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

Briana Ballis* and Katelyn Heath†

May 20, 2023

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 capping Black over-representation in SpEd led to small gains in high school completion and college attainment for Black students in special and general education. Overall, our results suggest that reductions in 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, University of California at Merced. Email:bballis@ucmerced.edu

†**Corresponding author.** Texas Policy Lab, Rice University. Email: katelyn.heath@rice.edu. Address: 478 Kraft Hall, Rice University, 6100 Main Street, Houston, TX 77005

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. We would like to also thank Francesca Molinari, Evan Riehl, Seth Sanders, and participants in the Cornell Labor workshop for their helpful comments. 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 would like to acknowledge 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.

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 pre-school 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, and currently over 14 percent of students participate. 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). Ultimately, there is no consensus in the literature on why Black students are over-represented in SpEd. It could be the case that Black students have a greater underlying need for SpEd services.[1] Race may be correlated with other factors (such as socioeconomic status) that contribute to the need for services. Alternatively, it could be the case that implicit (or explicit) racial biases in the SpEd referral or evaluation process result in some Black students being misidentified for SpEd.[2] While a large literature has explored the disproportionate representation of Black students in SpEd, there is little evidence on whether it helps or harms students.

In this paper, we examine the direct and spillover effects of policy that limited over-representation of Black and Hispanic students in SpEd in Texas. Although SpEd services aim

[1] Interestingly, after conditioning on important confounds such as prior academic achievement and socioeconomic status, recent 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; Morgan, Farkas, Hillemeier, & Maczuga, 2017; Shifrer et al., 2011; Hibel et al., 2010).

[2] As we will discuss in Section 2, for many high-incidence 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).

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 are also costs associated with participation. For example, SpEd can impose stigma costs, lower expectations from teachers and/or parents, reduce students own self-perceptions and aspirations, reduce students' exposure to the general curriculum, or result in negative spillover effects 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 in Section 2.3 with a theoretical framework, 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, it is likely that that the costs of SpEd will outweigh the benefits. In such cases, reducing over-representation could lead to improved outcomes. However, if SpEd is reduced among a population that genuinely requires services at higher rates, they are likely to be negatively impacted. Despite the important implications that differences in SpEd placement may have for Black and Hispanic students' long-run trajectories, to our knowledge, no prior study has been able to explore how reducing over-representation influences long-run trajectories.

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, then the district was deemed out of compliance and had to reduce access to SpEd for Black (or Hispanic) students.[3] In 2004, the statewide average Black disproportionality rate was 2.7%. Thus, we expect the policy led to strong pressure to reduce SpEd enrollment among Black students. In contrast, in 2004, the statewide average Hispanic disproportionality rate was -0.05%, well below the 1% threshold.[4] As the majority

[3]Along with this policy change, Texas also implemented a cap on overall SpEd enrollment at 8.5%. We discuss how the SpEd enrollment cap interacts with the disproportionality caps in detail in Sections 2 and 4. Reassuringly, controlling for the 8.5% SpEd enrollment cap does not affect our estimates of the disproportionality caps. Furthermore, although in theory a district could increase the percent of other races in SpEd in order to reduce disproportionality, this was very unlikely in light of the cap on overall SpEd enrollment.

[4]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

of districts were already in compliance with the Hispanic disproportionality cap, we do not expect the policy to have a strong impact on Hispanic SpEd students.

We rely on this exogenous policy change, along with administrative data from Texas, to study how SpEd removal due to the disproportionality caps impacted longer-run outcomes. We use a dose-response difference-in-differences estimation strategy that utilizes variation across districts in their rates of Black and Hispanic disproportionality prior to policy implementation, and across cohorts in the amount of time they spent in school under the policy. We focus on students in SpEd as of 5th grade prior to policy implementation to avoid endogenous changes in the underlying ability distribution of SpEd students in the post-policy period.[5] Our findings throughout represent intent-to-treat estimates of the impact of the disproportionality caps on all Black or Hispanic students in SpEd as of 5th grade prior to policy implementation. We note that effects are likely to be larger for the subset of students who were actually removed from SpEd as a result of the caps.

Ultimately, we find that capping Black disproportionality reduced the likelihood that Black students remained in SpEd at 9th grade by about 1.5%, but *increased* their high school graduation by 2.0% and college-going by 4.6%.[6] In contrast, since Hispanic students are, on average, under-represented in SpEd in Texas, the Hispanic disproportionality cap was significantly less binding. We do not find a statistically significant or economically meaningful impact of the Hispanic disproportionality cap on Hispanic students' SpEd participation.[7]

Although our data prevent us from being able to precisely determine the mechanisms behind the positive impacts of SpEd removal for Black students, we find several pieces of

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.

[5]We justify this sample selection in Section 4 and note that our results are robust to assigning SpEd status as of 4th or 6th grade instead (available upon request).

[6]These reflect average effect sizes for students at the average district, who are exposed to the policy in every year between 5th and 9th grade.

[7]These estimates reflect the impact of SpEd removal, as opposed to the impacts of never receiving SpEd. While it is challenging to identify the students who never receive SpEd in the absence of the policy, in Section 5.2 we look at low-achieving GE students who faced reductions in later SE access. These students experienced positive impacts on educational attainment, and this may provide insights into the effects of never receiving SpEd.

evidence consistent with the misclassification of students for whom the costs of SpEd outweigh the benefits. First, we find that the positive impact of the Black disproportionality cap on long-run outcomes is driven by districts who we predict to be over-classifying Black SpEd students. That is, districts where Black students are more likely to be in SpEd even after conditioning on confounding characteristics (such as free and reduced-price lunch status and early achievement measures). Second, we find suggestive evidence that the positive impacts are concentrated in districts with lower teacher experience and fewer Black teachers, who may be more like to misclassify Black students (Redding, 2019; Gershenson et al., 2016; Dee, 2005). Finally, we find that the policy resulted in students with relatively more mild disabilities being removed from SpEd post-policy, compared to those removed pre-policy. This suggests that districts were selecting SpEd removals based on criteria other than disability severity post-policy, which is consistent with districts removing students who were misidentified for SpEd.

We also explore the effects of reducing disproportionality on General Education (GE) students. We define our sample analogously for GE students as we did for SpEd students, estimating effects on the sample of Black or Hispanic students in GE as of 5th grade prior to policy implementation. These students may have been directly affected by the policy by being less likely to be placed in SpEd in later grades, or indirectly affected if their peers losing access to SpEd led to spillovers in the classroom. A-priori, the effects on GE students could be negative or positive. If the removal of SpEd supports within the GE classroom lead to greater classroom disruption, less time and attention from the GE teacher, or less supports (like teacher's aides) for all students, this could have negative effects on GE students.[8] However, the policy had a positive impact on Black SpEd students, improving their high school completion and college enrollment. Thus, if these SpEd students are now performing better in the GE classroom, this could result in positive peer-to-peer spillovers. Additionally, if the policy led to a reduction in perceived racial bias, this could relieve fears of racial discrimination in terms of future SpEd placement for GE students.

For Black GE students, we find small but significant declines in the likelihood of SpEd

[8]Our data does not include classroom-level information, so we are unable to track movements in and out of the GE classroom. However, the effects of the disproportionality caps are driven by SpEd students who spend the majority of their day in GE classrooms. Thus, we view our results as reflecting changes in supports within GE classrooms, rather than changes in classroom composition.

placement at 9th grade, and positive effects on their high school completion and college enrollment. Although we are not able to directly measure the extent to which these long-run impacts reflect direct vs. spillover effects, the magnitude of the effect on SpEd placement among 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, we find significant improvements in college enrollment among relatively higher achievers, for whom we do not find an effect on SpEd placement in later grades. This suggests that the positive effects for these students are likely driven by positive peer-to-peer effects and/or a reduction in received racial bias as a result of the policy. For Hispanic GE students, despite the fact that we did not find reductions in SpEd placement for SpEd students, we do estimate that capping Hispanic disproportionality led to a very small declines in SpEd placement at 9th grade. This could be driven by the fact that it is likely easier to never place a student in SpEd rather than remove existing services from a population that is, on average, under-represented to begin with. However, we do not find subsequent impacts on high school completion or college enrollment for Hispanic GE students.

Our study offers three primary contributions to the literature. First, we provide novel estimates of the long-term impacts of limiting Black student disproportionality. 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 (Elder et al., 2021; Morgan, Farkas, Hillemeier, & Maczuga, 2017; Morgan, Farkas, Cook, et al., 2017; Gordon, 2017; Morgan et al., 2016; Shifrer et al., 2011; Hibel et al., 2010; Hosp & Reschly, 2003; Donovan & Cross, 2002; Oswald et al., 1999). Most recently, Elder et al. (2021) investigate gaps in SpEd placement across race in Florida. The authors find that Black and Hispanic students tend to be conditionally over-represented in SpEd (compared to White students) in districts with relatively small minority shares and conditionally under-represented in schools with large minority shares. However, no previous studies have been able to evaluate whether there are long-term impacts associated with this difference in SpEd representation across race. We overcome the identification challenges this literature has faced by leveraging a unique policy change in Texas that led to exogenous changes in whether minority students were over-represented in SpEd, which were unrelated to changes in the *need* for SpEd. 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 an exogenous policy-driven reduction in Black disproportionality.

Second, our paper contributes to the small but growing literature on the effectiveness of SpEd programs. Not only has participation in SpEd grown significantly over time, SpEd services are an increasingly expensive component of school expenditures. It is estimated that educating a SpEd student costs roughly twice as much, on average, as educating a GE student (Parrish et al., 2004). 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). In this paper, contrary to most of the prior literature, we find that SpEd removal improves long-run outcomes. Our results suggest that while some students benefit from SpEd services, other students are worse-off in the long-run, potentially as a result of misclassification. Our results are consistent with a conceptual framework we present in Section 2.3, which illustrates that when there is significant racial bias in SpEd identification and placement practices, the costs of SpEd can outweigh the benefits. Overall, we find several pieces of empirical evidence that point to the misclassification of Black students in SpEd, which is consistent with the presence of racial bias in these districts. Ultimately, our findings point to the importance of carefully examining disability evaluation criteria to ensure that students of all races are appropriately evaluated for placement in SpEd. More generally, 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 contributing to racial gaps during adulthood, to our knowledge, there is no existing literature on how SpEd programs may be impacting long-run racial gaps. We know from Ballis and Heath (2021) that SpEd programs can have large and important impacts on long-run outcomes. Yet, how SpEd placement among marginal Black students contributes to disparities across race in adulthood is largely unknown. Our paper helps to fill this gap by addressing the extent to which disproportionality contributes to gaps in educational attainment in adulthood.

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 learning disabilities (by far the largest category), followed by speech impairments, intellectual disabilities, and emotional disturbance (see Appendix Table A.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. IEPs are *individualized* and may vary widely so that each student receives a different set or combination of services depending on the student's disability and the school they attend.

For the majority of students, the process for determining eligibility for SpEd is complex. For a number of disability types, including learning impairments, the tests used to evaluate students vary widely. It is a subjective process to determine which tests to use and whether a student's academic achievement is sufficiently hindered by their disability to qualify for SpEd. 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."

Furthermore, it is difficult to measure the costs and benefits of SpEd services for students on the margin of placement. On the one hand, SpEd is a program specifically designed to provide one-on-one or small group instruction and supports aimed at boosting academic achievement. On the other hand, students may do worse academically as a result of being placed in SpEd (Prenovitz, 2017; Morgan, Frisco, Farkas, & Hibel, 2010; Markussen, 2004).

SpEd may lead students to spend less time on the GE curriculum, reviewing lower-than-grade-level topics that are unnecessary for them. In addition, students may suffer socially from the stigma of a disability label, in a way that harms their persistence and engagement in school (Shifrer, 2013). Finally, students may experience negative spillover effects from their SpEd peers if, for example, their peers exhibit behavioral challenges that negatively influence them.

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).

Despite this policy implemented by the federal government, to our knowledge there were no formal systems in place to monitor disproportionality in Texas until 2004 when the Texas Education Agency introduced the Performance-Based Monitoring Analysis System (PBMAS).[9] 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. 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. Appendix Figures A.1 and A.2 show tables from the 2004-2005 PBMAS Policy Manual illustrating the performance levels associated with varying levels of district disproportionality for Black and Hispanic students.

It is important to note that PBMAS also introduced monitoring of the overall SpEd en-

[9]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.

rollment rate. Any district with over 8.5 percent of students in SpEd was out of compliance under PBMAS standards. Appendix Figure A.3 shows the table from the 2004-2005 PBMAS Policy Manual illustrating the performance levels associated with various rates of SpEd enrollment. In a separate paper, we study in depth the impacts of this SpEd enrollment cap (Ballis & Heath, 2021). The SpEd enrollment cap led to significant reductions in SpEd access, which 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. Furthermore, we reconcile in Appendix B the differences between the negative impacts of the SpEd enrollment cap found in Ballis and Heath (2021) and the positive impacts of the Black disproportionality cap here. 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.

Figure 1a illustrates the percent of students in SpEd in Texas relative to the rest of the U.S. Prior to the policy's implementation, in the 2003-2004 school year, the average percent of students in SpEd was around 12%. By 2016-2017, the average had fallen to about 8.5%. This is in contrast to the national average, which remained approximately steady at around 13.5% from 2004 to 2016. In Figure 1b, we show trends in the percent of students in SpEd in Texas overall and by race. Again, we see a dramatic decrease in SpEd enrollment after 2004 for each race. Across all years we see that 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.

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),

[10]The monitored outcomes for Bilingual/English as a Second Language and Migrant students do not include

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.

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 pressures 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

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.

[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.

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 may vary by the severity of a student's disability and the types of services they receive):

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

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

[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.

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 is predicted to benefit the marginal student.

The presence of racial bias in the model helps to explain how Black students may benefit from disproportionality monitoring. Our results show positive long-run impacts of the Black disproportionality cap. In Section 5.1.2 we present evidence consistent with the misclassification of these Black SpEd students, which also supports a theory of racial bias. In contrast, Hispanic students are not over-represented in SpEd in Texas, on average, and we do not find statistically significant impacts of the Hispanic disproportionality cap. Elder et al. (2021) find that minority students tend to be over-represented in SpEd in schools with relatively small minority shares, and under-represented in SpEd in schools with relatively large minority shares. Hispanic students are not in the statistical minority in Texas, and this may help to explain why they are not over-represented, on average.

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.[13] 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 college attendance in the state of Texas. 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, & 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.

[13]Since we do not have reliable data on measures of dropout in Texas, we estimate impacts on high school completion.

[14]We currently have 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 students in SpEd as of 5th grade between 1994 and 2000 who are Black or Hispanic we find that 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 the number of years it takes to earn 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.

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 18% relative to Hispanic students at 14%. Our final analysis sample consists of 72,196 Black students in SpEd at 5th grade and 153,098 Hispanic students in SpEd at 5th grade. For SpEd students, we only have information on disability type (described further in Appendix Table A.1) and classroom setting.[15] Among all races, learning disabilities is the most common disability type, followed by speech impairments. The vast majority (roughly 90%) of students spend greater than 50% of their day in the GE classroom.

In Appendix Table A.2, we illustrate 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. Additionally, districts above the Black disproportionality threshold 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 Black disproportionality threshold, and for those that do the differences are very 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 disproportionality threshold. We account for differences in baseline characteristics in our empirical strategy by including controls for each of these demographic variables at the individual, grade, and district level. In addition, in Section 5.1.1 we demonstrate that our results are robust to controlling for district-level time-trends in the baseline levels of the demographic variables.

[15]The data do not contain any information beyond this for SpEd students (such as what services or accommodations they receive or classroom size).

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. 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.

Appendix Figures A.4a and A.4b illustrate the intuition behind our treatment variables. We sort districts by their 2004 Black or Hispanic disproportionality rate. The bottom series in each figure, denoted with circles, shows the average disproportionality rate from 1994 to 2017 for districts already below the 1% threshold in 2004. In the top three series, districts are split into terciles based on their 2004 disproportionality rate, conditional on having a rate greater than 1%. Comparing the top most series, denoted with x's, to the bottom series illustrates that districts with the highest rates of disproportionality made the largest reductions across the post-period in their disproportionality rates, indicating that they are more treated by the policy relative to those already meeting or nearly meeting the threshold. However, we find much less response in the first and second terciles above the 1% threshold, especially for Hispanic disproportionality. In addition, the rates of disproportionality are much lower for Hispanic students than for Black students. Overall, we do not expect the Hispanic disproportionality cap to have the same effect on students as the Black disproportionality cap, since the Hispanic disproportionality cap is much less binding.

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.[16] Thus, 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

[16]To reduce their SpEd rate, districts must decide which students will be removed from SpEd and which students will not be placed in SpEd to begin with. These decisions will necessarily impact the underlying ability distribution of the students who remain in SpEd.

grade.[17] Similarly, to estimate effects on GE students we focus on 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:

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

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.[18] The term $Disp_{2004,d}$ represents the 2004 district-level Black or Hispanic disproportionality rate. Rather than interact this with an indicator for the post-policy period, we interact it with $Exposure_c$, which measures the number of years an individual is in school under the policy. This captures the fact that students in school for longer under the policy are more likely to experience greater reductions in access to SpEd. For the outcome of SpEd status in 9th grade, $Exposure_c$ is the number of years each cohort was exposed to the policy between 5th and 9th grade, and for high school graduation and post-secondary enrollment $Exposure_c$ is the number of years each cohort is exposed between 5th and 12th grade.[19] The main coefficient of interest is β_1, which estimates the effect of reducing disproportionality among Black or Hispanic students.

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

[17] Appendix Figure A.5 illustrates the percent of all students entering SpEd by grade, and shows that the fraction of new entries levels off around 4th grade and drops each year after that. When we use students in SpEd as of 4th or 6th grade prior to policy implementation instead of 5th grade, results remain qualitatively and quantitatively similar. These results are available upon request.

[18] 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 in order to avoid endogenous changes in grade repeating.

[19] Exposure is based on the expected number of years in school under the policy, rather than actual years to avoid endogenous changes in exposure driven by grade-repeating.

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. Students are assigned the district in which they are observed in 2004 and their cohort year corresponds to the year they were in kindergarten.[20]

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 percent of students in SpEd in 2004 in each district and $Exposure_c$. 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 (for Black or Hispanic students) provide an accurate counterfactual for trends among districts with high disproportionality rates. We test these assumptions directly by implementing an event study analysis of the following form:

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

where 9th grade cohort indicator variables are interacted with $Disp_{2004,d}$ and $SpEd_{2004,d}$. For the outcome of SpEd status at 9th grade, we exclude the 2004 9th grade cohort, since this is the last cohort to be fully unexposed to the policy during 9th grade. For the long-run outcomes, the 2001 9th grade cohort is excluded 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

[20]If students are not observed in the data 2004, they are assigned the district in which they are first observed. If students are not observed in kindergarten we use the year and grade of their first observation to compute the kindergarten cohort they would have been in. If a student repeats a grade, she remains assigned to her original cohort, to avoid endogenous changes in cohort year. We note that our results are robust to using the last district individuals were observed in, if they are not in the data in 2004. These results are available upon request.

the whole, we do not find evidence of pre-treatment trends.[21] 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.

5 Results

5.1 Black Disproportionality on Special Education Students

We begin by focusing on the effect of the Black disproportionality cap on Black SpEd students. Starting with the event study estimates, Figure 3 plots the coefficients of indicator variables for each 9th grade cohort year interacted with the pre-policy Black disproportionality rate. Consistent with our identification assumption, in each subfigure we see that cohorts in higher disproportionality rate districts enrolled in 9th grade before 2004 did not experience differential trends in their outcomes relative to students in districts with lower disproportionality rates.

For cohorts in 9th grade after the policy's introduction in 2004, Figure 3a shows a downward trend in the likelihood of continuing in SpEd in 9th grade. The cohort with the most years of policy exposure experienced the largest declines in SpEd enrollment. In the long-run, the same cohorts that experienced declines in SpEd participation experienced significant increases in educational attainment. Figure 3b shows an upward trend in high school completion for those exposed to the policy during high school, with the largest and most significant increases for those entering 9th grade after 2004. Figure 3c shows that the disproportionality cap improved college enrollment, with the largest and most significant impacts for those entering 9th grade after 2007 (who were first exposed to the policy in 7th grade).

Turning to our main table of results, Columns (1-3) of Table 2 present estimates of the impact of the Black disproportionality cap for 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 on SpEd participation at 9th grade increases as we add controls. Controlling for the classroom setting at 5th grade has

[21] 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).

the largest impact on the significance of the coefficients. Intuitively, this makes sense, as the amount of time one spends in separate classrooms can be thought of as a proxy for the severity of the disability. We expect the policy to have a larger impact on students with relatively more mild and malleable disabilities.

In the fully specified model in Column (3), for Black SpEd students a 1 p.p. increase in a district's 2004 Black disproportionality rate led to a 0.09 p.p decline in the likelihood of continuing in SpEd at 9th grade for each year of policy exposure. We scale our estimates to give an average effect size for students exposed to the policy in every year between 5th and 9th grade at the average district. To do so, we multiply the coefficient by 3.2, the average district's distance above the 1% Black disproportionality threshold at baseline and by 4, the number of years between 5th and 9th grade. This implies that the likelihood of continuing in SpEd at 9th grade fell by 1.2 p.p. (or 1.5%) for Black SpEd students. In the long-run, we find that the Black disproportionality cap improved Black SpEd student's outcomes. For those in the average district who were exposed to the policy in every year after 5th grade, the fully specified model suggests that the likelihood of completing high school increased by 1.2 p.p. (2.0%) and college enrollment increased by 1.5 p.p. (4.6%). This increase in college enrollment is driven by increases in 2-year college (rather than 4-year college).[22] Throughout, we find larger impacts on college enrollment compared to high school completion. This could be due to the fact that SpEd allows modifications to high school graduation requirements, which may make it relatively easier to graduate high school in general. If anything, we believe this would lead us to underestimate the positive impacts on high school completion. The increases in high school completion and even larger increases in college enrollment provide compelling evidence that educational attainment improved significantly.

To account for multiple inference, we also 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 completion and college enrollment (Kling, Liebman, & Katz, 2007). The results using this summary measure, shown in the bottom panel of Table 2, also indicate an improvement in the long-run outcomes of Black SpEd students. Finally, as can be seen in the second row of each table, we additionally control for the cap on the overall SpEd rate at 8.5%. Although not the main focus of this paper, we will discuss in detail the

[22]These results are available upon request.

effects of the SpEd cap on Black (and Hispanic) students in Section 5.4.[23]

In Appendix Table A.3, we investigate how the effect of the Black disproportionality cap varies by disability type. We find that the declines in SpEd participation are driven by students with specific learning disabilities (SLD) in Column (2). Since there is relatively more subjectivity in the evaluation criteria for SLD compared to more severe or physical disability types, these are precisely the students for whom we would expect to see the greatest changes in SpEd participation. For these students we also find positive long-run impacts on high school completion and college enrollment. As expected, we do not find significant impacts for those with physical impairments since it is more difficult to deny SpEd services to students with relatively more severe and objective disability types.

Appendix Table A.4 demonstrates that while the disproportionality cap led to similar declines in SpEd participation across gender, the positive long-run effects are driven by males. By income, we find statistically significant declines in SpEd participation for FRL students. The effect on non-FRL students is similar in magnitude, although not precisely estimated, which may be driven by the much smaller sample size. In the long-run, we find statistically significant increases in the likelihood of high school completion and college enrollment for FRL students, as well as increases in college enrollment for non-FRL students.[24]

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

[23]As mentioned previously and as illustrated in Appendix Table B.1, neither controlling for nor omitting the SpEd cap affects the results of the disproportionality cap.

[24]If parents' expectations of their children increased subsequent SpEd removal, they could have invested more in college preparation, such as SAT courses and extra-curricular activities. The fact that we find larger positive impacts on college enrollment among higher income students suggests household investments could have played a role. However, lower income students still benefited from SpEd removal, despite their families being less likely to be able to change their household investments, pointing to the importance of changes occurring at school, such as a reduction in stigma or less exposure to disruptive peers. We discuss mechanisms further in Section 5.1.2.

on the demographics that were statistically significantly different across districts above and below the black disproportionality threshold, as shown in Appendix Table A.5. We largely find that our results are robust to including these trends. The one exception is that the positive impact on high school completion documented in our main specification is no longer statistically significant when we include the baseline FRL rate interacted with cohort year. However, the effect size for the model including this trend (1.30% increase) is quite similar in magnitude to our baseline specification (1.96% increase).

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 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.6, we directly test whether 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 5th and 9th grade.

In the bottom panel of Appendix Table A.6 we find increases in the likelihood of being enrolled at 9th grade for FRL students, but do not find a statistically significant impact on the likelihood that non-FRL students were enrolled by 9th grade as a result of the disproportionality cap. The fact that we do not find a decrease in enrollment by 9th grade for non-FRL students provides suggestive 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), as non-FRL students are more likely to have families that have the resources to move them in response to the policy. Instead, we conclude that this reflects the fact that lower-income 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.[25] Furthermore, increases in lower-income and potentially lower-achieving 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.[26]

[25]In Appendix Table A.6, we additionally demonstrate that there were no significant changes in the likelihood of being enrolled in grades 6 or 7 for the full sample, and only a marginally significant positive impact on being enrolled at 7th grade for FRL students as a result of the Black disproportionality cap.

[26]In addition to investigating whether there is attrition from the sample, it is natural to investigate whether the policy led to changes in the likelihood of switching districts. Importantly, excessive district switching does

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 to boost the performance ratings of the SpEd subgroup, which is an incentive working in the opposite direction of the disproportionality and SpEd enrollment caps aimed at reducing access to SpEd programs in our setting.

Finally, to rule out channels other than SpEd removal that could potentially be driving the results, we investigate the extent to which districts altered their spending. In Appendix Table A.7 we estimate district-level changes in SpEd and GE spending on SpEd and GE students. Overall, we find reductions in total district-level SpEd spending (as expected after a large drop in SpEd enrollment), but do not find changes in the level of SpEd spending per SpEd students or GE spending per GE students.

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 male students and those with specific learning disabilities. Given that prior literature generally finds that SpEd participation positively impacts students, the fact that Black students benefit from SpEd removal due to

not pose a threat to identification since we assign treatment based on each student's pre-policy district. However, district switching may have important implications for the mechanisms of the effect on long-run outcomes, and thus we discuss in detail the effect of the policy on district switching in Section 5.1.2.

capping disproportionality is perhaps surprising. Ultimately, we present several pieces of evidence that point toward a story of initial misclassification among Black students who benefit from the policy-induced SpEd removal.

First, we investigate whether the effects of the policy differ based on whether districts, on average, conditionally over- or under-represent Black students in SpEd. As discussed previously, Black students in Texas (and in the U.S. in general) are on average placed in SpEd at higher rates than White students. However, previous literature has shown that after conditioning on important background characteristics, such as income and academic achievement, Black students may be under-represented in SpEd relative to White students. In theory, districts should be placing students in SpEd until the marginal cost of providing services exceeds the marginal benefit to the student. Therefore, if districts with a *conditional* over-representation of Black students in SpEd have placed students in SpEd whose marginal cost exceeds the marginal benefit, this would imply that these districts are misidentifying students for SpEd. And, removal from SpEd would improve their long-run outcomes. Likewise, if districts with a *conditional* under-representation of Black students in SpEd have placed students in SpEd whose marginal benefit exceeds the marginal cost, these students would benefit from SpEd and potentially do worse in the long-run if removed.

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. If districts with a conditional over-representation of Black students in SpEd are districts where Black students benefit from SpEd removal due to the disproportionality cap, this would provide suggestive evidence in favor of a story of Black students benefiting from SpEd removal due to initial misclassification. 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 pre-

[27]The results of this logit model are presented in Appendix Table A.8. SpEd status is predicted as of 5th grade, using baseline covariates measured as of 3rd grade. One important caveat of this analysis is that predicting SpEd participation with the limited variables in our dataset is difficult. In particular, the R-squared from our logit model predicting SpEd participation for White students is 0.221. Thus, our prediction model

dict 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.

Our estimates for the impact of the policy separately by districts conditionally over- or under-representing Black SpEd students are presented in Table 3. In line with our prior, the negative impact of the disproportionality cap on the likelihood of SpEd placement is driven by districts with a conditional over-representation of Black students in SpEd. Additionally, we find a statistically significant increase in the likelihood of high school completion and college enrollment for Black students in districts predicted to over-identify Black students for SpEd. In contrast, the impact of the disproportionality cap is not statistically significant in districts predicted to under-identify Black students for SpEd. These results suggest that the positive impacts of SpEd removal are being driven by those who were initially misidentified for SpEd.

Next, 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 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 4 demonstrates that capping Black disproportionality increased the likelihood that the Black students removed from SpEd post-policy were relatively higher performing on the reading exam and in less restrictive classroom settings (i.e. spending less than 50% of the day in resource rooms). 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

will be less optimal than that of Elder et al. (2021), who are able to link birth certificate records to educational data to improve the performance of their model.

evidence that districts were removing those who had been previously misidentified for SpEd.

Finally, we investigate whether the impacts of the policy differed by district-level teacher experience and racial composition.[28] Intuitively, it may be the case that districts with lower teacher experience may have been more likely to misclassify Black students for SpEd pre-policy, and thus, Black students in these districts might benefit the most from the pressure to reduce Black disproportionality. In Table 5 Columns (1-2) we present effects for Black students in districts whose mean level of teacher experience is above or below the statewide average level of experience of 11.7 years. Consistent with our theory, we find that the Black disproportionality cap has a somewhat larger positive impact on educational attainment in districts with below-average teacher experience. It is also possible that same-race teachers may have been less likely to misidentify Black students for SpEd. Table 5 Columns (3-4) present estimates for Black SpEd students in districts with above or below the 90th percentile proportion of Black teachers (i.e., 37.5 percent) at baseline.[29] Consistent with our theory, the effects of the Black disproportionality cap are driven by districts with below the 90th percentile proportion of Black teachers (i.e. districts with low proportions of same-race teachers).[30]

Thus far, we have presented evidence that is consistent with a story of SpEd misclassification among Black students. Although we can not directly measure the source of the misidentification given the nature of our data, 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

[28]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.

[29]The average percent of Black teachers in each district is 9.2% and the median is 3.6%. Thus, we split our sample into districts with above or below the 90th percentile proportion of Black teachers in order to have a sizable number of observations in both groups, as well as a reasonably large proportion of Black teachers in the "high" same-race teacher category.

[30]Our results are similar if we split districts into above or below the average (9.2%) percent of Black teachers, with effect sizes on college enrollment that are larger and more statistically significant in districts with below-average proportions of Black teachers. These results are available upon request.

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).

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.9 shows a statistically significant increase in the likelihood of switching districts for Black SpEd students. While we are not able to look precisely at the timing of switching districts relative to exiting SpEd, since each are only measured at one point in time per year, we find that the grade in which students move out of SpEd between 5th and 9th grade and the grade in which students switched districts roughly coincide.[31] Thus, it is difficult to determine whether students lost SpEd prior to switching. However, previous literature finds that moving districts is disruptive to student learning (Welsh, 2017; Gasper et al., 2012; Hanushek et al., 2004). Thus, we would expect increased switching to reduce long-run outcomes, which would attenuate our estimates. Indeed, as shown in Appendix Table A.10 Column (2), when we re-estimate our main specification including a control for district switching we find that the coefficient on switching predicts a reduction in high school completion and college enrollment.[32] Additionally, the effect of the disproportionality cap remains quantitatively and qualitatively very similar when we include the control for district switching. This provides compelling evidence that the positive effects of the disproportionality cap are a result of the change in SpEd status, rather than the result of switching districts.

Finally, we estimate effects on absences, suspensions, expulsions, grade repeating, and whether students took the 8th grade math and reading exams in Appendix Table A.11.[33] In the top panel, we do not find any statistically significant changes in the percent of days absent, truant, or the likelihood of being suspended or expelled. In the bottom panel, we do not find a statistically significant impact on the likelihood of repeating a grade between 5th and 9th grade. Although we do not produce estimates of performance on the standardized

[31]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 loosing SpEd and switching districts.

[32]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).

[33]Our data does not contain GPA or course grades, so we are not able to investigate these outcomes.

exams given the limitations of the exams for SpEd students (i.e., not all SpEd students take the exam and those who do often take modified/accommodated versions), we do estimate effects on *taking* the 8th grade math and reading exams. We find significant increases in the likelihood that Black students took the 8th grade math and reading exams as a result of the Black disproportionality cap. Intuitively, it makes sense that Black students would be more likely to take the exams after being removed from SpEd since they would no longer be able to qualify for test-taking exemptions. Additionally, this 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. Finally, we note that we cannot test whether SpEd removal had an impact on non-cognitive skills. In particular, reductions in misclassification could lead to increases in motivation, persistence, less stress, or better ability to focus.

5.2 Black Disproportionality on General Education Students

Next, we turn to estimating the impact of the Black disproportionality cap on GE students. While the cap directly targeted SpEd students, the policy may have also directly or indirectly affected GE students. Direct effects could be driven by reductions in the likelihood that GE students received SpEd in later grades. Indirect effects could be driven by spillovers from their peers who are no longer in SpEd.[34] These spillovers could be positive or negative. On the one hand, the increase in SpEd removal could lead to negative effects since there are no longer resources, such as teacher's aides within the GE classroom, that may have previously been available to both SpEd and GE students. On the other hand, spillovers could lead to positive effects since Black students could feel less at risk of being targeted by racially motivated policies or could have benefited from their SpEd peers improved outcomes.

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 throughout the pre-period across districts more or less treated by the Black disproportionality cap. In Table 6 Column (1), we present estimates for Black GE students.[35] For those in the

[34]About 85% of Black and Hispanic 5th grade SpEd students spend less than 50% of their day in separate classrooms. Thus, indirect effects would be operating through impacts within the classroom rather than compositional changes in who is in the GE classroom.

[35]We now include an estimate of the Hispanic disproportionality cap on Black GE students to investigate

average district exposed to the policy in every year after 5th grade, the Black disproportionality cap reduced the likelihood of participating in SpEd in 9th grade by 0.40 p.p. (9%). 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 for Black GE students. In particular, for students at the average district exposed to the policy in every year after 5th grade the Black disproportionality cap increased the likelihood of completing high school by 1.0 p.p. (1.5%) and enrolling in college by 1.3 p.p. (2.3%). This also holds up to concerns of multiple inference. In the bottom panel of Table 6 we find a statistically significant positive impact on a summary index of long-run outcomes for Black GE students.

To better understand which GE students are driving the improvements in educational attainment, we estimate effects by baseline achievement quintiles in Appendix Table A.12. 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 two quintiles. Since at least some of the positive long-run impacts for these students is likely 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 long-run educational attainment. This implies that the positive effects experienced by those in the highest quintiles will be predominately driven by spillovers, since these students did not experience impacts on the likelihood of being placed in SpEd later on. Mechanisms for these students may be driven by positive peer-to-peer spillover effects and/or a perceived 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 spillover effect of Hispanic disproportionality on Black GE students. The effect of this cap will be discussed in Section 5.3.

the effect of the Black disproportionality cap on White and Hispanic students, proportionally the next largest racial groups. We now incorporate all three treatment variables additively into our model.[36] For Hispanic GE students in Column (2) of Table 6, we find that the Black disproportionality cap led to a 0.68 p.p. (1%) increase in high school completion. There is no statistically significant effect of the Black disproportionality cap on SpEd placement for Hispanic GE students, leading us to conclude that the positive effect on high school completion is likely driven by positive peer-to-peer spillovers. For White GE students in Column (3), we do not find a statistically significant impact of the Black disproportionality cap.

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 -0.05%. 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, Figure 5 illustrates, on the whole, a lack of differential trends during the pre- *and* post-policy periods.

Columns (4-6) of Table 2 present the difference-in-differences estimates of the impact of the Hispanic disproportionality cap on Hispanic SpEd students. Across all specifications, we do not find a statistically significant effect on the likelihood of SpEd in 9th grade. In the long-run, our point estimate suggests a decrease in the likelihood of high school completion by 0.028 p.p., a very small (almost negligible) effect. In addition, we do not find a significant impact on college enrollment. The bottom panel of Table 2 demonstrates that there is very little impact of the cap on a summary index of long-run outcomes for Hispanic SpEd students. It is significantly smaller in magnitude than the positive impact we find for Black students. Furthermore, Appendix Table A.13 demonstrates the impact non-parametrically, across terciles of the 2004 Hispanic disproportionality rate. Here the impact of the cap on high school completion is no longer statistically significant, even in the most impacted districts.

Turning to Hispanic GE students, in Column (2) of Table 6 we find that a 1 p.p. increase in the 2004 Hispanic disproportionality rate for those exposed to the policy in every year

[36] Appendix Figure A.6 illustrates a lack of correlation between the 2004 district-level Black and Hispanic disproportionality rates, motivating why they are incorporated additively in our model.

after 5th grade led to a 0.06 p.p. (1.9%) decline in SpEd participation.[37] 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, consistent with the impacts we found for Hispanic SpEd students, we do not find a statistically significant impact of the Hispanic disproportionality cap on Hispanic GE students.

Again, we investigate spillover effects on other races in Table 6. The Hispanic disproportionality cap had negative impacts on high school completion for Black GE students (Column (1)) as well as negative effects on high school completion and college enrollment for White GE students (Column (3)). 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 GE teachers may be working to compensate for this loss of services among the Hispanic students in their classroom in a way that leads to negative spillover effects on other students in the classroom.

5.4 SpEd Enrollment Cap

Throughout this paper, we have controlled for the effect of 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 carefully investigating any differences across Black and Hispanic students separately in response to the SpEd enrollment cap.

[37]Since the Hispanic disproportionality is already below the 1% threshold, the effect size for students at the average district would be even smaller.

Special Education Students- Appendix Figure A.7 illustrates the event study analysis for the impact of the SpEd enrollment cap on Black SpEd students. On the whole, we find that the pre-period trends do not significantly deviate from 0. In the second row of Table 2 Column (3), we find that the the likelihood of continuing in SpEd at 9th grade fell by 3.4 p.p.[38] 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. However, our event study figures appear to show the beginning of a downward trend in long-run outcomes for cohorts most exposed to the policy, although the estimates are not statistically significant.

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 the 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. Recall from Section 5.1.2, we used a logistic regression to predict the likelihood that Black students are over-represented in SpEd compared to White students. On average, schools above the 1% Black disproportionality threshold have significantly higher (about 9 times) conditional over-representation of Black students in SpEd compared to districts above the 8.5% SpEd enrollment threshold. Although conditionally over-representing Black students in SpEd does not perfectly coincide with racial bias, this provides suggestive evidence in favor of differences in racial bias helping to explain the differences in effects across the two caps.

Turning to Hispanic SpEd students, in Column (6) of Table 2, we find that the SpEd enrollment cap led to reductions in the likelihood of continuing in SpEd by 3.2 p.p. (4.1%).[39] In the long-run, the SpEd cap reduced the likelihood of high school completion by 2.2 p.p. (3.7%) and college enrollment by 1.6 p.p. (5.6%). The SpEd enrollment cap worsened long-run outcomes for Hispanic SpEd students, likely as a result of reducing SpEd enrollment for

[38]We scale the coefficients to give an effect size for students exposed to the policy in each year after 5th grade at the average district, which was 3.2 p.p. above the 8.5% SpEd enrollment threshold in 2004.

[39]Event study estimates for Hispanic SpEd students are presented in Appendix Figure A.9.

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.

General Education Students- For Black GE students, event study figures illustrating the effect of the SpEd enrollment cap are shown in Appendix Figure A.8. In Table 6, we see for the fully exposed Black GE student at the average district, SpEd participation at 9th grade fell by 0.85 p.p. (19%), although the SpEd enrollment cap did not have a statistically significant impact on long-run outcomes. We can see suggestive evidence of the beginning of a downward trend for cohorts most exposed to the policy, particularly for college enrollment. Although the estimates are once again noisy. This could either be because the SpEd cap did not have a meaningful impact on Black students' long-run outcomes or we may be underpowered to estimate an effect. For Hispanic GE students, event study estimates for Hispanic GE students are presented in Appendix Figure A.10. In Column (2) of Table 6, we find a 0.55 p.p. (17%) decline in SpEd participation. We also find significant negative impacts on high school completion (1.4%) and college enrollment (2.6%). This is consistent with the negative impacts we find for Hispanic SpEd students in Table 2. This result reflects a combination of spillover effects from SpEd students as well as a direct effect of the reduction in the likelihood of Hispanic GE students receiving SpEd services later on.

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 the 2004-05 school year, it impacted districts differentially based on their pre-treatment disproportionality rates.

We estimate the impact of the Black and Hispanic disproportionality caps separately for

students in SpEd or GE as of 5th grade prior to policy implementation. Overall, we find that the Black disproportionality cap led to meaningful reductions in the likelihood of receiving SpEd services among Black students previously enrolled in SpEd. We find positive effects of the Black disproportionality cap on long-run outcomes for Black students in SpEd and GE. The Hispanic disproportionality cap did not have a meaningful impact on Hispanic SpEd students, since the majority of districts were already in compliance with this cap before it went in to place.

We explore several potential mechanisms behind the positive effect of the Black disproportionality cap on Black students. We find evidence consistent with a story of misclassification of Black students for SpEd in districts with high rates of disproportionality. In particular, we find that the positive effects of the disproportionality cap are driven by students in districts that over-classify Black students in SpEd, relative to White students. 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. Finally, we find that the positive impacts are concentrated in districts that may be more likely to misclassify Black students for SpEd (i.e., those with lower teacher experience and fewer Black teachers).

The impacts we find for GE students are consistent with the impacts we find for SpEd students, with Black GE students experiencing gains in the long-run resulting from the cap on Black disproportionality. Given that we find negative impacts on the likelihood that GE students were enrolled in SpEd at 9th grade, these effects likely represent a combination of direct and spillover effects. GE students themselves may be less likely to be misclassified for SpEd in later grades and GE students may benefit from from a reduction in misclassification among their peers. To the extent that the Black disproportionality cap alleviates racial bias in schools, this could help justify the positive impact for both SpEd and GE student outcomes.

Overall, our findings have meaningful implications for all public school students. Students who require SpEd services greatly benefit from them in the long-run (Ballis & Heath, 2021). 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. Ultimately, we caution against the interpretation that capping Black disproportionality is

necessarily 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.

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.

Cameron, S., & Heckman, J. (2001). The Dynamics of Educational Attaniment 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.

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.

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*.

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.

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).

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).

Markussen, E. (2004). Special Education: does it help? A study of special education in Norwegian upper secondary schools. *European Journal of Special Needs Education*, *19*(1), 33–48.

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., Frisco, M., Farkas, G., & Hibel, J. (2010). A Propensity Score Matching Analysis of the Effects of Special Education Services. *The Journal of Special Education*, *43*(4), 236–254.

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.

Parrish, T., Harr, J., Wolman, J., Anthony, J., Merickel, A., & Esra, P. (2004). State Special Educaion Finance Systems, 1999-2000 Part II: Special Education Revenues and Expenditures. *Center for Special Education Finance*.

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.

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

Redding, C. (2019). A Teacher Like Me: A Review of the Effect of Student-Teacher Racial/Ethnic Matching on Teacher Perceptions of Students and Student Academic and Behavioral Outcomes. *Review of Educational Research*, *89*(4), 499–535.

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*.

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.

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.

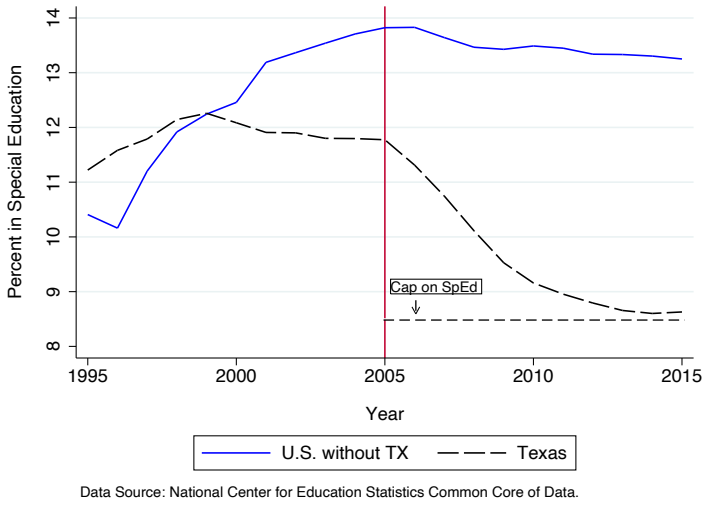
Tobias, S., Zibrin, M., & Menell, C. (1983). Special Education Referrals: Failure to Replicate Student-Teacher Ethnicity Interaction. *Journal of Education Psychology*, *75*(5), 705–707.

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

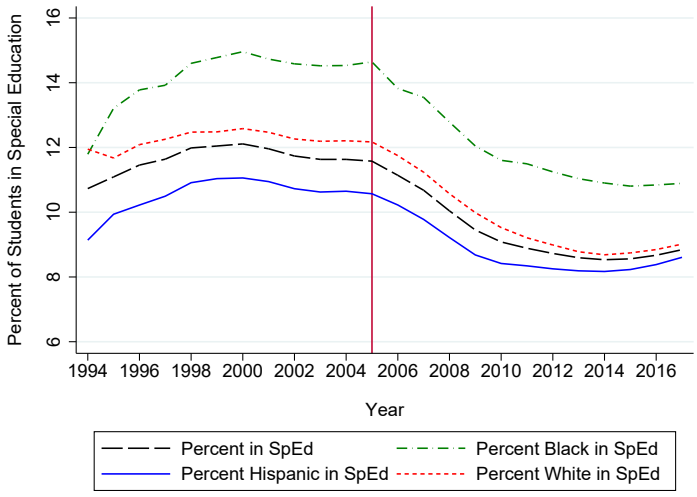
Figure 1

(a) Special Education Rate



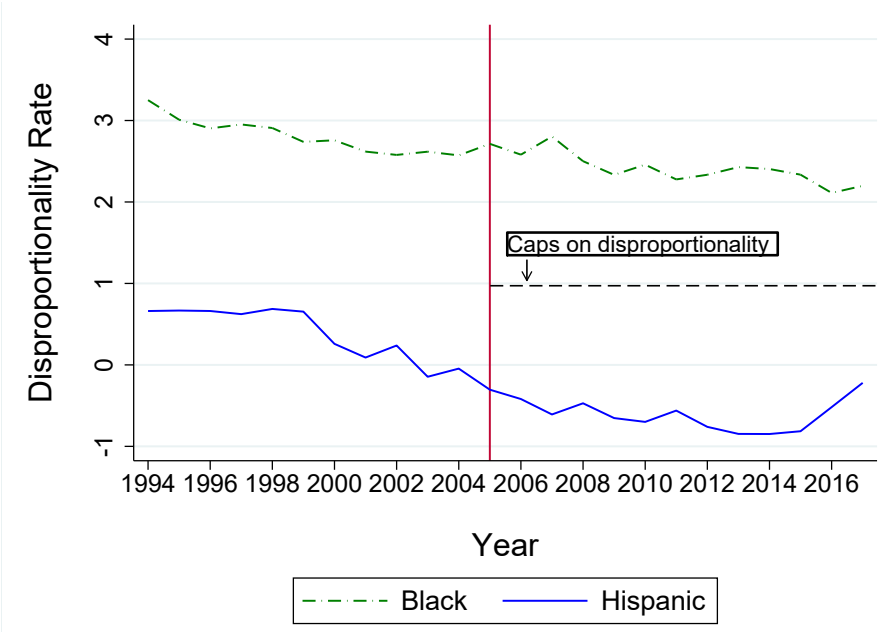
Data Source: National Center for Education Statistics Common Core of Data.

(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.

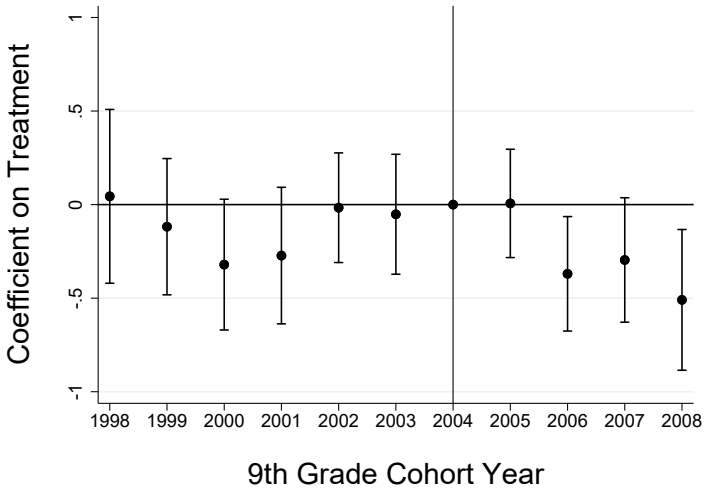
Figure 2 Disproportionality Rate for Black and Hispanic Students



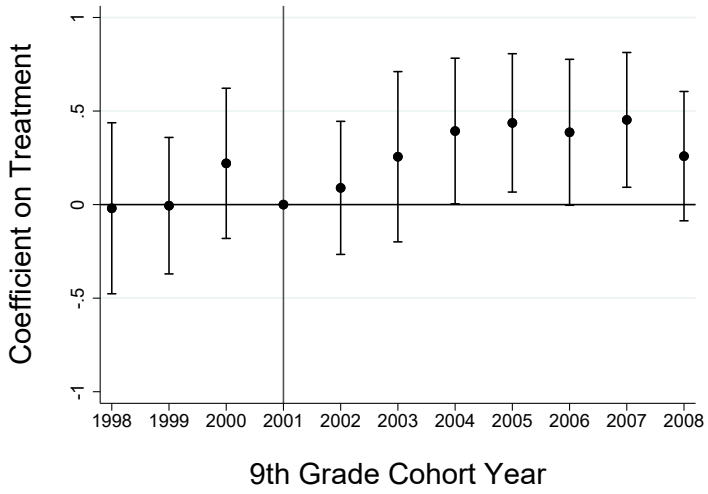
The disproportionality rate is measured as the percent of Black or Hispanic students in special education minus the percent of Black or Hispanic students in a given district.

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

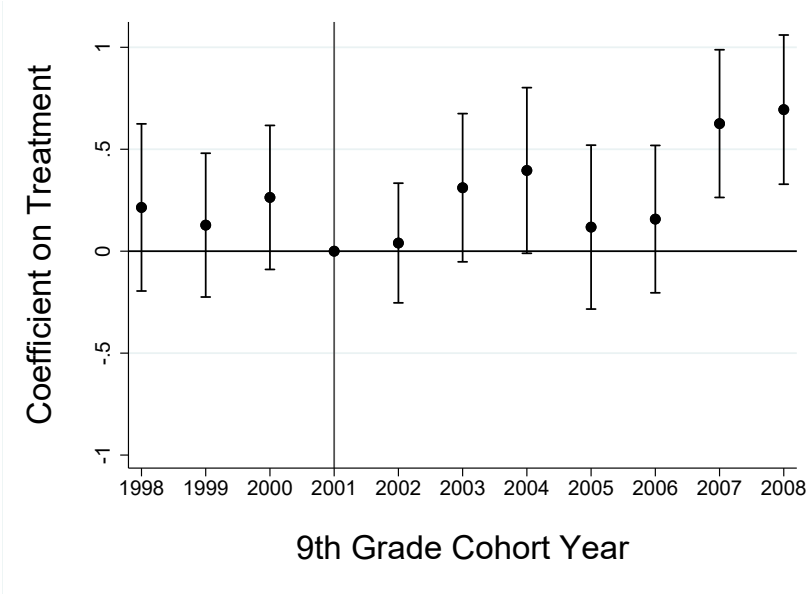
(a) Grade 9 Special Education Status



(b) High School Completion



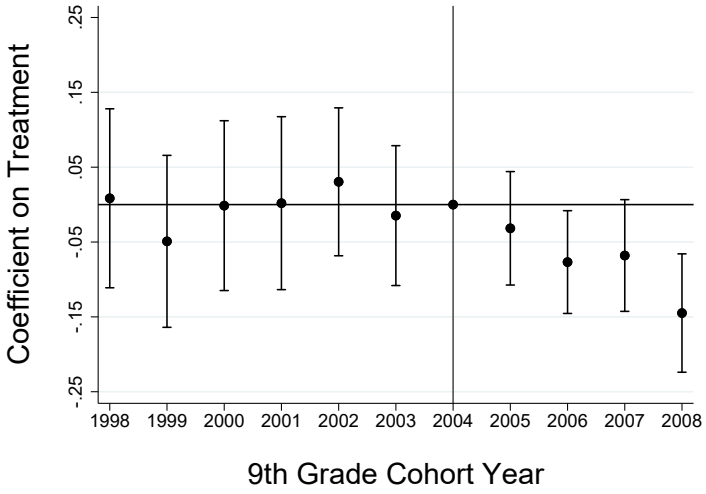
(c) College Enrollment



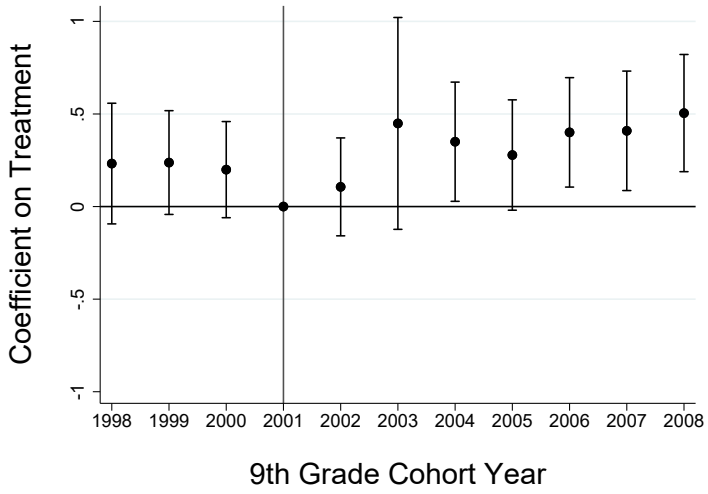
In each graph, the points represent the average district-level Black disproportionality rate in 2004 interacted with indicators for each 9th grade cohort year. The coefficients are measured in percentage points, such that the scale on the y-axis ranges from -1 to 1 percentage point. The length of the vertical bars denote the 95% confidence intervals. Regressions include controls for individual-level disability type, classroom setting, gender, FRL, ESL, gifted, and Title I status, along with district-cohort level gender, race, ESL, FRL, Title I, and gifted composition. We additionally include district and cohort fixed effects, and robust standard errors are clustered at the district level. In Panel (a) the vertical line is placed at 2004, since this is one year prior to when 9th graders are first exposed to the policy. In Panels (b) and (c) the vertical line is placed at 2001 since individuals in 9th grade in 2001 would have been in 12th grade in 2004.

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

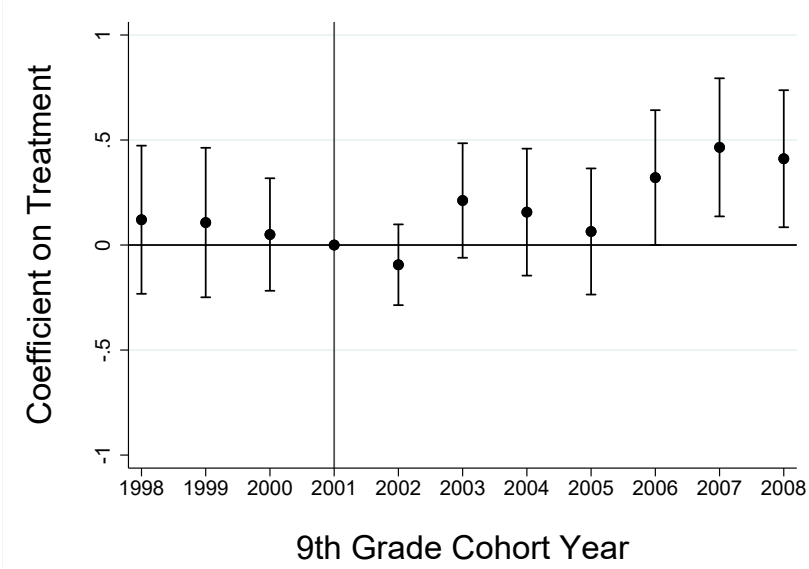
(a) Grade 9 Special Education Status



(b) High School Completion



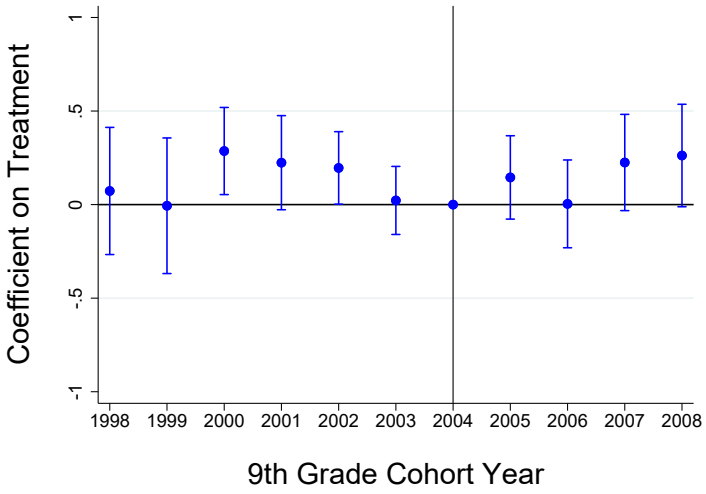
(c) College Enrollment



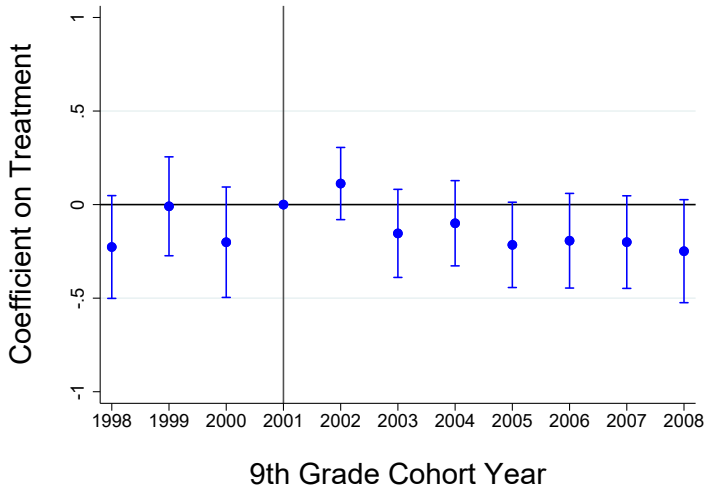
In each graph, the points denote the average district-level Black disproportionality rate in 2004 interacted with indicators for each 9th grade cohort year. The vertical bars denote the 95% confidence intervals. Regressions include controls for 5th grade math and reading exam performance, gender, FRL, ESL, gifted, and Title I status, along with district-cohort level gender, race, ESL, FRL, Title I, and gifted composition. We additionally include district and cohort fixed effects, and robust standard errors are clustered at the district level.

Figure 5 Event Study Estimates of Hispanic Disproportionality Cap for Hispanic Special Education Students

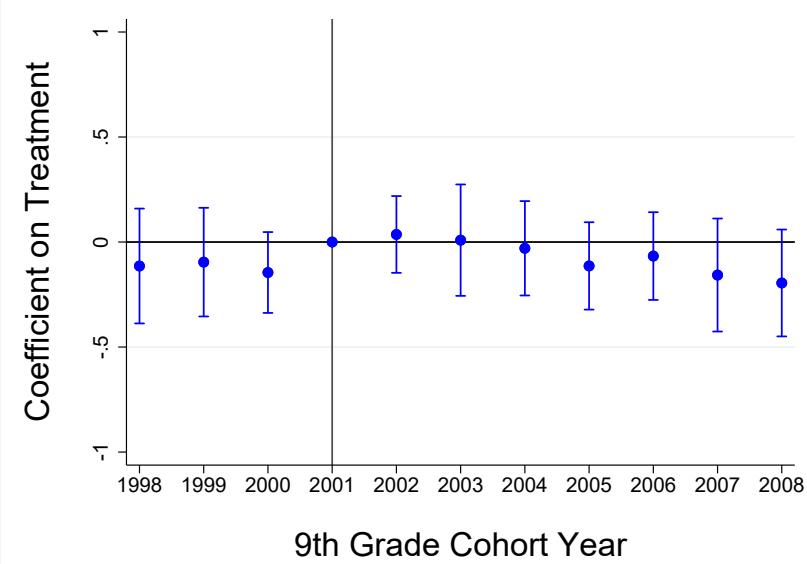
(a) Grade 9 Special Education Status



(b) High School Completion



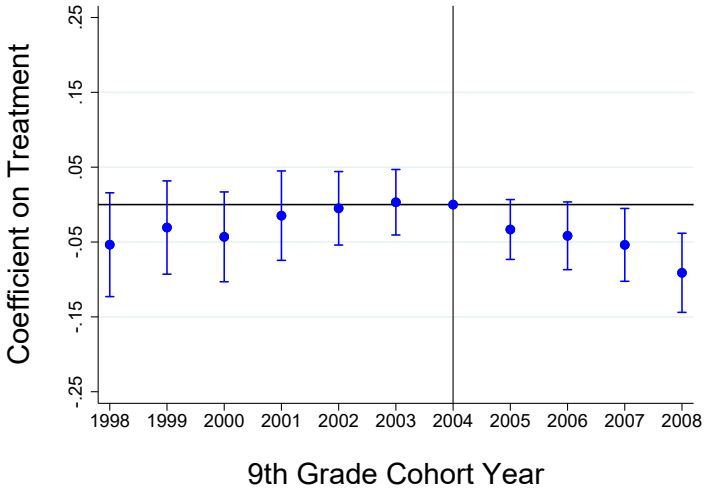
(c) College Enrollment



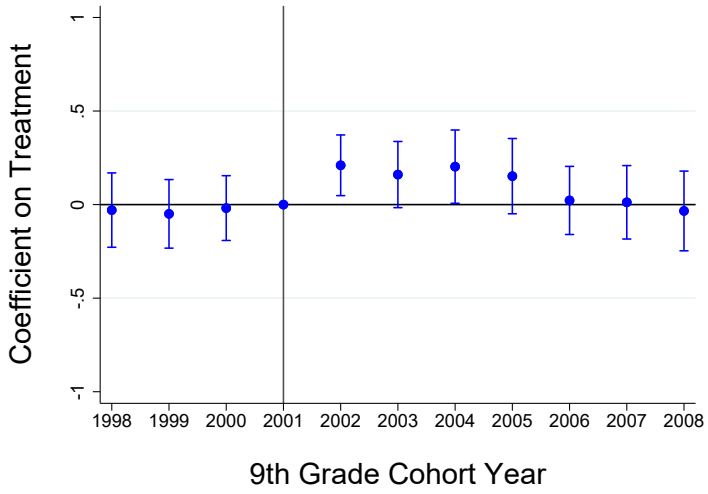
In each graph, the points denote the average district-level Hispanic disproportionality rate in 2004 interacted with indicators for each 9th grade cohort year. The vertical bars denote the 95% confidence intervals. See Figure 3 for the full set of controls used in each regression.

Figure 6 Event Study Estimates of the Hispanic Disproportionality Cap for Hispanic General Education Students

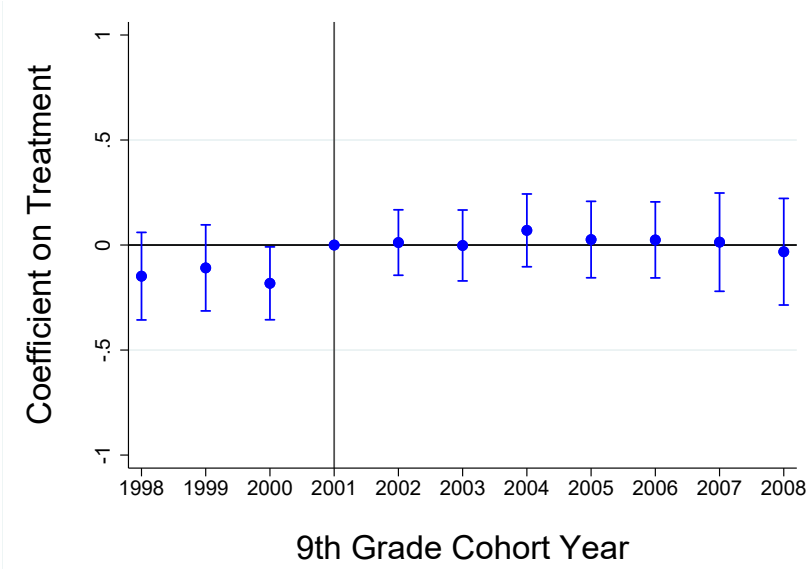
(a) Grade 9 Special Education Status



(b) High School Completion



(c) College Enrollment



In each graph, the points denote the average district-level Hispanic disproportionality rate in 2004 interacted with indicators for each 9th grade cohort year. The vertical bars denote the 95% confidence intervals. See Figure 4 for the full set of controls used in each regression.

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	0.510	0.503	0.509	0.657	0.654	0.656
FRL	0.512	0.700	0.769	0.629	0.801	0.837
ESL	0.036	0.003	0.080	0.040	0.002	0.093
Gifted	0.105	0.068	0.074	0.014	0.006	0.008
Title I	0.543	0.634	0.768	0.565	0.627	0.754
White	0.442	·	·	0.421	·	·
Black	0.140	·	·	0.182	·	·
Hispanic	0.392	·	·	0.386	·	·
Other	0.027	·	·	0.012	·	·
Took G5 Math Exam	0.812	0.768	0.756	0.412	0.283	0.356
Took G5 Reading Exam	0.804	0.761	0.744	0.362	0.245	0.294
Math G5 Z-score	0.042	-0.422	-0.143	-0.672	-1.184	-0.920
Reading G5 Z-score	0.035	-0.322	-0.226	-0.682	-1.100	-1.017
G5 SpEd Rate	0.141	0.183	0.139	·	·	·
Malleable Disability	·	·	·	0.861	0.797	0.872
$\geq$ 50% of day in GE Class	·	·	·	0.911	0.878	0.917
Long-run Outcomes						
High School Diploma	0.713	0.665	0.656	0.625	0.597	0.589
Enroll College	0.567	0.526	0.477	0.347	0.323	0.290
Enroll 2 year College	0.459	0.393	0.405	0.312	0.284	0.269
Enroll 4 year College	0.124	0.149	0.084	0.041	0.044	0.026
N	2,808,992	394,404	1,102,470	396,358	72,196	153,098

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. Enrollment in college is measured 6 years after expected high school graduation, and is not conditional on high school diploma.

Table 2 Effect of Policy on SpEd Students

	Black Students			Hispanic Students		
SpEd Status G9	(1)	(2)	(3)	(4)	(5)	(6)
$Disp_{d,2004} \times Exposure$	-0.0363	-0.0744*	-0.0931**	0.0027	0.0193	0.0264
	(0.038)	(0.039)	(0.039)	(0.029)	(0.028)	(0.030)
$SpEd_{d,2004} \times Exposure$	-0.2947***	-0.2988***	-0.2674***	-0.2855***	-0.2423***	-0.2471***
	(0.069)	(0.061)	(0.067)	(0.078)	(0.081)	(0.081)
Mean Dept Var	0.778	0.778	0.778	0.763	0.763	0.763
High School Completion						
$Disp_{d,2004} \times Exposure$	0.0475**	0.0612***	0.0533**	-0.0208*	-0.0236*	-0.0281**
	(0.024)	(0.023)	(0.022)	(0.012)	(0.012)	(0.014)
$SpEd_{d,2004} \times Exposure$	-0.0183	-0.0057	-0.0077	-0.1051***	-0.1055***	-0.0973***
	(0.034)	(0.034)	(0.037)	(0.034)	(0.034)	(0.034)
Mean Dept Var	0.597	0.597	0.597	0.589	0.589	0.589
College Enrollment						
$Disp_{d,2004} \times Exposure$	0.0494***	0.0704***	0.0670***	-0.0059	-0.0158	-0.0155
	(0.016)	(0.017)	(0.018)	(0.012)	(0.014)	(0.016)
$SpEd_{d,2004} \times Exposure$	0.0289	0.0435	0.0265	-0.0588**	-0.0659**	-0.0720**
	(0.036)	(0.038)	(0.038)	(0.028)	(0.033)	(0.035)
Mean Dept Var	0.323	0.323	0.323	0.290	0.290	0.290
Summary Index						
$Disp_{d,2004} \times Exposure$			0.1237***			-0.0466*
			(0.036)			(0.025)
$SpEd_{d,2004} \times Exposure$			0.0184			-0.1751***
			(0.062)			(0.061)
Mean Dept Var			-0.278			-0.319
Observations	72,196	72,196	72,196	153,098	153,098	153,098
Individual Controls		X	X		X	X
District-Cohort Controls			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 prior to policy implementation. $Disp_{d,2004} \times Exposure$ denotes the coefficient on the 2004 district-level Black or Hispanic disproportionality rate interacted with exposure (the number of years and individual was in school under the policy).
$SpEd_{d,2004} \times Exposure$ is the 2004 district-level SpEd rate interacted with exposure. 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, college enrollment, and associate's and bachelor's degree attainment 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. 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, including indicators for high school graduation and college enrollment. Then, we create one summary index by averaging across the standardized long-run outcomes for each individual. Regressions include 5th grade cohorts from 1994 (when the data begins) to 2004 (the year prior to policy implementation).

Table 3 Effect of Policy by Prediction of District Over- or Under-Representation of Black Students

	Black Students	
	Conditionally Over-represented	Conditionally Under-represented
SpEd Status	(1)	(2)
$Disp_{d,2004} \times Expo$	-0.1059**	-0.0688
	(0.045)	(0.059)
$SpEd_{d,2004} \times Expo$	-0.2088***	-0.2529***
	(0.072)	(0.091)
Mean Dept Var	0.803	0.760
High School Completion		
$Disp_{d,2004} \times Expo$	0.0620**	0.0379
	(0.029)	(0.032)
$SpEd_{d,2004} \times Expo$	-0.0938	0.0425
	(0.066)	(0.051)
Mean Dept Var	0.599	0.595
College Enrollment		
$Disp_{d,2004} \times Expo$	0.0862***	0.0449
	(0.024)	(0.085)
$SpEd_{d,2004} \times Expo$	-0.0116	0.0370
	(0.051)	(0.047)
Mean Dept Var	0.318	0.327
Observations	30,811	41,547

*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors are clustered at the district level. Regressions include district and cohort fixed effects, along with individual and cohort-district level controls. See Table 2 for full set of controls. The category "Over" implies Black students are over-represented in SpEd, that is, predicted to be more likely to be in SpEd relative to observationally-equivalent White peers. Likewise, "Under" implies under-representation in SpEd relative to White peers. See Appendix Table A.8 for the probability model.

Table 4 District-Level Changes in the Composition of Black Students Who Lose SpEd

	Male	**ESL**	**FRL**	**Took Math**	**Took Reading**	**Math Score**	**Reading Score**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$Disp_{d,2004} \times Exposure$	-0.0828	0.0014	-0.0187	0.1272	0.1037	-0.0869	0.2432**
	(0.070)	(0.002)	(0.048)	(0.078)	(0.076)	(0.096)	(0.122)
$SpEd_{d,2004} \times Exposure$	0.0182	-0.0029	0.0537	-0.1421	-0.1211	-0.2179	-0.2642*
	(0.082)	(0.003)	(0.065)	(0.097)	(0.091)	(0.153)	(0.151)
Mean Dept Var	-0.038	0.0002	-0.037	0.254	0.262	0.214	0.213
Observations	2,595	2,595	2,595	2,595	2,595	1,995	1,978
	RR<50%	**Malleable**	**SLD**	**Speech**	**ED**	**OHI**	**Autism**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$Disp_{d,2004} \times Exposure$	0.1102***	0.0440	-0.0006	0.0797	-0.0146	0.0108	0.0087
	(0.040)	(0.036)	(0.074)	(0.065)	(0.029)	(0.033)	(0.007)
$SpEd_{d,2004} \times Exposure$	-0.0132	-0.0949**	0.0512	-0.2304***	0.0288	0.0492	0.0312***
	(0.066)	(0.037)	(0.090)	(0.080)	(0.039)	(0.044)	(0.010)
Mean Dept Var	0.112	0.070	-0.086	0.185	-0.008	-0.020	-0.008
Observations	2,595	2,595	2,595	2,595	2,595	2,595	2,595

*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors are clustered at the district level. 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 exposure and the 2004 district-level SpEd rate interacted with exposure, 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 is a set of disability types we deem as being relatively more subjective in their evaluation criteria and include specific learning disabilities (SLD), speech impairments, emotional disturbance (ED), and other health impairment (OHI).

Table 5 Effect of Policy on Black SpEd Students by Teacher Experience and Racial Composition

	Experience		**Racial Composition**	
	Above Average	Below Average	Above 90th Percentile	Below 90th Percentile
SpEd Status	(1)	(2)	(3)	(4)
$Disp_{d,2004} \times Expo$	-0.1379***	-0.0749	-0.6033***	-0.1047***
	(0.041)	(0.070)	(0.186)	(0.039)
$SpEd_{d,2004} \times Expo$	-0.3920***	-0.1266	-0.6602**	-0.1618**
	(0.075)	(0.087)	(0.303)	(0.064)
Mean Dept Var	0.808	0.741	0.782	0.777
High School Completion				
$Disp_{d,2004} \times Expo$	0.0178	0.1092***	-0.0332	0.0409*
	(0.026)	(0.033)	(0.140)	(0.022)
$SpEd_{d,2004} \times Expo$	-0.0176	-0.0484	0.2129	0.0094
	(0.052)	(0.051)	(0.252)	(0.042)
Mean Dept Var	0.590	0.606	0.505	0.625
College Enrollment				
$Disp_{d,2004} \times Expo$	0.0726***	0.0864***	0.0260	0.0685***
	(0.023)	(0.032)	(0.147)	(0.018)
$SpEd_{d,2004} \times Expo$	0.0926**	-0.0571	-0.1767	0.0160
	(0.049)	(0.047)	(0.185)	(0.043)
Mean Dept Var	0.285	0.371	0.244	0.347
Observations	40,177	32,020	16,877	55,319

*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors are clustered at the district level. Regressions include district and cohort fixed effects. See Table 2 for full set of controls. Black students are split by whether their average district-level teacher experience is above or below the statewide average teacher experience of 11.7 years in columns (1) and (2). And Black SpEd students are split by whether the district-level composition of Black teachers is above or below the 90th percentile (i.e., 37.5%) in columns (3) and (4).

Table 6 Effect of Policy on GE Students

	Black	Hispanic	White	All Races
SpEd Status G9	(1)	(2)	(3)	(4)
$DispBlack_{d,2004} \times Exposure$	-0.0315***	-0.0067	0.0021	-0.0034
	(0.011)	(0.007)	(0.005)	(0.005)
$DispHispanic_{d,2004} \times Exposure$	-0.0171	-0.0157***	-0.0006	-0.0050
	(0.013)	(0.006)	(0.004)	(0.005)
$SpEd_{d,2004} \times Exposure$	-0.0666***	-0.0430***	-0.0402***	-0.0386***
	(0.020)	(0.009)	(0.006)	(0.006)
Mean Dept Var	0.045	0.032	0.030	0.032
High School Completion				
$DispBlack_{d,2004} \times Exposure$	0.0454***	0.0302**	-0.0067	-0.0068
	(0.016)	(0.015)	(0.008)	(0.009)
$DispHispanic_{d,2004} \times Exposure$	0.0256	-0.0031	-0.0322***	-0.0272***
	(0.022)	(0.012)	(0.007)	(0.009)
$SpEd_{d,2004} \times Exposure$	-0.0180	-0.0416**	-0.0058	-0.0195*
	(0.027)	(0.021)	(0.010)	(0.011)
Mean Dept Var	0.680	0.667	0.788	0.728
College Enrollment				
$DispBlack_{d,2004} \times Exposure$	0.0561***	0.0036	-0.0178	-0.0131
	(0.018)	(0.026)	(0.013)	(0.016)
$DispHispanic_{d,2004} \times Exposure$	-0.0473**	0.0129	-0.0308**	-0.0451***
	(0.022)	(0.019)	(0.012)	(0.015)
$SpEd_{d,2004} \times Exposure$	-0.0223	-0.0592*	0.0439**	-0.0014
	(0.038)	(0.030)	(0.019)	(0.021)
Mean Dept Var	0.571	0.507	0.687	0.603
Summary Index				
$DispBlack_{d,2004} \times Exposure$	0.1073***	0.0194	-0.0237	-0.0224
	(0.032)	(0.027)	(0.017)	(0.018)
$DispHispanic_{d,2004} \times Exposure$	-0.0202	0.0090	-0.0571***	-0.0710***
	(0.034)	(0.024)	(0.016)	(0.017)
$SpEd_{d,2004} \times Exposure$	0.0187	-0.0227	0.0313	0.0121
	(0.048)	(0.043)	(0.023)	(0.025)
Mean Dept Var	0.115	0.069	0.326	0.208
Observations	280,462	770,520	991,246	2,100,665

*** 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 general education (GE) as of 5th grade prior to the policy (in 5th grade cohorts 1994 to 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.