisons within schools, then school segregation may create different identification patterns. Minority students may not stand out as needing services when compared to peers in high minority share schools (leading to under-representation), but may appear to need more support relative to peers in high White share schools (leading to over-representation).

[29]Note that Elder et al. (2021) are combining Black and Hispanic students in their paper into one "minority" group. We chose not to do the same in our setting, since the shares of Hispanic students are much higher in Texas than in Florida, and thus, are not in the minority.

Next, we investigate whether impacts for Black SLD students differ by district-level teacher racial composition and experience.[30] It may be the case that Black teachers are less likely to misidentify Black students for SpEd. For example, Hart and Linday (2024) find that Black students matched to Black teachers are less likely to be identified for SpEd, on average. In Table 5 columns (1)-(2), we present estimates for Black SLD students in districts with above or below the 90th percentile proportion of Black teachers (i.e., 39.5 percent) at baseline.[31] As hypothesized, the positive long-run impacts of the Black disproportionality cap are driven by districts with below the 90th percentile proportion of Black teachers. We find positive impacts of the policy on educational attainment in districts with "low" proportions of Black teachers and negative (although not statistically significant) impacts on educational attainment in districts with high proportions of Black teachers.[32] This provides further evidence consistent with the theory that Black students are benefiting from SpEd removal due to misclassification for SpEd. Further, we investigate whether districts with lower teacher experience were more likely to misclassify Black students for SpEd pre-policy. In Table 5 columns (3)-(4), we present effects for Black SLD students in districts whose mean level of teacher experience is above or below the statewide average level of experience (11.7 years). We find that the Black disproportionality cap has a larger positive impact on high school completion in districts with below-average teacher experience. However, the negative impacts on SpEd placement and positive impacts on college enrollment are driven by districts with above-average teacher experience. Thus, we do not find strong evidence that low teacher experience is driving the results.

Finally, we investigate whether the types of students removed from SpEd before vs. after the cap was implemented differed in observable ways. This can provide insights into whether districts changed the types of students selected for SpEd removal after being pressured to

[30]Our data only contain district-level staffing data, and thus, we are not able to look at teacher experience or race at the student-level.

[31]The average percent of Black teachers in each district is 9.3%, the median is 3.6%, and the 75th percentile is 8.9%. We split our sample into districts with above or below the 90th percentile proportion of Black teachers in order to have a reasonably large proportion of Black teachers in the "high" same-race teacher category.

[32]Note that we find larger negative impacts on SpEd placement at 9th grade in districts with high shares of Black teachers compared to those with "low" shares. Districts with high shares of Black teachers also have higher shares of Black students and Black disproportionality at baseline. Thus, these may be districts that experienced pressure to reduce Black SpEd enrollment to comply with the disproportionality cap, but may not be places that were likely to have misidentified students, and therefore didn't benefit from this SpEd reduction.

reduce the over-representation of Black SpEd students. We estimate a district-level regression for Black students in SpEd as of 5th grade. We compute the outcome for each district within each cohort as the difference between the percent of students with a particular attribute (e.g. male, ESL, FRL) who are not in SpEd at 9th grade (given SpEd at 5th grade), minus the total percent of students with that attribute in SpEd in 5th grade. Table 6 demonstrates that capping Black disproportionality increased the likelihood that the Black students removed from SpEd post-policy were relatively higher performing in reading and in less restrictive classroom settings (i.e., spending less than 50% of the day in resource rooms) at baseline. Thus, students with relatively more mild conditions were being removed from SpEd post-policy, rather than those with increasingly more severe conditions (which is what we would expect a priori if districts did not have any misidentified students). This provides further evidence that districts were removing those who had been misidentified for SpEd.

Ultimately, given the nature of our data, we can not identify the source behind the apparent misidentification. However, we propose several possibilities. First, there may be implicit or explicit racial bias from teachers in the SpEd referral process (Dever et al., 2016; Sabine et al., 2015; Tobias et al., 1983, 1982). Second, there may be racial bias inherent in the evaluation process, either resulting from the questions on the test used to evaluate students or resulting from biases on the part of the test administrator (Artiles et al., 2002; de la Cruz, 1996; Rose & Huefner, 1984). Finally, there may be biases driven by differences in other characteristics that are correlated with race, such as income, which could lead to bias in the referral and evaluation process for SpEd (Dever et al., 2016; Podell & Soodak, 1993).

Indirect Effects - Next, we turn to investigating indirect effects of the policy. Again, the magnitude of the impacts in the long-run are larger than what could plausibly be explained by changes in SpEd status alone. First, removing SpEd from misclassified students likely improved their classroom behaviors, generating positive spillover effects on other students. Since 80% of Black SpEd students spend the majority of their day in GE classrooms, behavioral improvements among formerly misclassified students likely has positive spillover effects on both GE and SpEd students. Second, a reduction in disruptive behaviors could have improved instructional quality, as both GE and SpEd teachers would have more time to spend on instruction rather than managing behaviors. Third, the policy could have freed up resources

for those who remain in SpEd. Although we do not observe changes in per-pupil SpEd spending (Appendix Table A.10), removing misidentified students could have resulted in a re-allocation of resources (Miles & Darling-Hammond, 1998). For example, the SpEd rate fell by 2.3 p.p. (or 13%) relative to what would be expected pre-policy at the 90th percentile district.[33] Thus, smaller caseloads would allow SpEd and GE teachers to focus more intensively on remaining SpEd students. Teachers could have shifted supports away from managing misidentified students toward instruction for remaining SpEd students in need of services.

Finally, the fact that benefits concentrate almost exclusively among Black students, with only minimal gains for White and Hispanic students, suggests that the policy may have prompted districts to address forms of racial bias specific to Black students.[34] This could operate through several channels. First, the fact that racial bias likely played an important role in the initial misclassification of Black SpEd students suggests that reductions in racial bias could play a role in explaining the positive impacts more broadly (H. Morgan, 2020). Being told to reduce SpEd placements for Black students could be viewed as a signal that these placements were racially biased. This may have prompted teachers to re-evaluate their perceptions of these students, reassess how similar behaviors are interpreted across race, and reset expectations for Black students more positively (Santiago-Rosario et al., 2021; Gershenson & Papageroge, 2018). Further, as teachers observe positive improvements in their Black students' performance (evidenced by their improvements in educational attainment), this may also lead them to reevaluate their perceptions of their Black students. Second, Black students may feel less at risk for being targeted by racial discrimination and less at risk for being placed in SpEd as a result of the policy. Black students may interpret this as an improvement in teachers' views of them, which could boost students' own confidence, improve daily interactions, and reduce perceptions of racial bias.

[33]To determine the share of 9th grade SpEd students absent the policy, we subtract the share of students who typically exit SpEd by 9th grade from the share in 5th grade, and add the share who newly enter by 9th grade. To determine the share of SpEd students under the policy, we adjust the annual exit and entry probabilities by our estimated policy effects, scaled by how far the district is above the disproportionality target. We report the effect for a district at the 90th percentile above the Black disproportionality cap.

[34]We find small marginally significant positive impacts of the Black disproportionality cap on White and Hispanic SpEd students' college enrollment. These results are roughly half the size of the impact on Black SpEd students. We also find a small marginally significant positive impact on Hispanic GE students' high school completion.

5.2 Black Disproportionality on General Education Students

We now turn to estimating the impact of the Black disproportionality cap on GE students. While the cap targeted SpEd students, the policy may have also directly or indirectly affected GE students. Direct effects could be driven by reductions in SpEd in later grades. Indirect effects could be driven by spillovers from SpEd peers (both those removed from and remaining in SpEd), as well as possible changes in teaching practices and racial bias. Since the majority of SpEd students spend most of their day in GE classrooms, indirect effects would operate through impacts within the classroom rather than compositional changes in who is in the classroom. The effects for GE students could be positive or negative. On the one hand, removing SpEd from peers could lead to negative effects if there are fewer resources, such as teacher's aides, available in the GE classrooms. On the other hand, effects could be positive if students experience positive spillovers from their SpEd peers and/or if the policy helped reduce racial disparities in SpEd placements and education policies more broadly.

Figure 4 presents event study estimates for the effect of the Black disproportionality cap on Black GE students. For each outcome we do not find evidence of differential trends in the pre-period across districts more or less treated by the Black disproportionality cap. In Table 7 column (1), we present estimates for Black GE students.[35] For those in the average district exposed to the policy every year between 5th and (expected) 9th grade, the Black disproportionality cap reduced the likelihood of participating in SpEd in 9th grade by 0.32 p.p. (7%). This implies that at least part of the impact on GE students will indeed be driven by the direct effect of being less likely to receive SpEd later on. However, the percentage point change on the impact of SpEd participation at 9th grade is much smaller for GE students than for SpEd students. This can be seen visually in the event study (Figure 4a), where both the point estimates and 95% confidence interval for SpEd participation among Black GE students are significantly smaller than for Black SpEd students (Figure 3a).

In the long-run, consistent with the effects we found for Black SpEd students, we find improvements in Black GE students' educational attainment. For students at the average district exposed to the policy every year between 5th and 12th grade, the Black disproportionality

[35]We include an estimate of the Hispanic disproportionality cap on Black GE students to investigate the spillover effect of Hispanic disproportionality on Black GE students. The effect of this cap will be discussed in Section 5.3.

cap increased the likelihood of completing high school by 0.51 p.p. (0.7%) and enrolling in college by 1.7 p.p. (3.1%).[36] The effects on college enrollment are again driven by 2 year, rather than 4 year college.

To better understand which GE students are driving the improvements in educational attainment, we estimate effects by baseline achievement quintiles in Appendix Table A.11. For baseline achievement in Math, we find that the reductions in SpEd participation are driven by those in the lowest quintile. For reading achievement, although the estimate on SpEd participation is no longer significant (perhaps due to a lack of statistical power), the magnitude of the effect is largest for those in the lowest quintile. Since at least some of the positive long-run impacts for these students is driven by the direct impact of being less likely to participate in SpEd, part of the mechanism may be in-line with a similar story of reduced misclassification, as we argue for SpEd students. Interestingly, all Black GE students experienced increases in college enrollment. This implies that the positive effects experienced by those in the highest quintiles will be driven by indirect mechanisms, since these students did not experience changes in SpEd placement later on. Mechanisms for these students are likely similar to the mechanisms discussed in Section 5.1.2. These include positive peer-to-peer spillover effects, improvements in teacher attention and instructional quality, and a reduction in racial bias felt by all students with the implementation of the disproportionality cap.

Finally, it is possible that a reduction in SpEd support services within the GE classroom for Black students could have impacted students of other races as well. We turn to investigating the effect of the Black disproportionality cap on White and Hispanic students, proportionally the next largest racial groups (Table 7). We now incorporate all three treatment variables additively into our model.[37] Overall, the largest impacts of the policy are experienced by Black GE students. We do not find any impact of the cap on White GE students (column 3), and we find a statistically significant impact only on high school completion for Hispanic GE students (column 2). This positive impact is likely driven by similar positive peer-to-peer spillovers as those experienced by Black GE students.

[36]This also holds up to concerns of multiple inference. We find a statistically significant positive impact on a summary index of long-run outcomes for Black GE students (available upon request).

[37]Appendix Figure A.3 illustrates a lack of correlation between the 2004 district-level Black and Hispanic disproportionality rates, motivating why they are incorporated additively in our model.

5.3 Hispanic Disproportionality

We now turn to the effects of the Hispanic disproportionality cap on Hispanic SpEd students. As previously noted, the statewide district-level average Hispanic disproportionality rate was already below the 1% threshold in 2004, at about -2.9%. Thus, Hispanic students were, on average, under-represented in SpEd prior to the caps on over-representation. Therefore, we do not anticipate finding much impact of the Hispanic disproportionality cap. Indeed, Appendix Figure A.4 illustrates, on the whole, a lack of differential trends during the pre- *and* post-policy periods. Appendix Table A.12 presents the impact of the Hispanic disproportionality cap on 5th grade Hispanic SpEd students. We do not find a statistically significant effect on the likelihood of SpEd participation in 9th grade. In the long-run, our point estimate suggests a decrease in the likelihood of high school completion by 0.44 p.p. Since the Hispanic disproportionality rate is already below the 1% threshold, the effect size for students at the average district would be even smaller. However, we do not find a significant impact on college enrollment. Furthermore, the effect on a summary index of high school completion and college enrollment is not statistically significant. Ultimately, given a lack of clear pattern in the event study in Figure A.4, as well as a lack of significance in the long-run summary index, we conclude that we are not able to identify the impacts of an over-representation of Hispanic SpEd students in this context.

Turning to Hispanic GE students, Appendix Figure A.5 presents event study estimates for the effect of the Hispanic disproportionality cap. We find a modest post-policy decline in 9th grade SpEd participation, statistically significant only for students with the greatest years of exposure. In column (2) of Table 7 we find a 0.18 p.p. decline in SpEd participation. Although districts did not remove Hispanic students from SpEd in response to the Hispanic disproportionality cap, we find that they felt at least some (albeit small) pressure to reduce the rate at which students were newly identified for SpEd. Intuitively, it makes sense that districts already meeting the Hispanic disproportionality threshold, who feel relatively less pressure to reduce SpEd enrollments, may reduce the rate at which they newly identify students for SpEd, but not need to remove existing students from SpEd. In the long-run, we do not find any statistically significant impacts of the Hispanic disproportionality cap on high school completion or college enrollment for Hispanic GE students.

Again, we investigate spillover effects on other races in Table 7. The Hispanic disproportionality cap had negative impacts on high school completion for White GE students, as well as negative impacts on college enrollment for Black and White GE students. Although we cannot directly test why the Hispanic disproportionality cap had negative impacts on Black and White students, if the Hispanic GE students less likely to be in SpEd at 9th grade are students who would have benefited from SpEd (since they are under-represented in SpEd at baseline, unlike Black students on average), then these students may be experiencing greater challenges in the classroom, leading to more disruptive behaviors and thus negative spillover effects on other students in the classroom. Overall, given a lack of impact on Hispanic SpEd students and the under-representation of Hispanic students in SpEd at baseline, we leave for future work any further potential explorations of the Hispanic disproportionality cap.

5.4 SpEd Enrollment Cap

Throughout this paper, we have controlled for the SpEd enrollment cap. As discussed in Appendix B, Table B.1 demonstrates that controlling for the SpEd enrollment cap does not significantly affect our estimates of the disproportionality cap. However, we control for the SpEd cap throughout given the fact that it was an important policy change introduced at the same time as the disproportionality caps. In Ballis and Heath (2021), we investigate in-depth the effects of the SpEd enrollment cap and find significant reductions in high school completion and college enrollment, which are driven by non-White and low-income students. Given our current focus on race, we turn to investigating any differences across Black and Hispanic students separately in response to the SpEd enrollment cap.

Black Students- In the second row of Table 2 column (3), we find that the the likelihood of continuing in SpEd at 9th grade fell by 2.1 p.p. (2.7%) for Black SpEd students exposed to the SpEd enrollment cap in every year between 5th and 9th grade. Again, we scale the coefficients to give an effect size for students at the average district, which was 3.3 p.p. above the 8.5% SpEd enrollment threshold in 2004. Despite the reductions in SpEd participation, we do not find statistically significant impacts of the SpEd enrollment cap on Black SpEd student's high school completion or college enrollment. For Black GE students, Table 7 illustrates that the SpEd enrollment cap led to a 0.96 p.p. (21%) decrease in SpEd participation at 9th grade, but again did not have a statistically significant impact on long-run outcomes.

This leads to the question of why the Black disproportionality cap had a statistically significant positive impact on Black students' long-run outcomes, but the SpEd enrollment cap did not. To investigate this, we turn back to our theoretical framework. It could be the case that districts with high disproportionality rates have greater racial bias, and thus have higher γ_s. In this case, there will be many more Black students in SpEd for whom $a_c > c$, implying that they do not benefit from SpEd. In contrast, districts relatively more impacted by the SpEd enrollment cap may have lower racial bias, and thus fewer students who are misclassified for SpEd.

Hispanic Students- Turning to Hispanic SpEd students, in column (7) of Appendix Table A.2 we find that the SpEd enrollment cap led to reductions in the likelihood of continuing in SpEd by 2.4 p.p. (3.1%). In the long-run, the SpEd cap reduced the likelihood of high school completion by 1.9 p.p. (3.2%) and college enrollment by 1.7 p.p. (6.0%). For Hispanic GE students in column (2) of Table 7, we find a 0.72 p.p. (22%) decline in SpEd participation, a 1.2 p.p. (1.8%) decline in high school completion, and a 1.2 p.p. (2.4%) decline in college enrollment. The SpEd enrollment cap worsened long-run outcomes for Hispanic SpEd and GE students, likely as a result of reducing SpEd enrollment for Hispanic students who needed services. From this analysis, we conclude that the negative impacts of the SpEd enrollment cap explored in Ballis and Heath (2021) were largely driven by Hispanic students. We expand on the comparison and reconciliation of the effects found in these two papers in Appendix B.

6 Conclusion

Under the Performance Based Monitoring Analysis System (PBMAS) introduced in 2004, Texas capped Black and Hispanic disproportionality rates, that is, the percent of Black and Hispanic students in SpEd relative to the percent of Black and Hispanic students in the district. These district-level caps allow us to quantify causal estimates of the effect of reducing disproportionality on long-run outcomes. We use cross-cohort and cross-district variation in how far districts were from meeting the cutoffs before PBMAS in a dose-response difference-in-differences estimation framework. When the policy went into effect in 2004, it impacted districts differentially based on their pre-treatment disproportionality rates.

Throughout the paper, we have focused primarily on Black students. Hispanic students

were under-represented in SpEd at baseline and the cap on over-representation did not have a significant impact on SpEd participation. For Black students, we find that the Black disproportionality cap led to meaningful reductions in the likelihood of receiving SpEd services and improvements in high school completion and college enrollment for those in SpEd and GE.

We explore several potential mechanisms behind the positive effects of the Black disproportionality cap on Black students. For Black SLD students removed from SpEd as a result of the policy, we find evidence consistent with a story of misclassification. In particular, we find that the positive impacts are larger in districts that may be more likely to misidentify Black students for SpEd, that is, those that over-classify Black students in SpEd relative to White students, those with low shares of Black students, and those with fewer Black teachers. We also find that Black students removed from SpEd post-policy are relatively higher performing with more mild disability types at baseline, compared to the students who lose SpEd pre-policy. Indirect effects of the policy for Black SpEd and GE students also play an important role. These effects are likely driven by a combination of factors including changes in resource allocation, teaching practices, peer-to-peer spillovers, and racial bias.

When considering the external validity of our estimates, it is important to interpret the impacts of the disproportionality cap within the broader context of Texas' reduction in SpEd enrollment. Our prior work showed that capping overall SpEd enrollment led to significant negative impacts on SpEd student outcomes. Although we show in Appendix B that our estimates are robust to the simultaneity of this cap, in theory our estimates of the disproportionality cap could still be biased downward. Given that we are finding positive impacts, if anything this would only attenuate our estimates. Thus, we hypothesize that if similar disproportionality caps were implemented in other states without a comparable reduction in SpEd enrollment the impacts could be even more positive. Furthermore, it is also important to interpret our results within the broader context of race in Texas. The extent to which Black students are over-identified in SpEd and Hispanic students are under-identified in SpEd likely varies across states, particularly depending on the level of implicit or explicit racial bias. Racial bias is difficult to quantify, making it challenging to determine how Texas' school system compares to other states. Ultimately, given Texas' size, data availability, and the uniqueness of this policy, this context provides a rare opportunity to obtain causal estimates of the impacts of limiting disproportionality in SpEd, an important topic nationwide.

Our findings have meaningful implications for all public school students. Students who require SpEd services greatly benefit from them in the long-run. However, those who are misclassified for SpEd can be significantly harmed in the long-run. Whether students are appropriately identified for SpEd has important long-run implications for all students in the classroom. SpEd is an intensive and costly intervention, and it is important both to schools and students that individuals be appropriately placed in SpEd. Importantly, we caution against the interpretation that capping Black disproportionality is the best policy intervention, and instead point to the importance of considering the eligibility criteria for SpEd services, particularly for Black students, to ensure that all students are appropriately classified for SpEd. Finally, our results suggest that future research on the impact disproportionality policies may have on broader changes in racial bias could be especially important.

References

Aisch, G., Gebeloff, R., & Quealy, K. (2014). Where We Came From and Where We Went, State by State. *New York Times*.

Artiles, A., Harry, B., Reschly, D., & Chinn, P. (2002). Over-Identification of Students of Color in Special Education: A critical overview. *Multicultural Perspectives*, *4*(1), 3–10.

Ballis, B., & Heath, K. (2021). The Long-Run Impacts of Special Education. *American Economic Journal: Economic Policy*, *13*(4), 72–111.

Baron, E. J., Hyman, J. M., & Vasquez, B. N. (2022). Public School Funding, School Quality, and Adult Crime. *National Bureau of Economic Research, Working Paper No. 29855*.

Bayer, P., & Charles, K. K. (2018). Divergent Paths: A New Perspective on Earnings Differences Between Black and White Men Since 1940. *The Quarterly Journal of Economics*, *133*(3), 1459–1501.

Blackorby, J., & Cameto, R. (2004). Changes in School Engagement and Academic Performance of Students with Disabilities. *Wave 1 Wave 2 Overview*.

Bound, J., Schoenbaum, M., & Waidmann, T. (1995). Race and Education Differences in Disability Status and Labor Force Attachment in the Health and Retirement Survey. *Journal of Human Resources*, *30*, 227–267.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics*, *225*(2), 200–230.

Cameron, S., & Heckman, J. (2001). The Dynamics of Educational Attainment for Black, Hispanic, and White Males. *Journal of Political Economy*, *109*(3), 455–499.

Card, D., & Rothstein, J. (2007). Racial Segregation and the Black-White Test Score Gap. *Journal of Public Economics*, *91*(11-12), 2158–2184.

Carrell, S. E., Hoekstra, M., & Kuka, E. (2018, November). The long-run effects of disruptive peers. *American Economic Review*, *108*(11), 3377–3415. Retrieved from `https://www.aeaweb.org/articles?id=10.1257/aer.20160763` doi: 10.1257/aer.20160763

Center for Parent Information & Resources. (2022). Evaluating School-Aged Children for Disability. Retrieved from `https://www.parentcenterhub.org/evaluation/`

Chetty, R., Hendren, N., Jones, M. R., & Porter, S. R. (2019). Race and Economic Opportunity in the United States: an Intergenerational Perspective. *The Quarterly Journal of Economics*, *135*(2), 711–783.

Cohen, J. (2007). Causes and Consequences of Special Education Placement: Evidence from Chicago Public Schools. *Brookings Institution Working Paper*.

Cutler, D., & Vogl, T. (2012). Socioeconomic Status and Health: Dimensions and Mechanisms. *Oxford Handbook of Health Economics*, 124–163.

Dee, T. (2005). A Teacher Like Me: Does Race, Ethnicity, or Gender Matter? *American Economic Review*, *95*(2), 158–165.

de la Cruz, R. E. (1996). Assessment-Bias Issues in Special Education: A Review of the Literature.

Dever, B. V., Raines, T. C., Dowdy, E., & Hostutler, C. (2016). Addressing Disproportionality in Special Education Using a Universal Screening Approach. *Journal of Negro Education*, *85*(1), 59–71.

Dhuey, E., & Lipscomb, S. (2011). Funding Special Education by Capitation: Evidence From State Finance Reforms. *Education Finance and Policy*, 168–201.

Donovan, S., & Cross, C. (2002). Minority Students in Special and Gifted Education. *National Academy Press*.

Elder, T., Figlio, D., Imberman, S., & Persico, C. (2021). School Segregation and Racial Gaps in Special Education Identification. *Journal of Labor Economics*.

Exceptional Lives. (2025). What are Special Education Services? Retrieved from `https://exceptionallives.org/guides/what-are-special-education-services/`

Fish, R. E. (2019). Standing Out and Sorting In: Exploring the Role of Racial Composition in Racial Disparities in Special Education. *American Educational Research Journal*, *56*(6), 2573–2608.

Gasper, J., DeLuca, S., & Estacion, A. (2012). Switching Schools: Revisiting the Relationship Between School Mobility and High School Dropout. *American Educational Research Journal*, *49*(3), 487–519.

Gershenson, S., Holt, S., & Papageorge, N. (2016). Who believes in me? The effect of student-teacher demographic match on teacher expectations. *Economics of Education Review*, *52*, 209–224.

Gershenson, S., & Papageroge, N. (2018). The Power of Teacher Expectations: How racial bias hinders student attainment. *Education Next*, *18*(1), 64–70.

Gordon, N. (2017). Race, Poverty, and Interpreting Overrepresentation in Special Education. *Brookings Report*.

Hanushek, E., Kain, J., & Rivkin, S. (2002). Inferring Program Effects for Special Populations: Does Special Education Raise Achievement for Students with Disabilities? *The Review of Economics and Statistics*, *84*(4), 584–599.

Hanushek, E., Kain, J., & Rivkin, S. (2004). Disruption versus Tiebout improvement: the costs and benefits of switching schools. *Journal of Public Economics*, *88*, 1721–1746.

Harrison, J. R., Bunford, N., Evans, S. W., & Owens, J. S. (2013). Educational Accommodations for Students with Behavioral Challenges: A Systematic Review of the Literature. *Review of Educational Research*, *83*(4).

Hart, C. M., & Linday, C. A. (2024). Teacher-Student Race Match and Identification for Discretionary Educational Services. *American Educational Reserach Journal*, *61*(3), 474–507.

Heckman, J. J., & Karapakula, G. (2019). The Perry Preschoolers at Late Midlife: A Study in Design-Specific Inference. *National Bureau of Economic Research, Working Paper No. 25888*.

Hibel, J., Farkas, G., & Morgan, P. (2010). Who is Placed into Special Education? *Sociology of Education*, *83*(4), 312–332.

Hosp, J., & Reschly, D. (2003). Referral Rates for Intervention or Assessment: A Meta-Analysis of Racial Differences. *Journal of Special Education*, *37*(2), 67–80.

Jencks, C., & Phillips, M. (1998). The Black-White Test Score Gap: An Introduction. *The Black-White Test Score Gap*, 1–51.

Kling, J., Liebman, J., & Katz, L. (2007). Experimental Analysis of Neighborhood Effects. *Econometrica*, *75*(1), 83–119.

Kolbe, T., Dhuey, E., & Doutre, S. M. (2023). Unequal and Increasingly Unfair: How Federal Policy Creates Disparities in Special Education Funding. *Exceptional Children*,

0(0).

Leive, A., & Christopher, R. (2022). Education Gradients in Mortality Trends by Gender and Race. *Journal of Human Capital*, *16*(1).

Luna, R. E. (2024). What Is Specific Learning Disorder. Retrieved from `https://www.psychiatry.org/patients-families/specific-learning-disorder/what-is-specific-learning-disorder`

Miles, K. H., & Darling-Hammond, L. (1998). Rethinking the Allocation of Teaching Resources: Some lessons from high-performing schools. *Educational Evaluation and Policy Analysis*, *20*(1), 9–29.

Morgan, H. (2020). Misunderstood and Mistreated: Students of Color in Special Education. *Voices of Reform: Educational Research to Inform and Reform*, *3*(2).

Morgan, P., Farkas, G., Cook, M., Strassfel, N., Hillemeier, M., Pun, W. H., & Schussler, D. (2017). Are Black Children Disproportionately Overrepresented in Special Education? A Best-Evidence Synthesis. *Exceptional Children*, *83*(2), 181–198.

Morgan, P., Farkas, G., Hillemeier, M., & Maczuga, S. (2017). Replicated Evidence of Racial and Ethnic Disparities in Disability Identification in U.S. Schools. *Educational Researcher*, *46*(6), 305–322.

Morgan, P., Hammer, C. S., Farkas, G., Hillemeier, M., Maczuga, S., Cook, M., & Morano, S. (2016). Who Receives Speech/Language Services by 5 Years of Age in the United States? *American Journal of Speech-Language Pathology*, *25*, 183–199.

Mountjoy, J. (2022). Community Colleges and Upward Mobility. *American Economic Review*, *112*(8), 2580–2630.

Office of Special Education and Rehabilitative Services. (2009). Questions and Answers on Disproportionality. Retrieved from `https://www2.ed.gov/policy/speced/guid/idea/disproportionality-q-a.pdf`

Oswald, D., Coutinho, M., Best, A., & Singh, N. (1999). Ethnic Representation in Special Education: The Influence of School-Related Economic and Demographic Variables. *Journal of Special Education*, *32*(4), 194–206.

Podell, D., & Soodak, L. (1993). Teacher Efficacy and Bias in Special Education Referrals. *The Journal of Educational Research*, *86*(4), 247–253.

Prenovitz, S. (2017). Accountability Incentives and Special Education. *Unpublished working paper*.

Prieto, A., & Zucker, S. (1981). Teacher Perception of Race as a Factor in the Placement of Behaviorally Disordered Children. *Behavioral Disorders*, *7*(1), 34–38.

Rawe, J. (2024). The difference between IEPs and 504 plans. *Understood.org*. Retrieved from `https://www.understood.org/en/articles/the-difference-between-ieps-and-504-plans`

Reardon, S. (2016). School Segregation and Racial Academic Achievement Gaps. *The Russell Sage Foundation Journal of the Social Sciences*, *2*(5), 34–57.

Rose, E., & Huefner, D. S. (1984). Cultural Bias in Special Education Assessment and Placement.

Rosenthal, B. (2016). Denied: How Texas Keeps Tens of Thousands of Children Out of Special Education. *Houston Chronicle*.

Rothstein, J., & Schanzenbach, D. W. (2022). Does Money Still Matter? Attainment and Earnings Effects of Post-1990 School Finance Reforms. *Journal of Labor Economics*, *40*(51), 5141–5178.

Rothstein, J., & Wozny, N. (2013). Permanent Income and the Black-White Test Score Gap. *Journal of Human Resources*, *95*(5), 1468–1497.

Sabine, G., Sabine, S.-K., & Pit-ten Cate, I. M. (2015). Are School Placement Recommendations Accurate? The effect of students' ethnicity on teachers' judgments and recognition memory. *European Journal of Psychology of Education*, *30*(2), 169–188.

Sallin, A. (2021). Estimating Returns to Special Education: Combining machine learning and text analysis to address confounding. *Working Paper*.

Sant'Anna, P. H., & Zhao, J. (2020). Doubly Robust Difference-in-Differences Estimators. *Journal of Econometrics*, *219*(1), 101–122.

Santiago-Rosario, M. R., Whitcomb, S. A., Pearlman, J., & McIntosh, K. (2021). Associations Between Teacher Expectations and Racial Disproportionality in Discipline Referrals. *Journal of School Psychology*, *85*, 80–93.

Schwartz, A. E., Hopkins, B. G., & Stiefel, L. (2021). The Effects of Special Education on the Academic Performance of Students with Learning Disabilities. *Journal of Policy Analysis and Management*, *40*(2), 480-520.

Shifrer, D. (2013). Stigma of a Label: Educational Expectations for High School Students Labeled with Learning Disabilities. *Journal of Health and Social Behavior*, *54*(4), 462–480.

Shifrer, D., Muller, C., & Callahan, R. (2011). Disproportionality and Learning Disabilities: Parsing Apart Race, Socioeconomic Status, and Language. *Journal of Learning Disabilities*, *44*(3), 246–257.

Skiba, R. J., Simmons, A. B., Ritter, S., Gibb, A. C., Rausch, M. K., Cuadrado, J., & Chung, C.-G. (2008). Achieving Equity in Special Education: History, Status, and Current Challenges. *Exceptional Children*, *74*(3), 264–288.

Stiefel, L., Fatima, S. S., & Cimpian, J. R. (2024). The Role of School Context in Explaining Racial Disproportionality in Special Education. *Educational Evaluation and Policy Analysis*.

Texas Education Agency. (2004). Performance Based Monitoring Analysis System 2004-2005 Manual. Retrieved from `https://tea.texas.gov/student-assessment/monitoring-and-interventions/rda/pbmas-pbm/0405pbmasmanual.pdf`

Texas Education Agency. (2025). Disproportionate Representation in Special Education. Retrieved from `https://tea.texas.gov/academics/special-student-populations/special-education/data-and-reports/disproportionate-representation-in-special-education`

Tobias, S., Cole, C., Zibrin, M., & Bodlakova, V. (1982). Teacher-Student Ethnicity and Recommendations for Special Education Referrals. *Journal of Education Psychology*, *74*(1), 72–76.

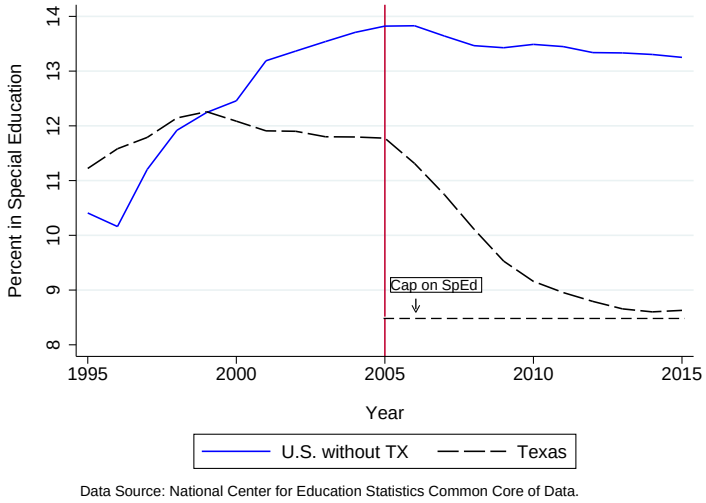
Tobias, S., Zibrin, M., & Menell, C. (1983). Special Education Referrals: Failure to Replicate Student-Teacher Ethnicity Interaction. *Journal of Education Psychology*, *75*(5), 705–707.

Welsh, R. (2017). School Hopscotch: A Comprehensive Review of K-12 Student Mobility in the United States. *Review of Educational Research*, *87*(3), 475–511.

Figures and Tables

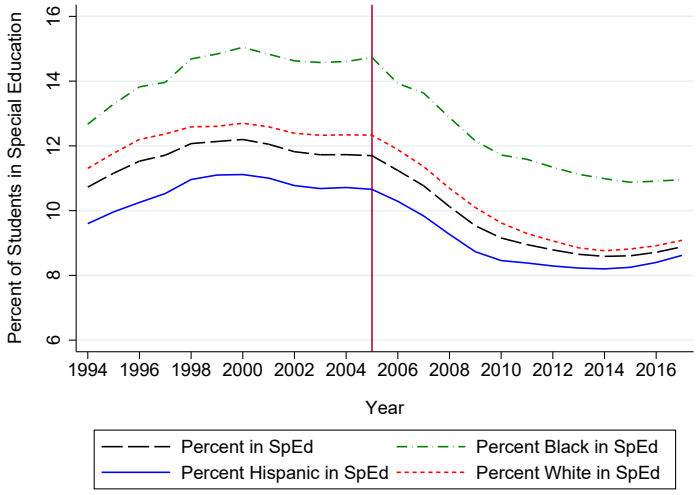
Figure 1

(a) Special Education Rate



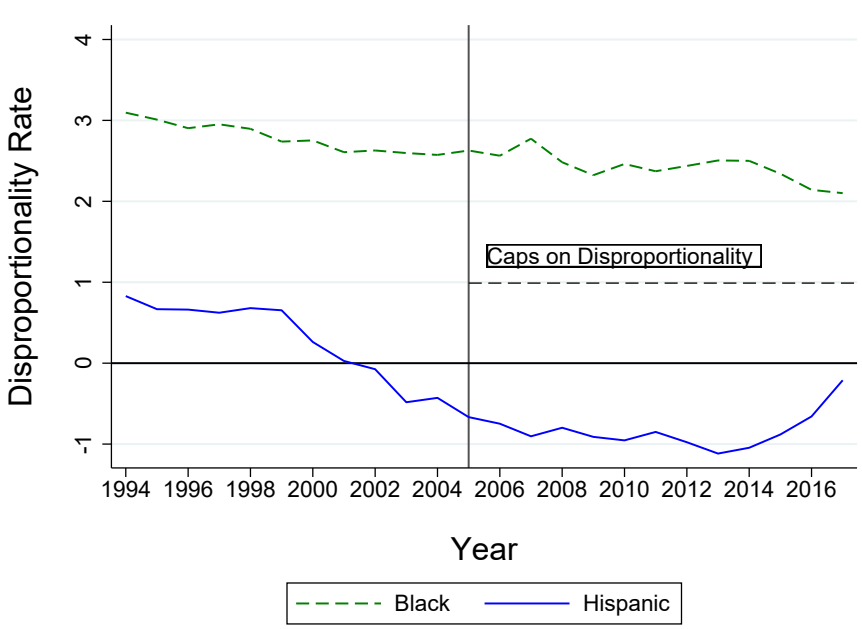
Data Source: National Center for Education Statistics Common Core of Data.

(b) Special Education Rate by Race



In figure (a), averages represent statewide population averages, that is, the number of students in a state in special education divided by the total number of students in that state. Figure (b) plots the percent of students in special education in Texas by race.

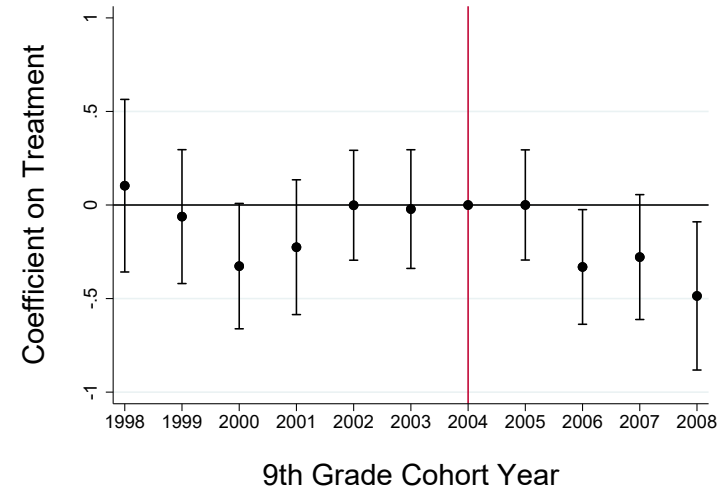
Figure 2 Disproportionality Rate for Black and Hispanic Students



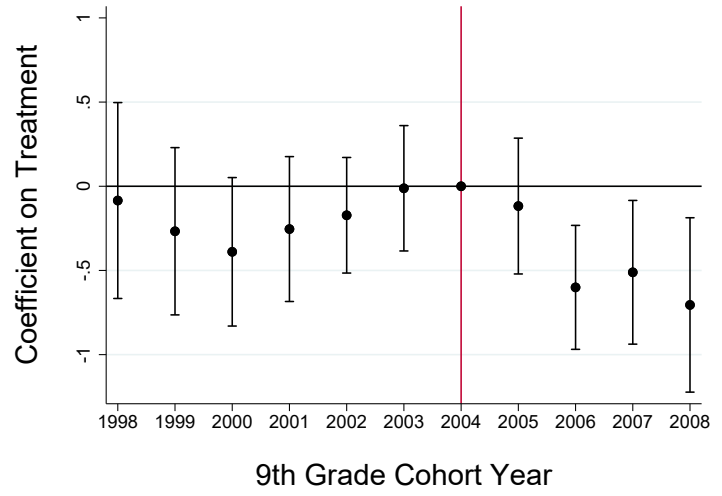
The disproportionality rate is measured as the percent of Black or Hispanic students in special education minus the percent of Black or Hispanic students in a given district.

Figure 3 Event Study Estimates of the Black Disproportionality Cap for Black Special Education Students

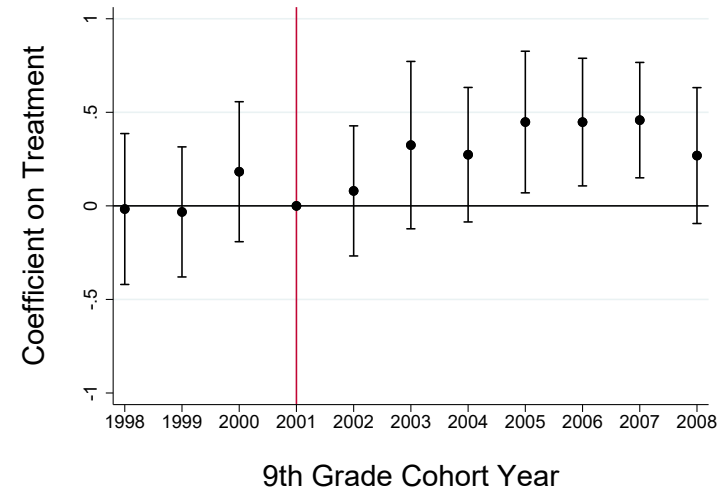
(a) All Students G9 SpEd Status



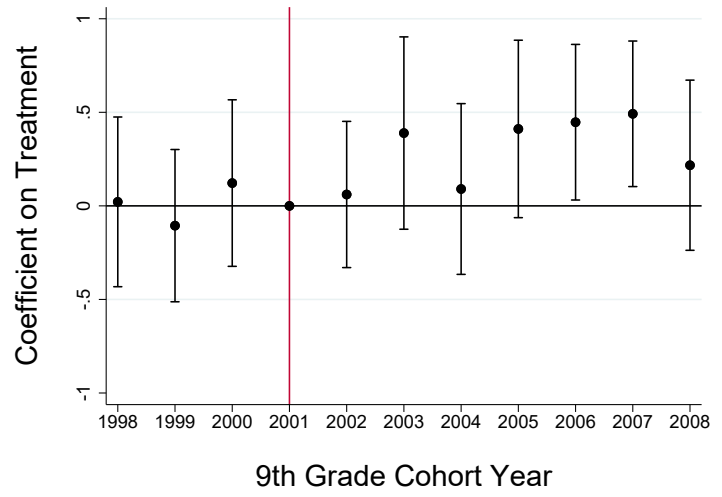
(b) SLD Students G9 SpEd Status



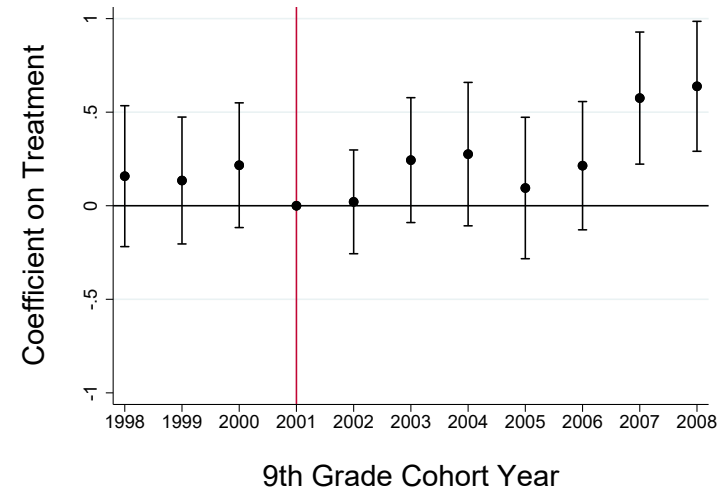
(c) All Students HS Completion



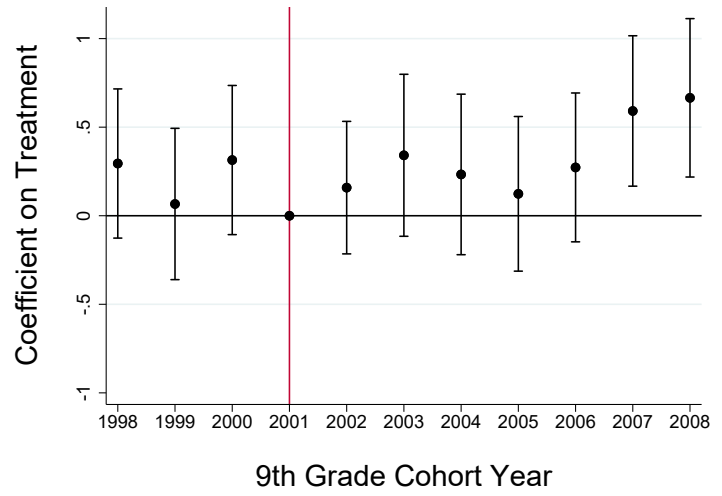
(d) SLD Students HS Completion



(e) All Students College Enrollment



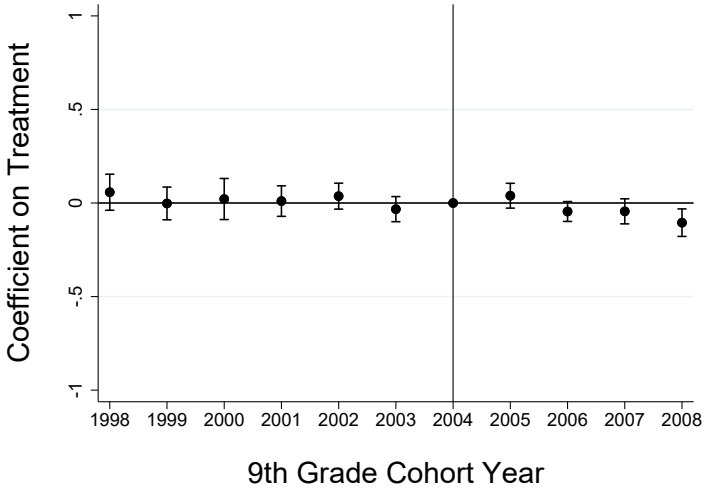
(f) SLD Students College Enrollment



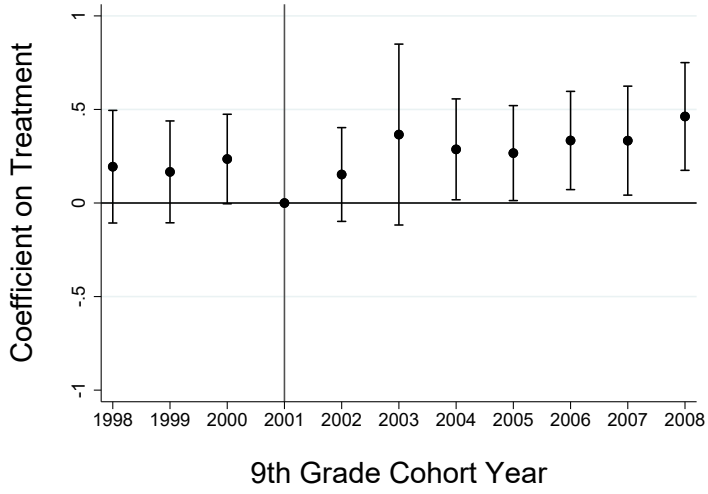
Points represent district-level Black disproportionality rate in 2004 interacted with indicators for each 9th grade cohort year. Y-axis ranges from -1 to 1 percentage point. The vertical bars denote the 95% confidence intervals. See Table 2 for full set of controls.

Figure 4 Event Study Estimates of the Black Disproportionality Cap for Black General Education Students

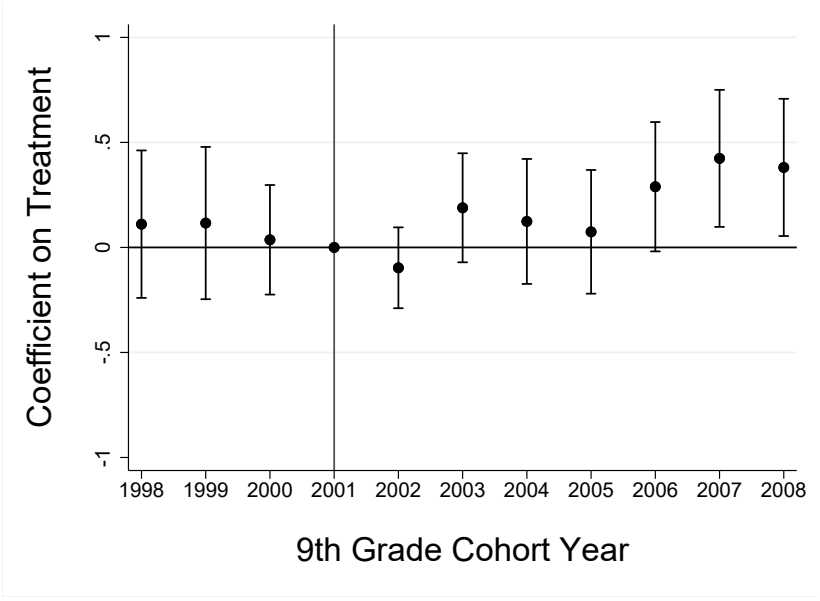
(a) Grade 9 Special Education Status



(b) High School Completion



(c) College Enrollment



Points represent district-level Black disproportionality rate in 2004 interacted with indicators for each 9th grade cohort year. Y-axis ranges from -1 to 1 percentage point. The vertical bars denote the 95% confidence intervals. See Table 7 for full set of controls.

Table 1 Descriptive Statistics for 5th Grade Cohorts between 1994 to 2004

	All Students			SpEd Students		
	All Races	Black Students	Hispanic Students	All Races	Black Students	Hispanic Students
	(1)	(2)	(3)	(4)	(5)	(6)
Covariates						
Male	0.512	0.506	0.512	0.659	0.658	0.659
FRL	0.522	0.704	0.798	0.632	0.803	0.837
ESL	0.036	0.003	0.080	0.039	0.002	0.093
Gifted	0.104	0.067	0.073	0.014	0.006	0.008
Title I	0.545	0.636	0.768	0.564	0.626	0.751
White	0.439	·	·	0.418	·	·
Black	0.142	·	·	0.185	·	·
Hispanic	0.394	·	·	0.387	·	·
Other	0.026	·	·	0.012	·	·
Took G5 Math Exam	0.808	0.762	0.752	0.406	0.277	0.352
Took G5 Reading Exam	0.800	0.755	0.740	0.357	0.240	0.291
Math G5 Z-score	0.034	-0.435	-0.152	-0.678	-1.189	-0.926
Reading G5 Z-score	0.027	-0.333	-0.234	-0.687	-1.102	-1.022
G5 SpEd Rate	0.143	0.185	0.140	·	·	·
Malleable Disability	·	·	·	0.911	0.879	0.918
$\geq$ 50% of day in GE Class	·	·	·	0.857	0.792	0.869
Long-run Outcomes						
High School Diploma	0.729	0.676	0.676	0.633	0.602	0.597
Enroll College	0.557	0.512	0.466	0.338	0.312	0.281
Enroll 2 year College	0.473	0.404	0.407	0.310	0.280	0.263
Enroll 4 year College	0.114	0.136	0.078	0.038	0.040	0.024
N	2,877,992	409,276	1,134,932	410,191	75,725	158,855
Disability Type						
Learning Disability				0.629	0.607	0.699
Speech Impairment				0.128	0.106	0.115
Emotional Disturbance				0.073	0.920	0.524
Intellectual Disability				0.051	0.889	0.479
Other Health Impairment				0.080	0.721	0.497
Autism				0.112	0.115	0.006
Auditory Impairment				0.011	0.008	0.001
Orthopedic Impairment				0.011	0.008	0.119
Visual Impairment				0.005	0.004	0.004
Traumatic Brain Injury				0.001	0.002	0.001
Deaf/Blind				0.0002	0.0003	0.0002

Numbers represent the proportion of students in each demographic category, on a 0 to 1 scale. FRL is an indicator for receiving free or reduced-price lunch. ESL is an indicator for participation in the English as a Second Language program. Gifted is a separately defined category from Special Education in Texas, and is a program for high achieving students. Malleable Disability refers to students with learning disabilities, speech impairments, other health impairments, and emotional disturbance. High School diploma is measured within 2 years of expected high school graduation, and conditional on being observed in the data in grade 9. College enrollment is measured 6 years after expected high school graduation, and is not conditional on high school diploma.

Table 2 Effect of Policy on **Black Special Education Students**

	All Black Students			SLD	Non-SLD
SpEd Status G9	(1)	(2)	(3)	(4)	(5)
$Disp_{d,2004} \times FracExpo$	-0.174	-0.331**	-0.376**	-0.548***	0.022
	(0.160)	(0.167)	(0.162)	(0.198)	(0.224)
$SpEd_{d,2004} \times FracExpo$	-0.743***	-0.716***	-0.627**	-1.308***	-0.189
	(0.264)	(0.251)	(0.250)	(0.296)	(0.215)
Mean Dept Var	0.774	0.774	0.774	0.807	0.722
High School Completion					
$Disp_{d,2004} \times FracExpo$	0.339**	0.423***	0.404***	0.425**	0.190
	(0.162)	(0.162)	(0.155)	(0.185)	(0.214)
$SpEd_{d,2004} \times FracExpo$	-0.008	0.144	0.127	0.111	0.058
	(0.194)	(0.190)	(0.204)	(0.345)	(0.188)
Mean Dept Var	0.602	0.602	0.602	0.614	0.584
College Enrollment					
$Disp_{d,2004} \times FracExpo$	0.341***	0.478***	0.476***	0.443***	0.294
	(0.122)	(0.130)	(0.131)	(0.157)	(0.224)
$SpEd_{d,2004} \times FracExpo$	0.195	0.294	0.170	0.390	-0.081
	(0.247)	(0.252)	(0.234)	(0.309)	(0.233)
Mean Dept Var	0.312	0.312	0.312	0.311	0.314
Observations	75,725	75,725	75,725	46,218	29,507
Individual Controls		X	X	X	X
District-Cohort Controls			X	X	X

*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors are clustered at the district level. All specifications include cohort fixed effects and district fixed effects. Regressions are run on students in SpEd as of 5th grade from 1994 (when the data begins) to 2004 (the year prior to policy implementation). SpEd status is measured 4 years after 5th grade, to correspond to expected 9th grade. Individual-level controls include disability type, classroom setting, ESL, FRL, Title I, and gifted status as of 5th grade. District-cohort level controls include gender, race, ESL, FRL, Title I, and gifted composition. High school diploma and college enrollment are conditional on being observed in Texas public schools as of 9th grade. Long-run outcomes are censored such that individuals have 2 years after expected high school completion to earn a high school diploma and 6 years after expected high school completion to enroll in college. Column (4) is estimated for Black SpEd students with Specific Learning Disabilities (SLD) as of 5th grade and column (5) is estimated for all other Black SpEd students as of 5th grade.

Table 3 Heterogeneous Effects of the Policy on **Black Special Education Students**

	Male	Female	FRL	Non-FRL
SpEd Status	(1)	(2)	(3)	(4)
$Disp_{d,2004} \times FracExpo$	-0.282*	-0.595**	-0.396**	-0.276
	(0.159)	(0.233)	(0.160)	(0.281)
$SpEd_{d,2004} \times FracExpo$	-0.562**	-0.743**	-0.715***	-0.425
	(0.250)	(0.330)	(0.260)	(0.344)
Mean Dept Var	0.787	0.748	0.791	0.703
High School Completion				
$Disp_{d,2004} \times FracExpo$	0.666***	-0.134	0.429**	0.142
	(0.184)	(0.201)	(0.173)	(0.278)
$SpEd_{d,2004} \times FracExpo$	0.207	-0.012	0.223	-0.588
	(0.266)	(0.232)	(0.240)	(0.427)
Mean Dept Var	0.584	0.636	0.577	0.704
College Enrollment				
$Disp_{d,2004} \times FracExpo$	0.580***	0.216	0.351**	0.986***
	(0.150)	(0.206)	(0.141)	(0.342)
$SpEd_{d,2004} \times FracExpo$	0.321	-0.062	0.161	0.458
	(0.252)	(0.286)	(0.232)	(0.481)
Mean Dept Var	0.293	0.350	0.276	0.462
Observations	49,817	25,908	60,791	14,934

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full set of controls. Gender and Free and Reduced-Price Lunch (FRL) are measured as of 5th grade.

Table 4 Effect of Policy on **Black Students with Specific Learning Disabilities** by Representation and Racial Composition

	Blinder-Oaxaca Decomp		Racial Composition	
	Cond Over-rep	Cond Under-rep	Below 20%	Above 80%
SpEd Status	(1)	(2)	(3)	(4)
$Disp_{d,2004} \times FracExpo$	-0.735***	-0.440	-0.813**	-0.022
	(0.254)	(0.299)	(0.364)	(3.302)
$SpEd_{d,2004} \times FracExpo$	-0.963**	-1.148***	-0.369	-1.399
	(0.436)	(0.408)	(0.374)	(1.223)
Mean Dept Var	0.831	0.796	0.802	0.682
High School Completion				
$Disp_{d,2004} \times FracExpo$	0.506*	0.327	0.849*	-7.280**
	(0.277)	(0.238)	(0.439)	(2.686)
$SpEd_{d,2004} \times FracExpo$	-0.554	0.554	0.173	-2.210
	(0.527)	(0.452)	(0.576)	(2.163)
Mean Dept Var	0.658	0.592	0.657	0.472
College Enrollment				
$Disp_{d,2004} \times FracExpo$	0.506***	0.336	0.913**	-0.395
	(0.190)	(0.270)	(0.436)	(2.205)
$SpEd_{d,2004} \times FracExpo$	0.267	0.408	-0.098	-0.065
	(0.487)	(0.441)	(0.433)	(1.048)
Mean Dept Var	0.330	0.302	0.342	0.249
Observations	14,850	31,466	17,661	721

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full set of controls. "Cond Over-rep" implies Black students are more likely to be in SpEd relative to observationally-equivalent White peers. Likewise, "Cond Under-rep" implies under-representation in SpEd relative to White peers. "Below 20%" refers to districts with fewer than 20% Black students and "Above 80%" is districts with greater than 80% Black students.

Table 5 Effect of Policy on **Black Students with Specific Learning Disabilities** by Racial Composition and Teacher Experience

	Racial Composition		**Experience**	
	Above 90th Pctile	Below 90th Pctile	Above Average	Below Average
SpEd Status	(1)	(2)	(3)	(4)
$Disp_{d,2004} \times FracExpo$	-2.571***	-0.600***	-0.786***	-0.541
	(0.942)	(0.207)	(0.207)	(0.445)
$SpEd_{d,2004} \times FracExpo$	-2.139**	-0.815**	-1.636***	-1.039**
	(0.916)	(0.344)	(0.362)	(0.481)
Mean Dept Var	0.807	0.807	0.842	0.763
High School Completion				
$Disp_{d,2004} \times FracExpo$	-1.695	0.436**	0.269	0.552*
	(1.131)	(0.194)	(0.236)	(0.317)
$SpEd_{d,2004} \times FracExpo$	-0.509	0.283	-0.007	-0.174
	(1.743)	(0.370)	(0.436)	(0.565)
Mean Dept Var	0.489	0.645	0.607	0.622
College Enrollment				
$Disp_{d,2004} \times FracExpo$	-1.595	0.454***	0.581***	0.424
	(0.988)	(0.157)	(0.187)	(0.321)
$SpEd_{d,2004} \times FracExpo$	-1.518	0.473	0.623*	0.198
	(1.352)	(0.360)	(0.376)	(0.529)
Mean Dept Var	0.226	0.333	0.276	0.355
Observations	9,351	36,867	25,706	20,512

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full set of controls. Black students are split by whether the district-level composition of Black teachers is above or below the 90th percentile (i.e., 39.5%) in columns (1) and (1); and whether their district-level average teacher experience is above or below the statewide average teacher experience of 11.7 years in columns (3) and (4).

Table 6 District-Level Changes in the Composition of Black Students Who Lose Special Education

	Male	**ESL**	**FRL**	**Took Math**	**Took Reading**	**Math Score**	**Reading Score**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$Disp_{d,2004} \times FracExpo$	-0.473	0.005	-0.014	0.409	0.335	-0.590	1.120**
	(0.293)	(0.010)	(0.225)	(0.320)	(0.312)	(0.389)	(0.540)
$SpEd_{d,2004} \times FracExpo$	0.225	-0.008	0.236	-0.700***	-0.702***	-0.848	-1.192**
	(0.228)	(0.008)	(0.193)	(0.265)	(0.248)	(0.530)	(0.597)
Mean Dept Var	-0.034	0.0002	-0.032	0.236	0.244	0.205	0.198
Observations	2,800	2,800	2,800	2,800	2,800	2,093	2,072
	RR<50%	**Malleable**	**SLD**	**Speech**	**ED**	**OHI**	**Autism**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$Disp_{d,2004} \times FracExpo$	0.343**	0.240	-0.052	0.135	0.086	0.071	0.028
	(0.169)	(0.163)	(0.310)	(0.275)	(0.132)	(0.138)	(0.031)
$SpEd_{d,2004} \times FracExpo$	-0.063	-0.253	0.163	-0.768***	0.130	0.222*	0.101***
	(0.175)	(0.167)	(0.251)	(0.238)	(0.160)	(0.124)	(0.030)
Mean Dept Var	0.105	0.068	-0.079	0.169	-0.003	-0.019	-0.007
Observations	2,800	2,800	2,800	2,800	2,800	2,800	2,800

*** p<0.01, ** p<0.05, * p<0.1 We regress the district-level difference between the percent of students with a particular attribute not in SpEd at grade 9, given SpEd at grade 5 and the percent of students with the attribute in SpEd at grade 5. This outcome is regressed on the 2004 district-level Black disproportionality rate interacted with fraction exposed and the 2004 district-level SpEd rate interacted with fraction exposed, along with cohort fixed effects. $RR < 50\%$ is an indicator for whether students spent less than 50% of their day in a resource room (outside the GE classroom). All outcomes are measured as of 5th grade. Malleable disability types include specific learning disabilities (SLD), speech impairments, emotional disturbance (ED), and other health impairment (OHI).

Table 7 Effect of Policy on **General Education Students**

	Black	Hispanic	White	All Races
SpEd Status G9	(1)	(2)	(3)	(4)
$DispBlack_{d,2004} \times FracExpo$	-0.101***	-0.003	0.005	-0.013
	(0.035)	(0.028)	(0.018)	(0.019)
$DispHispanic_{d,2004} \times FracExpo$	-0.164	-0.179***	0.010	-0.042*
	(0.123)	(0.044)	(0.025)	(0.025)
$SpEd_{d,2004} \times FracExpo$	-0.292***	-0.219***	-0.189***	-0.182***
	(0.082)	(0.037)	(0.025)	(0.026)
Mean Dept Var	0.047	0.033	0.031	0.033
High School Completion				
$DispBlack_{d,2004} \times FracExpo$	0.159*	0.171*	0.020	-0.0002
	(0.094)	(0.095)	(0.050)	(0.069)
$DispHispanic_{d,2004} \times FracExpo$	-0.087	-0.108	-0.202***	-0.169**
	(0.369)	(0.131)	(0.071)	(0.078)
$SpEd_{d,2004} \times FracExpo$	0.069	-0.368***	-0.080	-0.185**
	(0.205)	(0.140)	(0.065)	(0.080)
Mean Dept Var	0.693	0.689	0.804	0.745
College Enrollment				
$DispBlack_{d,2004} \times FracExpo$	0.546***	0.015	-0.020	0.116
	(0.134)	(0.180)	(0.082)	(0.124)
$DispHispanic_{d,2004} \times FracExpo$	-1.317***	0.089	-0.178*	-0.240**
	(0.458)	(0.164)	(0.098)	(0.106)
$SpEd_{d,2004} \times FracExpo$	-0.154	-0.367*	0.303**	-0.071
	(0.261)	(0.205)	(0.135)	(0.145)
Mean Dept Var	0.557	0.496	0.681	0.593
Observations	288,626	788,729	1,004,472	2,140,896

*** p<0.01, ** p<0.05, * p<0.1 Regressions are run on students in general education (GE) as of 5th grade between 1994 and 2004. Outcome variables and controls are as defined in Table 2, except that we omit controls for disability type and classroom setting, and include controls for 5th grade math and reading standardized exam scores.

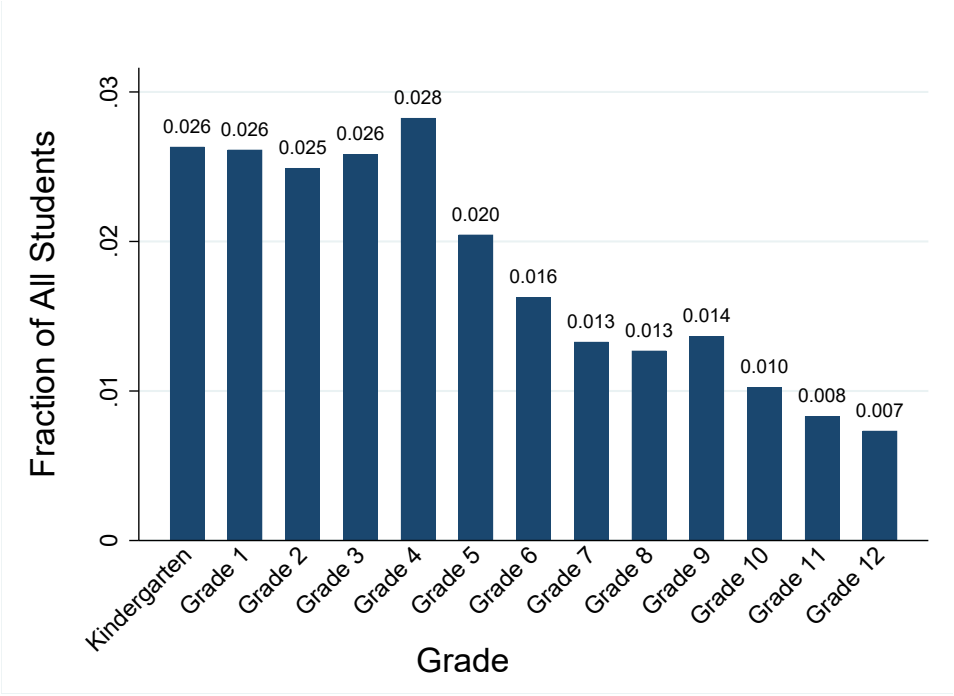
Online Appendix

The Long-Run Impacts of Reducing Racial Gaps in Special Education

Briana Ballis and Katelyn Heath

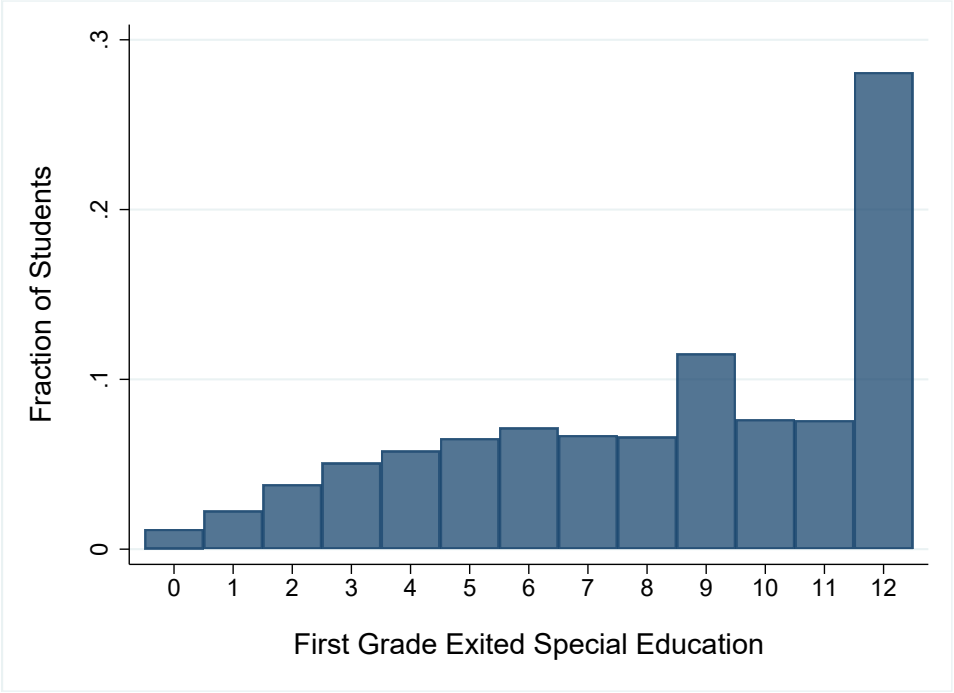
Appendix A

Figure A.1 Fraction of All Students Entering SpEd in Each Grade



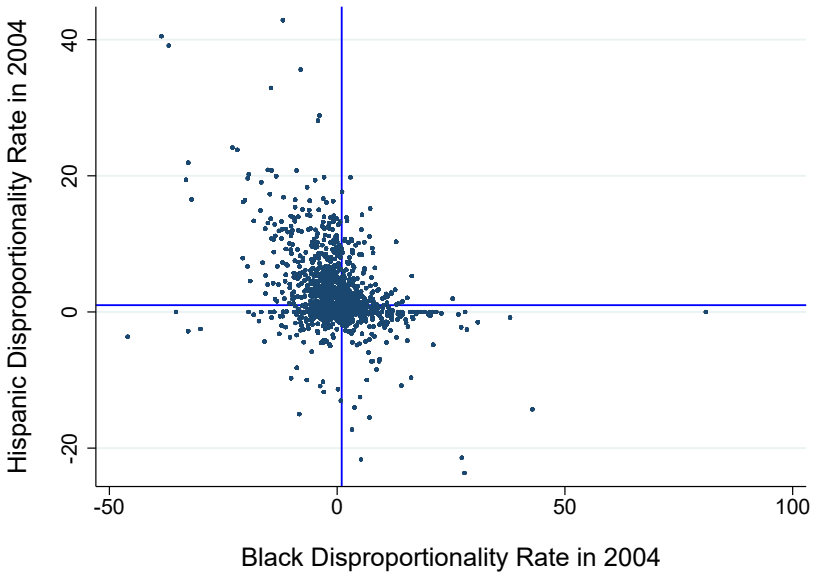
Each bar represents the fraction of students entering SpEd in each grade, out of the total number of students in each grade. This figure includes data from 1994 to 2017.

Figure A.2 Fraction of Students Exiting SpEd in Each Grade



Each bar represents the fraction of students who exited SpEd (for the first time) in each grade, conditional on ever being in SpEd. Students who don't exit before 12th grade are assumed to exit in 12th grade. This figure utilizes pre-policy cohorts, that is, those in 5th grade prior to 2004.

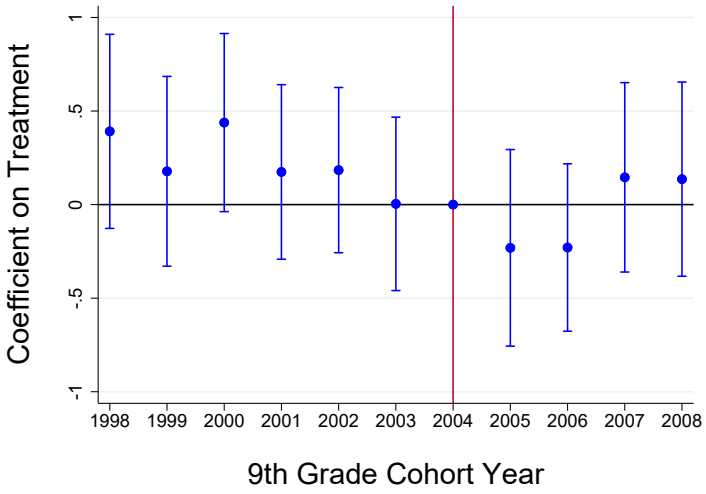
Figure A.3 District-Level Treatment Variation in Black and Hispanic Disproportionality



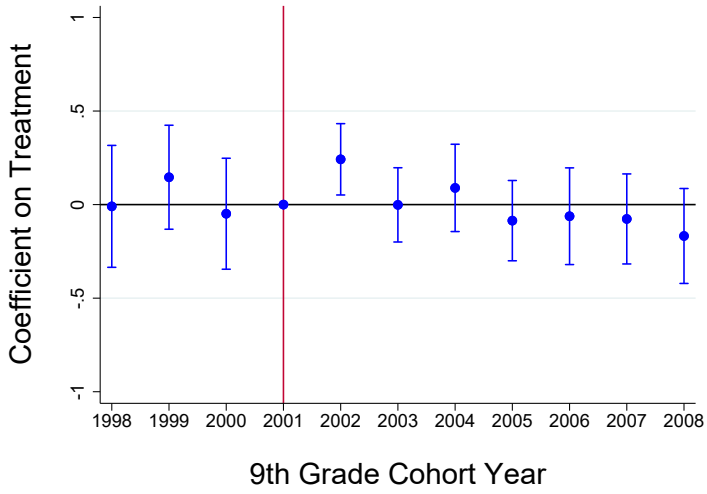
Each dot of the scatter plot represents a district. The x-axis is the 2004 district-level Black disproportionality rate and the y-axis is the 2004 district-level Hispanic disproportionality rate. The correlation coefficient is -0.3506***.

Figure A.4 Event Study Estimates of Hispanic Disproportionality Cap for Hispanic Special Education Students

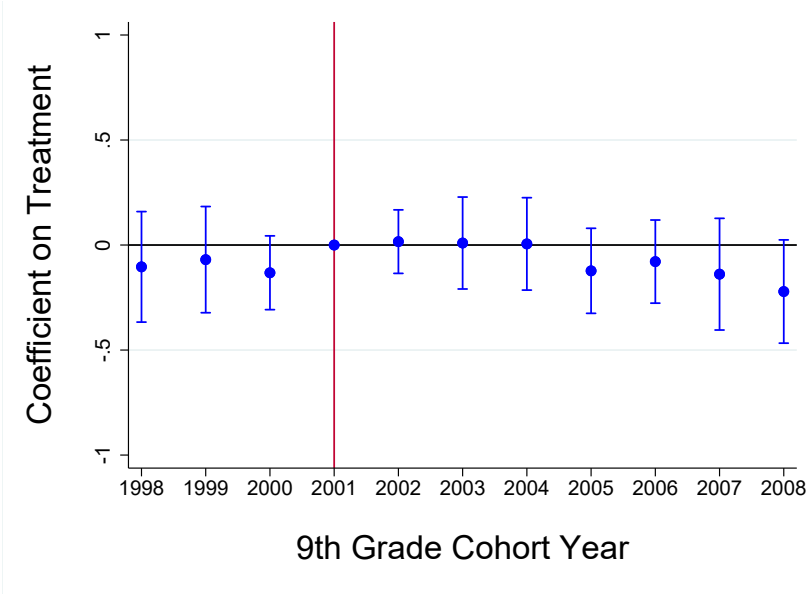
(a) Grade 9 Special Education Status



(b) High School Completion



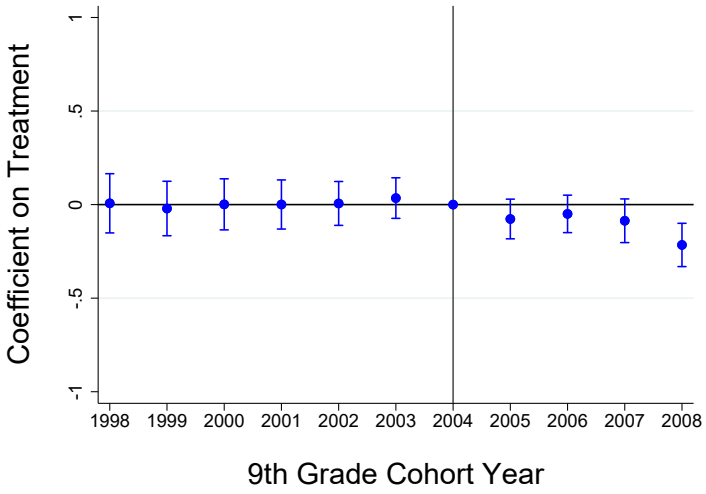
(c) College Enrollment



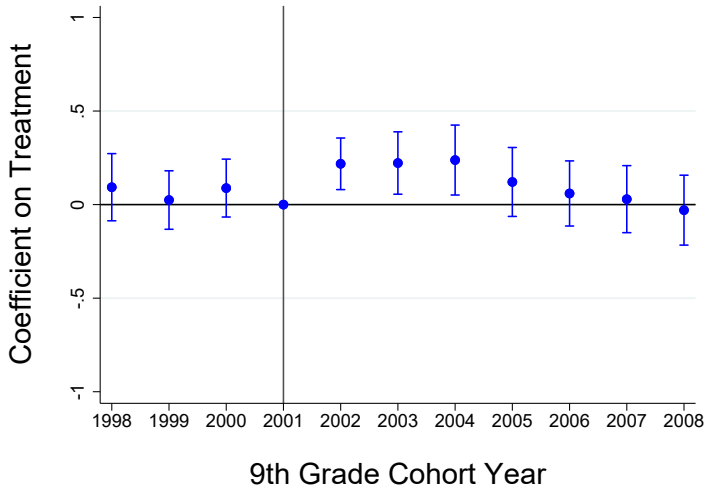
Points represent district-level Hispanic disproportionality rate in 2004 interacted with indicators for each 9th grade cohort year. Y-axis ranges from -1 to 1 percentage point. The vertical bars denote the 95% confidence intervals. See Table 2 for full set of controls used in each regression.

Figure A.5 Event Study Estimates of the Hispanic Disproportionality Cap for Hispanic General Education Students

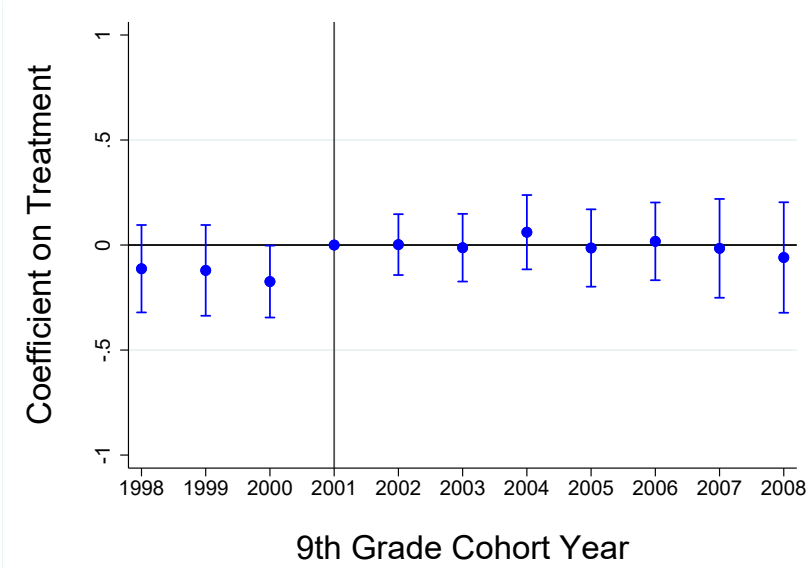
(a) Grade 9 Special Education Status



(b) High School Completion



(c) College Enrollment



Points represent district-level Hispanic disproportionality rate in 2004 interacted with indicators for each 9th grade cohort year. Y-axis ranges from -1 to 1 percentage point. The vertical bars denote the 95% confidence intervals. See Table 7 for the full set of controls used in each regression.

Table A.1 Difference in Means Between Districts Above and Below Disproportionality Caps in 2004

	Black Disprop. Rate			Hispanic Disprop. Rate		
	Less than 1pp	Greater than 1pp	Difference	Less than 1pp	Greater than 1pp	Difference
Male	0.518	0.517	0.001	0.518	0.516	0.002
White	0.541	0.556	-0.015	0.538	0.568	-0.031
Black	0.075	0.157	-0.082***	0.144	0.070	0.074***
Hispanic	0.371	0.272	0.099***	0.305	0.349	-0.044**
Other	0.012	0.014	-0.002	0.013	0.013	0.000
FRL	0.543	0.504	0.040***	0.538	0.497	0.041**
ESL	0.048	0.041	0.007*	0.044	0.045	0.001
Title I	0.720	0.655	0.065**	0.695	0.673	0.022
Gifted	0.066	0.063	0.003	0.061	0.070	-0.010***
Special Ed	0.137	0.142	-0.005	0.140	0.138	0.002
Standardized Math	-0.071	-0.033	-0.038	-0.099	0.028	-0.127***
Standardized Reading	-0.006	0.018	-0.024	-0.037	0.078	-0.015***
Urban	0.217	0.268	-0.051*	0.286	0.170	0.116***
N	592	611		760	443	

*** p<0.01, ** p<0.05, * p<0.1 This table provides descriptive statistics on all students in 2004 for grades K to 12 in districts with less than a 1% disproportionality rate for Black/Hispanic students and districts with greater than a 1% disproportionality rate for Black/Hispanic students. N represents the number of districts.

Table A.2 Effect of Policy on Special Education Students by Grade

	Black Students				Hispanic Students			
	3rd	4th	5th	6th	3rd	4th	5th	6th
SpEd Status	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Disp_{d,2004} \times FracExpo$	-0.116	-0.343**	-0.376**	-0.312**	-0.136	-0.160	-0.153	-0.270
	(0.146)	(0.147)	(0.162)	(0.151)	(0.225)	(0.241)	(0.209)	(0.201)
$SpEd_{d,2004} \times FracExpo$	-0.601***	-0.593***	-0.627**	-0.471**	-0.750***	-0.776***	-0.721***	-0.575***
	(0.193)	(0.193)	(0.250)	(0.197)	(0.261)	(0.263)	(0.242)	(0.222)
Mean Dept Var	0.802	0.742	0.774	0.827	0.775	0.765	0.760	0.823
HS Completion								
$Disp_{d,2004} \times FracExpo$	0.009	0.311*	0.404***	0.516***	-0.375	-0.399**	-0.441**	-0.278
	(0.207)	(0.180)	(0.155)	(0.147)	(0.303)	(0.199)	(0.173)	(0.179)
$SpEd_{d,2004} \times FracExpo$	0.327	0.181	0.127	0.036	-0.723**	-0.634**	-0.575**	-0.543**
	(0.269)	(0.257)	(0.204)	(0.219)	(0.295)	(0.253)	(0.231)	(0.228)
Mean Dept Var	0.635	0.623	0.602	0.603	0.634	0.618	0.597	0.591
College Enrollment								
$Disp_{d,2004} \times FracExpo$	0.459***	0.494***	0.476***	0.295**	0.103	-0.001	0.043	-0.072
	(0.175)	(0.137)	(0.131)	(0.145)	(0.268)	(0.242)	(0.208)	(0.225)
$SpEd_{d,2004} \times FracExpo$	-0.104	0.172	0.170	0.225	-0.604**	-0.608**	-0.512**	-0.358*
	(0.283)	(0.253)	(0.234)	(0.243)	(0.283)	(0.258)	(0.219)	(0.195)
Mean Dept Var	0.310	0.316	0.312	0.298	0.291	0.290	0.281	0.263
Observations	43,851	60,812	75,725	60,498	96,687	131,408	158,855	125,364

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full set of controls. Each sample contains estimates for students in SpEd as of 4th, 5th, or 6th grade prior to policy implementation.

Table A.3 Effect of the Policy on Black Special Education Students by Disability Type

	Black Students						
	All	SLD	Speech	ED	OHI	ID	Physical
SpEd Status	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$Disp_{d,2004} \times FracExpo$	-0.376**	-0.548***	-0.293	-0.031	-0.586*	-0.246*	0.342
	(0.162)	(0.198)	(0.492)	(0.337)	(0.316)	(0.139)	(0.637)
$SpEd_{d,2004} \times FracExpo$	-0.627**	-1.308***	-1.816**	0.140	-0.309	0.072	-0.410
	(0.250)	(0.296)	(0.873)	(0.166)	(0.574)	(0.306)	(0.313)
Mean Dept Var	0.774	0.807	0.326	0.791	0.851	0.953	0.920
High School Completion							
$Disp_{d,2004} \times FracExpo$	0.404***	0.425**	-0.190	0.614	0.221	0.599	-0.519
	(0.155)	(0.185)	(0.407)	(0.424)	(0.517)	(0.373)	(0.861)
$SpEd_{d,2004} \times FracExpo$	0.127	0.111	-0.230	0.056	0.094	1.409*	0.532
	(0.204)	(0.345)	(0.876)	(0.258)	(0.868)	(0.736)	(0.548)
Mean Dept Var	0.602	0.614	0.663	0.448	0.630	0.577	0.655
College Enrollment							
$Disp_{d,2004} \times FracExpo$	0.476***	0.443***	0.557	0.323	1.001**	-0.078	0.477
	(0.131)	(0.157)	(0.455)	(0.447)	(0.470)	(0.329)	(0.917)
$SpEd_{d,2004} \times FracExpo$	0.170	0.390	0.313	-0.440	-0.845	1.805***	0.107
	(0.234)	(0.309)	(0.887)	(0.273)	(0.866)	(0.643)	(0.487)
Mean Dept Var	0.312	0.311	0.473	0.263	0.371	0.124	0.396
Observations	75,725	46,218	8,232	6,618	5,460	6,677	1,535

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full set of controls. Disability type is measured as of 5th grade, and include Specific Learning Disability (SLD), Speech Impairments (Speech), Emotional Disturbance (ED), Other Health Impairments (OHI), and intellectual disability (ID). Physical combines orthopedic impairments, auditory and visual impairments, and deaf/blindness.

Table A.4 Heterogeneous Effects of the Policy for Black Special Education Students

			District Size			
	Urban	Rural	Quartile 1	Quartile 2	Quartile 3	Quartile 4
SpEd Status	(1)	(2)	(3)	(4)	(5)	(6)
$Disp_{d,2004} \times FracExpo$	-0.539**	-0.142	-0.060	-0.666***	-0.964***	-0.278
	(0.233)	(0.232)	(0.209)	(0.231)	(0.295)	(0.635)
$SpEd_{d,2004} \times FracExpo$	-0.549	-0.491*	-0.064	-1.634***	0.021	-0.896
	(0.346)	(0.265)	(0.170)	(0.572)	(0.905)	(1.311)
Mean Dept Var	0.753	0.834	0.821	0.777	0.741	0.759
High School Completion						
$Disp_{d,2004} \times FracExpo$	0.279	0.405*	0.348	0.211	-0.248	0.512
	(0.222)	(0.222)	(0.232)	(0.260)	(0.488)	(0.846)
$SpEd_{d,2004} \times FracExpo$	0.501	-0.245	-0.109	1.604***	-0.284	0.183
	(0.324)	(0.344)	(0.193)	(0.520)	(0.922)	(1.490)
Mean Dept Var	0.574	0.680	0.668	0.618	0.625	0.543
College Enrollment						
$Disp_{d,2004} \times FracExpo$	0.640***	0.164	0.179	0.487**	0.264	0.890
	(0.173)	(0.211)	(0.233)	(0.243)	(0.384)	(0.544)
$SpEd_{d,2004} \times FracExpo$	0.129	-0.046	-0.229	1.760**	0.627	-1.471*
	(0.297)	(0.367)	(0.219)	(0.713)	(0.672)	(0.710)
Mean Dept Var	0.317	0.298	0.296	0.326	0.369	0.288
Observations	55,883	19,842	16,642	18,601	12,328	28,154

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full set of controls. Urban/Rural and district size are measured as of 2004.

Table A.5 District Switching for Black Special Education Students

	Black Students		
	All	FRL	Non-FRL
Switch Districts G9	(1)	(2)	(3)
$Disp_{d,2004} \times FracExpo$	0.553***	0.473**	1.134***
	(0.198)	(0.211)	(0.310)
$SpEd_{d,2004} \times FracExpo$	-0.457**	-0.414*	-0.224
	(0.191)	(0.223)	(0.406)
Mean Dept Var	0.262	0.267	0.239
Observations	75,725	60,791	14,934

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full set of controls. We estimate the likelihood of switching districts between 5th and expected 9th grade. In column (2) we condition on students being FRL as of 5th grade and in column (3) we condition on non-FRL eligibility as of 5th grade.

Table A.6 Effect of Policy on Black Special Education Students Controlling for District Switching

	Original	Controlling for Switching
SpEd Status G9	(1)	(2)
$Disp_{d,2004} \times FracExpo$	-0.376**	-0.311*
	(0.162)	(0.165)
$SpEd_{d,2004} \times FracExpo$	-0.627**	-0.677***
	(0.250)	(0.257)
$Switch_{id}$		-0.117***
		(0.010)
Mean Dept Var	0.774	0.774
High School Completion		
$Disp_{d,2004} \times FracExpo$	0.404***	0.418***
	(0.155)	(0.159)
$SpEd_{d,2004} \times FracExpo$	0.127	0.070
	(0.204)	(0.203)
$Switch_{id}$		-0.103***
		(0.006)
Mean Dept Var	0.602	0.602
College Enrollment		
$Disp_{d,2004} \times FracExpo$	0.476***	0.481***
	(0.131)	(0.132)
$SpEd_{d,2004} \times FracExpo$	0.170	0.153
	(0.234)	(0.233)
$Switch_{id}$		-0.031***
		(0.007)
Mean Dept Var	0.312	0.312

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full set of controls. The term $Switch_{id}$ is an indicator variable for whether a student switched districts between 5th and 9th grade.

Table A.7 Effect of Policy on Black Special Education Students' Intermediate Outcomes

	Behavioral Outcomes					
	% Days Absent	**3+ Truant**	**10+ Truant**	**Suspended**	**Mult. Suspended**	**Expulsion**
$Disp_{d,2004} \times FracExpo$	-3.603	-0.016	0.029	0.031	0.062	0.038
	(2.599)	(0.033)	(0.025)	(0.181)	(0.155)	(0.031)
$SpEd_{d,2004} \times FracExpo$	4.237	0.053	-0.050	-0.200	-0.223	0.008
	(3.343)	(0.066)	(0.046)	(0.314)	(0.229)	(0.044)
Mean Dept Var	7.755	0.002	0.002	0.409	0.282	0.008
Observations	70,593	81,169	81,169	81,169	81,169	81,169

	Academic Outcomes		
	Repeat Grade	**Took G8 Math**	**Took G8 Reading**
$Disp_{d,2004} \times FracExpo$	0.042	0.834***	0.824***
	(0.065)	(0.210)	(0.222)
$SpEd_{d,2004} \times FracExpo$	0.047	0.086	0.195
	(0.099)	(0.201)	(0.234)
Mean Dept Var	0.061	0.335	0.336
Observations	78,815	78,815	78,815

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full set of controls. All outcomes in the top panel are measured as of 9th grade. % Days Absent is the percent of school days an individual was absent in expected 9th grade. 3+ Truant indicates that a student had 3 or more unexcused absences. 10+ Truant indicates that a student had 10 or more unexcused absences. Suspended is an indicator for whether a student was suspended in expected 9th grade, including both in-school and out-of-school suspensions. Mult. Suspended is an indicator for being suspended multiple times. Expulsion is an indicator for being expelled or otherwise displaced from school (e.g. this includes placement in a juvenile justice setting). Grade repeating is measured as an indicator equal to one if an individual repeated a grade between 5th and 9th grade. Took G8 math or reading indicates whether the individual took the 8th grade math or reading exam.

Table A.8 Inclusion of Demographic Trends for Effect of Policy on Special Education Students

	Black Students							
	Original	Black Comp.	Hispanic Comp.	FRL Comp.	ESL Comp.	Title I Comp.	Math Comp.	Reading Comp.
SpEd Status G9	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Disp_{d,2004} \times FracExpo$	-0.376**	-0.486***	-0.367**	-0.600***	-0.471***	-0.501***	-0.438***	-0.478***
	(0.162)	(0.173)	(0.158)	(0.152)	(0.145)	(0.153)	(0.156)	(0.152)
$SpEd_{d,2004} \times FracExpo$	-0.627**	-0.598**	-0.502**	-0.621***	-0.465**	-0.648***	-0.639***	-0.598***
	(0.250)	(0.241)	(0.225)	(0.206)	(0.236)	(0.229)	(0.229)	(0.207)
High School Completion								
$Disp_{d,2004} \times FracExpo$	0.404***	0.369**	0.364**	0.298*	0.416**	0.349**	0.404**	0.401**
	(0.155)	(0.163)	(0.153)	(0.159)	(0.163)	(0.160)	(0.158)	(0.159)
$SpEd_{d,2004} \times FracExpo$	0.127	0.136	0.124	0.115	0.098	0.105	0.111	0.110
	(0.204)	(0.204)	(0.211)	(0.207)	(0.203)	(0.207)	(0.207)	(0.206)
College Enrollment								
$Disp_{d,2004} \times FracExpo$	0.476***	0.435***	0.440***	0.455***	0.453***	0.435***	0.470***	0.477***
	(0.131)	(0.142)	(0.129)	(0.139)	(0.130)	(0.134)	(0.128)	(0.129)
$SpEd_{d,2004} \times FracExpo$	0.170	0.183	0.130	0.171	0.216	0.157	0.187	0.191
	(0.234)	(0.235)	(0.245)	(0.238)	(0.248)	(0.235)	(0.241)	(0.244)

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full list of controls. In column (1) we have our original main results from column (3) of Table 2. In columns (2) through (8), we include indicator variables for each cohort year interacted with baseline demographics, including racial composition, FRL, ESL, Title I, and math and reading performance.

Table A.9 Effect of Policy on Enrollment for Black Special Education Students

	Black Students		
	All	FRL	Non-FRL
Enrolled G6	(1)	(2)	(3)
$Disp_{d,2004} \times FracExpo$	0.046	0.021	0.174
	(0.056)	(0.059)	(0.194)
$SpEd_{d,2004} \times FracExpo$	-0.150**	-0.125	-0.143
	(0.071)	(0.084)	(0.150)
Mean Dept Var	0.943	0.946	0.930
Enrolled G7			
$Disp_{d,2004} \times FracExpo$	0.058	0.046	0.208
	(0.083)	(0.085)	(0.173)
$SpEd_{d,2004} \times FracExpo$	-0.172	-0.164	-0.062
	(0.115)	(0.121)	(0.201)
Mean Dept Var	0.920	0.924	0.895
Enrolled G8			
$Disp_{d,2004} \times FracExpo$	0.181	0.153	0.367**
	(0.114)	(0.121)	(0.179)
$SpEd_{d,2004} \times FracExpo$	-0.333***	-0.299**	-0.406*
	(0.128)	(0.145)	(0.210)
Mean Dept Var	0.900	0.903	0.885
Enrolled G9			
$Disp_{d,2004} \times FracExpo$	0.229*	0.239*	0.203
	(0.127)	(0.128)	(0.187)
$SpEd_{d,2004} \times FracExpo$	-0.462***	-0.412**	-0.610**
	(0.164)	(0.168)	(0.260)
Mean Dept Var	0.874	0.876	0.867
Observations	87,357	70,029	17,328

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full set of controls. We estimate the likelihood of being enrolled as of expected 6th, 7th, 8th, and 9th grades for Black SpEd students (given that they were enrolled in 5th grade). In column (2) we condition on FRL status as of 5th grade and in column (3) we condition on non-FRL as of 5th grade.

Table A.10 Effect of Policy on General and Special Education Spending

	Special Education Spending			
	(1) SpEd Spending Per All	(2) SpEd Spending Per SpEd	(3) Instr. SpEd Spending Per All	(4) Instr. SpEd Spending Per SpEd
$Disp\,Black_{d,2004} \times Post$	-248*	322	-193*	229
	(141)	(1,514)	(103)	(1,151)
$Disp\,Hispanic_{d,2004} \times Post$	-770**	3,115	-446**	2,904
	(352)	(2,731)	(221)	(1,951)
$SpEd_{d,2004} \times Post$	-628*	434	-523**	-904
	(349)	(2,165)	(226)	(1,619)
Mean Dept Var	773	10,221	773	7,635
	General Education Spending			
	(1) GE Spending Per All	(2) GE Spending Per GE	(3) Instr. GE Spending Per All	(4) Instr. GE Spending Per GE
$Disp\,Black_{d,2004} \times Post$	176	49	100	-21
	(340)	(397)	(220)	(255)
$Disp\,Hispanic_{d,2004} \times Post$	-772*	-1,650***	180	-368
	(452)	(543)	(293)	(337)
$SpEd_{d,2004} \times Post$	57	-1,381**	3	-916**
	(510)	(585)	(346)	(399)
Mean Dept Var	4,281	4,794	3,468	3,881

*** p<0.01, ** p<0.05, * p<0.1 Estimates are for district-level expenditures and run using K-12 panel data. Regressions include controls for individual, grade, and district-level gender, ESL, FRL, title I, gifted, and racial composition, as well as district and year fixed effects. Robust standard errors are clustered at the district level. Instr. stands for instructional expenditures.

Table A.11 Heterogeneity by Baseline (4th Grade) Achievement for Black General Education Students

	Math Quintiles				
	≤ 20	20-40	40-60	60-80	≥ 80
SpEd Status G9	(1)	(2)	(3)	(4)	(5)
$Disp_{d,2004} \times FracExpo$	-0.191*	-0.035	-0.018	-0.021	-0.037
	(0.113)	(0.053)	(0.040)	(0.029)	(0.028)
$SpEd_{d,2004} \times FracExpo$	-0.750***	-0.360***	-0.101	-0.087	-0.073
	(0.190)	(0.113)	(0.085)	(0.070)	(0.073)
Mean Dept Var	0.109	0.031	0.015	0.009	0.005
High School Completion					
$Disp_{d,2004} \times FracExpo$	0.270	0.106	0.067	0.264*	0.081
	(0.217)	(0.154)	(0.119)	(0.141)	(0.124)
$SpEd_{d,2004} \times FracExpo$	-0.029	0.479	0.084	0.071	0.012
	(0.307)	(0.302)	(0.248)	(0.270)	(0.229)
Mean Dept Var	0.557	0.703	0.770	0.826	0.868
College Enrollment					
$Disp_{d,2004} \times FracExpo$	0.558**	0.616***	0.436**	0.586**	0.711***
	(0.217)	(0.202)	(0.190)	(0.229)	(0.187)
$SpEd_{d,2004} \times FracExpo$	-0.166	-0.020	0.064	-0.363	0.024
	(0.349)	(0.329)	(0.328)	(0.421)	(0.361)
Mean Dept Var	0.423	0.558	0.630	0.689	0.725
Observations	62,240	55,258	51,327	45,525	38,866

	Reading Quintiles				
	≤ 20	20-40	40-60	60-80	≥ 80
SpEd Status G9	(1)	(2)	(3)	(4)	(5)
$Disp_{d,2004} \times FracExpo$	-0.181	-0.032	-0.051	-0.035	-0.040
	(0.112)	(0.051)	(0.039)	(0.036)	(0.026)
$SpEd_{d,2004} \times FracExpo$	-0.543**	0.063	-0.150**	0.047	-0.072
	(0.220)	(0.120)	(0.076)	(0.071)	(0.053)
Mean Dept Var	0.113	0.033	0.016	0.010	0.006
High School Completion					
$Disp_{d,2004} \times FracExpo$	0.204	0.304**	0.041	-0.020	0.180
	(0.224)	(0.154)	(0.152)	(0.138)	(0.125)
$SpEd_{d,2004} \times FracExpo$	0.025	-0.454	0.455	-0.154	0.304
	(0.302)	(0.318)	(0.296)	(0.273)	(0.240)
Mean Dept Var	0.565	0.689	0.756	0.815	0.863
College Enrollment					
$Disp_{d,2004} \times FracExpo$	0.631***	0.542**	0.592***	0.733***	0.447**
	(0.234)	(0.211)	(0.177)	(0.181)	(0.205)
$SpEd_{d,2004} \times FracExpo$	-0.132	-0.280	-0.317	-0.304	-0.415
	(0.329)	(0.376)	(0.384)	(0.359)	(0.332)
Mean Dept Var	0.412	0.542	0.623	0.687	0.730
Observations	57,739	52,323	51,379	46,817	44,651

*** p<0.01, ** p<0.05, * p<0.1 See Table 7 for full set of controls. Results are estimated for Black GE students separately across quintiles of 4th grade math and reading performance.

Table A.12 Effect of Policy on Hispanic Special Education Students

	Hispanic Students		
SpEd Status G9	(1)	(2)	(3)
$Disp_{d,2004} \times FracExpo$	-0.201	-0.096	-0.153
	(0.205)	(0.186)	(0.209)
$SpEd_{d,2004} \times FracExpo$	-0.910***	-0.713***	-0.721***
	(0.237)	(0.242)	(0.242)
Mean Dept Var	0.760	0.760	0.760
High School Completion			
$Disp_{d,2004} \times FracExpo$	-0.430**	-0.470***	-0.441**
	(0.169)	(0.168)	(0.173)
$SpEd_{d,2004} \times FracExpo$	-0.590***	-0.581***	-0.575**
	(0.217)	(0.224)	(0.231)
Mean Dept Var	0.597	0.597	0.597
College Enrollment			
$Disp_{d,2004} \times FracExpo$	0.072	0.047	0.043
	(0.188)	(0.200)	(0.208)
$SpEd_{d,2004} \times FracExpo$	-0.410**	-0.503**	-0.512**
	(0.172)	(0.213)	(0.219)
Mean Dept Var	0.281	0.281	0.281
Summary Index			
$Disp_{d,2004} \times FracExpo$			0.043
			(0.208)
$SpEd_{d,2004} \times FracExpo$			-1.429***
			(0.462)
Observations	158,855	158,855	158,855
Individual Controls		X	X
District-Cohort Controls			X

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full set of controls. In the Summary Index panel, the outcome variable is a summary measure of high school completion and college enrollment. We standardize each outcome to have mean 0 and standard deviation 1. Then, we create one summary index by averaging across the standardized long-run outcomes for each individual.

Appendix B

B.1 Robustness of Controlling for SpEd Enrollment

In this section, we expand on the discussion of the inclusion of the SpEd enrollment cap in our main specification. We begin with Appendix Figure B.1, which presents an analogous figure for the district-level SpEd rate to graphically illustrate the intuition behind including the $SpEd_{2004,d} * FracExpo_c$ term in our specification. Districts are sorted based on their 2004 SpEd rate. The bottom series in each figure (denoted with circles) shows the average SpEd rate from 1994 to 2017 for districts al-

ready below the 8.5% threshold in 2004. In the top three series, districts are split into terciles based on their 2004 SpEd rate, conditional on being above 8.5%. The figure illustrates that districts with the highest rates of SpEd made the largest reductions across the post-period in their SpEd rates, indicating that they are more treated by the policy relative to those already meeting or nearly meeting the threshold.

In our main specification, in equation (1), we include the disproportionality and SpEd caps additively in our model. We provide several pieces of evidence to justify this functional form. First, as Appendix Figures B.2a and B.2b show, there is no correlation between the two treatment variables. The correlation coefficient between the SpEd rate and Black disproportionality shown in Figure B.2a is 0.0022 and the coefficient between the SpEd rate and Hispanic disproportionality shown in Figure B.2b is 0.0310. Second, as mentioned previously in Section 4, Appendix Table B.1 shows that the effect of the disproportionality caps remain quantitatively and qualitatively similar when we do not control for the SpEd enrollment cap. Finally, we show in Appendix Table B.2 that the impacts of the disproportionality caps remain fairly similar across districts with higher vs. lower SpEd rates. Specifically, we estimate the impact of the disproportionality cap in districts with above and below the median level of SpEd enrollment. Although our estimates for Black students in districts with below median SpEd rates lose significance for SpEd enrollment, the magnitude of the results for SpEd enrollment remains similar.

B.2 Replicating and Reconciling Ballis and Heath (2021)

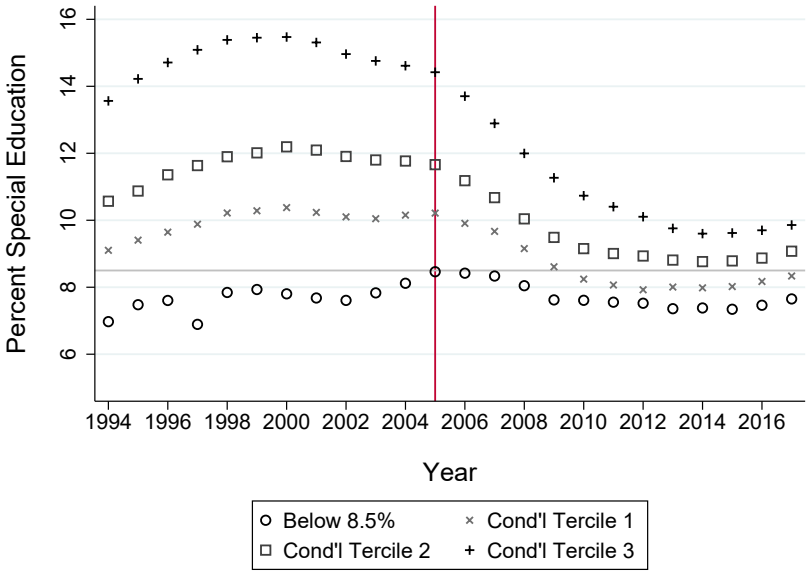
In Ballis and Heath (2021), we find large negative impacts of SpEd removal driven by the SpEd enrollment cap, and conclude that these effects are driven by low-income minority students. While seemingly at odds with the positive effects of SpEd removal for Black students in this paper, the results of these two papers are not inconsistent. In Appendix Table B.3 Column (1) below, we replicate Column (3) Table 5 of Ballis and Heath (2021) by re-estimating the effect of the SpEd enrollment cap on minority students (i.e., Black and Hispanic students combined). Note that the magnitude of the coefficients are different across the two tables, due to minor differences in specification choices. In this paper, we use a longer sample (back to 1994-95 rather than back to 2000-01) and do not use a "high impact" sample (which was defined as students with

so called malleable disability types in the previous paper). These minor differences in sample and variable specification do not ultimately affect the economic significance of the results or conclusions that we can draw from them. In both cases, we find large negative impacts on the likelihood of remaining in SpEd at 9th grade, the likelihood of completing high school, and the likelihood of college enrollment for minority students as a result of the SpEd enrollment cap.

Furthermore, in Appendix Table B.3 column (2) we include controls for the Black and Hispanic disproportionality caps. Reassuringly, the conclusions we draw from the SpEd cap remain similar, even after including the disproportionality controls. The one exception is the coefficient on the SpEd enrollment cap for college completion, which becomes smaller in magnitude and no longer statistically significant when we control for the disproportionality caps. However, the effect is still negative. In columns (3) and (4) we reproduce the main results for Black and Hispanic students from Table 2. This comparison makes clear that the positive effects of the Black disproportionality cap are driven by Black SpEd students, whereas the negative effects of the SpEd enrollment cap are driven by Hispanic SpEd students.

Recall from our conceptual framework in Section 2.3 that differences in SpEd placement across race can help predict differences in whether students will benefit or be harmed by SpEd removal. The negative impacts resulting from the SpEd enrollment cap are driven by Hispanic students, who are not over-represented in SpEd on average. Thus, this is consistent with these students having $\beta > a_i - c$, and for whom there isn't an additional γ_s bias term. In contrast, Black students, on average, benefit from the cap on Black disproportionality. Thus, our model implies that there exists some additional cost, γ_s, such that Black students are placed in SpEd at much higher than optimal rates. When these Black students are removed from SpEd, they perform better. Through our empirical work, we find several pieces of evidence that support a story of misclassification among Black students removed from SpEd due to the Black disproportionality cap.

Figure B.1 Percent of Students in SpEd by District SpEd Rate at Baseline

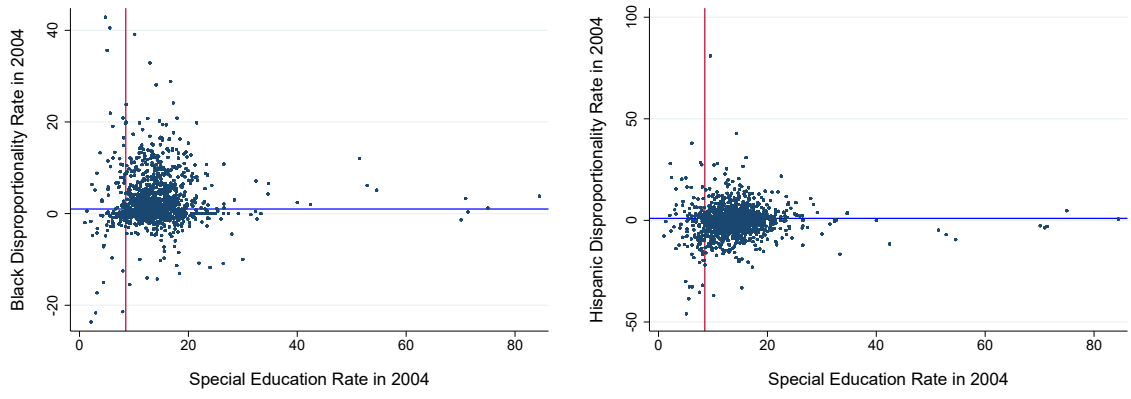


This figure plots the percent of students in SpEd from 1994 to 2017. Districts are split into four groups. The bottom series consists of districts with an average SpEd rate already below 8.5% prior to 2004. The top three series split the remaining districts above 8.5% into terciles based on the pre-period percent of students in SpEd.

Figure B.2

(a) District-Level Treatment Variation for Black Students

(b) District-Level Treatment Variation for Hispanic Students



Each dot of the scatter plots represents a district. The x-axis is the 2004 district-level SpEd rate, and the y-axis is the 2004 district-level Black or Hispanic disproportionality rate. The correlation coefficient in Figure (a) is 0.0022 and in Figure (b) is 0.0310***.

Table B.1 Impact of Disproportionality Cap with and without controlling for the Special Education Enrollment Cap on Special Education Students

	Black Students			Hispanic Students		
SpEd Status G9	(1)	(2)	(3)	(4)	(5)	(6)
$Disp_{d,2004} \times FracExpo$	-0.376**	-0.335*		-0.153	-0.295	
	(0.162)	(0.175)		(0.209)	(0.218)	
$SpEd_{d,2004} \times FracExpo$	-0.627**		-0.581**	-0.721***		-0.736***
	(0.250)		(0.247)	(0.242)		(0.241)
High School Completion						
$Disp_{d,2004} \times FracExpo$	0.404**	0.397**		-0.441**	-0.534***	
	(0.155)	(0.154)		(0.173)	(0.175)	
$SpEd_{d,2004} \times FracExpo$	0.127		0.083	-0.575**		-0.615***
	(0.204)		(0.208)	(0.231)		(0.231)
College Enrollment						
$Disp_{d,2004} \times FracExpo$	0.476***	0.467***		0.043	-0.040	
	(0.131)	(0.132)		(0.208)	(0.203)	
$SpEd_{d,2004} \times FracExpo$	0.170		0.119	-0.512**		-0.508**
	(0.234)		(0.239)	(0.129)		(0.215)

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full set of controls. Columns (1) and (4) control for both the disproportionality cap and the SpEd cap. Columns (2) and (5) control only for the relevant disproportionality cap. And, columns (3) and (6) control only for the SpEd enrollment cap.

Table B.2 Effect of Disproportionality Across Levels of Special Education on Special Education Students

	Black Students		Hispanic Students	
	Above Median	Below Median	Above Median	Below Median
SpEd Status G9	(1)	(2)	(3)	(4)
$Disp_{d,2004} \times FracExpo$	-0.445***	-0.368	-0.366	0.304
	(0.154)	(0.311)	(0.263)	(0.374)
Mean Dept Var	0.805	0.743	0.789	0.725
High School Completion				
$Disp_{d,2004} \times FracExpo$	0.415**	0.0476	-0.531**	-0.509*
	(0.191)	(0.250)	(0.229)	(0.293)
Mean Dept Var	0.638	0.565	0.606	0.586
College Enrollment				
$Disp_{d,2004} \times FracExpo$	0.463***	0.393	0.108	-0.320
	(0.148)	(0.259)	(0.240)	(0.290)
Mean Dept Var	0.322	0.303	0.273	0.292
Observations	37,914	37,811	87,916	70,939

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full set of controls. Columns (1) and (3) estimate the impact of the disproportionality cap in districts with above the median level of special education enrollment (which is 11.7%), while columns (2) and (4) estimate the impact of the disproportionality cap in districts with below the median level of special education enrollment.

Table B.3 Replicating Column 3 in Table 5 of Ballis and Heath (2021): Impact of SpEd Enrollment Cap and Disproportionality Caps on Minority Special Education Students

	Minority Students		Black Students	Hispanic Students
SpEd Status G9	(1)	(2)	(3)	(4)
$Disp\,Black_{d,2004} \times FracExpo$		-0.291**	-0.376**	
		(0.136)	(0.162)	
$Disp\,Hispanic_{d,2004} \times FracExpo$		-0.261		-0.153
		(0.199)		(0.209)
$SpEd_{d,2004} \times FracExpo$	-0.702***	-0.745***	-0.627**	-0.721***
	(0.202)	(0.211)	(0.250)	(0.242)
High School Completion				
$Disp\,Black_{d,2004} \times FracExpo$		0.0496	0.404***	
		(0.109)	(0.155)	
$Disp\,Hispanic_{d,2004} \times FracExpo$		-0.407**		-0.441**
		(0.176)		(0.173)
$SpEd_{d,2004} \times FracExpo$	-0.372**	-0.336*	0.127	-0.575**
	(0.187)	(0.192)	(0.204)	(0.231)
College Enrollment				
$Disp\,Black_{d,2004} \times FracExpo$		0.421***	0.476***	
		(0.099)	(0.131)	
$Disp\,Hispanic_{d,2004} \times FracExpo$		0.002		0.043
		(0.199)		(0.208)
$SpEd_{d,2004} \times FracExpo$	-0.327*	-0.235	0.170	-0.512**
	(0.177)	(0.172)	(0.234)	(0.219)
Observations	234,519	234,519	75,725	158,855

*** p<0.01, ** p<0.05, * p<0.1 See Table 2 for full set of controls. Minority students refers to Black and Hispanic students.