ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 21-364

Direct and Spillover Effects of Limiting Minority Student Access to Special Education

Briana Ballis
University of California at Merced

Katelyn Heath
Rice University

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, the impacts of the disproportionate representation of minority students in SpEd remains unclear. Using administrative data from Texas, we find that capping black disproportionality 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.

VERSION: March 2021

Suggested citation: Ballis, Briana, and Katelyn Heath. (2021). Direct and Spillover Effects of Limiting Minority Student Access to Special Education. (EdWorkingPaper: 21-364). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/m2m5-0t83

Direct and Spillover Effects of Limiting Minority Student Access to Special Education

Briana Ballis* and Katelyn Heath†

March 5, 2021

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, the impacts of the disproportionate representation of minority students in SpEd remains unclear. Using administrative data from Texas, we find that capping black disproportionality 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.

JEL Codes: I24, I26, J14, J15

*Department of Economics, University of California at Merced. Email:bballis@ucmerced.edu

†Texas Policy Lab, Rice University. Email: katelyn.heath@rice.edu.

We would like to thank Maria Fitzpatrick, Michael Lovenheim, Doug Miller, Jordan Matsudaira, Rodney Andrews, Scott Carrell, Paco Martorell, Marianne Page, and Marianne Bitler for their advice and guidance. 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

Black students are about one and a half times more likely to be receiving special education (SpEd) services in public school relative to white students (Gordon, 2017; Hosp & Reschly, 2003; Donovan & Cross, 2002; Oswald et al., 1999).[1] However, after conditioning on important confounds such as prior academic achievement and socioeconomic status, 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; Morgan, Farkas, Cook, et al., 2017; Morgan et al., 2016; Shifrer et al., 2011; Hibel et al., 2010). At the heart of this research is the concern that differences in SpEd classification between white and minority students imply that some minority students are inappropriately placed in SpEd. For example, Blanchett (2006) writes that "the disproportionate referral and placement of African American students in special education has become a discursive tool for exercising white privilege and racism." While it is true that racism could play a role in the misclassification of minority students for SpEd, estimates of the representation of minority students in SpEd are not sufficient to guide policy. Prior research is not able to provide causal evidence on the impacts of policy limiting minority student representation in SpEd, and therefore cannot speak to whether minority students are appropriately served by SpEd.

Although the effects of limiting the disproportionate representation of minority students in SpEd are unknown, policy that monitors disproportionality has been in place for many years. In 1997, the federal government amended the Individuals with Disabilities Education Act (IDEA) to require that public schools monitor significant disproportionality of minority students in SpEd (Office of Special Education and Rehabilitative Services, 2009).[2] Texas, the focus of this paper, implemented a policy in 2004 that caps disproportionality at the district

[1] Hispanic students are about equally likely to be in SpEd relative to white students.

[2] The federal government defines disproportionality as the percent of a particular race in SpEd divided by the overall SpEd rate for all races. In the re-authorization of IDEA in 2004 the federal government strengthened the pressure for districts to reduce disproportionality, since the 1997 amendments did not appear to have a significant impact on disproportionality. The re-authorization of IDEA in 2004 made disproportionality one of three priority areas for monitoring and enforcement by states (Strassfeld, 2017).

level. Texas requires that the percent of black or Hispanic students in SpEd be no greater than 1 percentage point higher than the percent of black or Hispanic students in a district overall.[3] These policies aimed at limiting disproportionality among minority students are based, at least in part, on the assumption that gaps between the classification of minority and non-minority students in SpEd are detrimental to minority student outcomes. To our knowledge, we are the first to test this assumption by causally estimating the impact of policy limiting disproportionality on minority students' outcomes in Texas.

A priori, the net benefit of limiting disproportionality is unclear. On the one hand, if minority students are misidentified for SpEd due to racial bias, then participating in SpEd could be harmful. SpEd participation could diminish achievement if it inhibits the growth and self-perceptions of students via the stigma of a disability label or by holding students back with instruction that interferes with time in the general classroom (Shifrer, 2013; Lackaye & Margalit, 2006; Bear, Clever, & Proctor, 1991). In this case, disproportionality remediation is likely to improve the long-run outcomes of minority students. On the other hand, if there is a greater need for SpEd services among minority students relative to white students, limiting their access to the individualized services that SpEd programs offer could harm student growth in school and later in life. While prior literature generally finds that marginal SpEd students benefit from SpEd services in the short-run (Hanushek, Kain, & Rivkin, 2002; Cohen, 2007; Prenovitz, 2017) and long-run (Ballis & Heath, 2021), this literature does not shed light on how minority students will be impacted by disproportionality remediation.

In this paper, we employ a dose-response difference-in-differences estimation strategy to causally estimate the effect of disproportionality remediation on minority students' long-run outcomes. This strategy exploits variation across districts in their rates of black and Hispanic disproportionality prior to policy implementation, and across cohorts in the amount of time students spent in school under the policy. Data come from the Texas Schools Project, a restricted-access administrative panel data set that allows researchers to link student-level

[3]Although disproportionality is defined differently in Texas than by the federal government, an algebraic transformation can be performed to show that Texas is technically monitoring disproportionality in accordance with the federal government's definition. However, they are doing so in a way that varies across districts based on the proportion of minority students and SpEd students in a given district.

records from the universe of Texas public K-12 students to public post-secondary school in Texas. Treatment effects are estimated separately for black and Hispanic students to determine the impact of limiting black disproportionality on black students, and the impact of limiting Hispanic disproportionality on Hispanic students. The 2004 Texas policy change also introduced a separate cap on district-level SpEd enrollment at 8.5 percent. In a separate paper, we study in depth the impacts of this SpEd enrollment cap on all SpEd students (Ballis & Heath, 2021).[4] However, we account for the SpEd enrollment cap in this paper to compare how minority students fare when access is limited through disproportionality remediation versus through policies that limit SpEd access for all students.[5]

We estimate the impact of limiting disproportionality and overall SpEd enrollment on the outcomes of black and Hispanic students already in SpEd prior to policy implementation.[6] For the fully exposed black student at the average district, receiving SpEd services as of 5th grade prior to the policy, we estimate a reduction in the likelihood of continuing in SpEd in 9th grade by about 1.6%.[7] In the long-run, we estimate increases in the likelihood of completing high school by 2.0% and enrolling in college by 4.6% for fully exposed black SpEd students. For Hispanic SpEd students, the Hispanic disproportionality cap did not have a statistically significant impact on the likelihood of SpEd participation, and we do not find robust evidence of a significant impact on long-run outcomes.

The cap on SpEd enrollment led to a roughly 4% decrease in the likelihood of remaining in SpEd at 9th grade for both black and Hispanic SpEd students. In the long-run, the SpEd

[4]As shown in (Ballis & Heath, 2021), the SpEd enrollment cap led to significant reductions in SpEd access, which generated significant reductions in educational attainment among Texas public school students.

[5]We incorporate additively the treatment variables for the SpEd enrollment cap and the disproportionality caps. We demonstrate in Section 4 a lack of correlation between treatment variables, lending motivation to incorporating treatment additively and not including an interaction term.

[6]We focus on students in SpEd as of 5th grade prior to policy implementation to avoid endogenous changes in the underlying ability distribution of students in SpEd in the post-policy period. We demonstrate in Section 5.5 that our results are robust to assigning SpEd status as of 4th or 6th grade as well.

[7]Average effect sizes for students fully exposed to the policy are computed by multiplying our estimates by the average district's distance above the 1% disproportionality cap (or the 8.5% enrollment cap, as applicable) and by the total number of years students who were fully exposed would have been in school under the policy. We estimate whether students remain in SpEd as of 9th grade, and use *expected* 9th grade, defined as SpEd status 4 years after 5th grade. The sample and variable selections are detailed further in Section 4.

enrollment cap did not have a statistically significant impact on black SpEd students. For fully exposed Hispanic students at the average district, we find a decrease in the likelihood of completing high school by 3.7%, enrolling in college by 5.6%, and earning an associate's degree by 16%. These effects are precisely estimated and imply meaningful negative impacts on long-run outcomes for Hispanic students in SpEd, as a result of capping district-level SpEd enrollment at 8.5%.

Despite the fact that the black disproportionality and SpEd enrollment caps both had a similar negative impact on SpEd participation, the disproportionality cap had a positive impact on black SpEd students, while the SpEd enrollment cap on the whole did not have a statistically significant impact on black SpEd students' outcomes (and at times had a suggestively negative impact on outcomes). This implies that there are meaningful differences across black SpEd students affected by each of the caps. Ultimately, we find that in districts with relatively higher SpEd rates, black SpEd students removed from SpEd by 9th grade were lower-achieving and less likely to have malleable disabilities at baseline.[8] This helps explain why removing these students from SpEd could potentially be harmful for long-run outcomes. In contrast, in districts with relatively higher disproportionality rates, black SpEd students removed post-policy were more likely to be higher-achieving and more likely to spend the majority of their time in General Education (GE) classrooms at baseline. If these relatively higher performing black students were misclassified for SpEd (for instance, due to racial bias), this would help justify why removing these students from SpEd improves their long-run outcomes. While we cannot discern what was driving the differences in SpEd removal decisions, we conclude that the differences in outcomes were likely driven by differences in the reasons for initial classification of black students into SpEd in districts with high levels of disproportionality compared to districts with high SpEd enrollment.

We additionally explore the potential spillover effects of the policy on GE students. The effects on GE students are broadly consistent with the direct effects on SpEd students.

[8]We define malleable disabilities as students with learning disabilities, speech impairments, other health impairments, or emotional disturbance. This is a collection of disability types which we have deemed as potentially less severe and/or more subjective in their evaluation criteria compared to other disability types.

Consistent with Ballis and Heath (2021), we find small negative impacts on high school completion and associate's degree attainment for GE students of all races resulting from the SpEd enrollment cap. This is consistent with the intuition that GE students do worse in school when greater numbers of SpEd students (who were for the most part already in the GE classroom for the majority of the school day) are no longer supported by additional services, such as teacher's aides within the GE classroom. Interestingly, removing black students from SpEd as a result of capping disproportionality has positive impacts on both SpEd and GE black students. These results represent a combination of direct and spillover effects on GE students, since we estimate significant declines in the likelihood that GE students themselves are in SpEd at expected 9th grade. Although we are not able to directly measure the extent to which these long-run impacts reflect direct vs. spillover effects, our results are consistent with a reduction in perceived racial bias in schools on the part of black GE students and/or an improvement in GE classroom instruction on the part of GE teachers.

Overall, our results suggest that while some minority students benefit from SpEd services, other students are worse-off in the long-run, potentially as a result of misclassification for SpEd. The cap on black disproportionality improved long-run outcomes for black SpEd students, and may be effective at reducing the misclassification of black students in SpEd. The policy also has meaningful impacts on GE students, whereby GE students' long-run outcomes are affected by whether minority students are appropriately placed in SpEd. We caution against the interpretation that disproportionality caps are the best intervention for reducing racial gaps in SpEd placement, and conclude that 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.

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)). This legislation introduced for the first time the requirement that schools provide a "free and appropriate" public education for all students regardless of physical or cognitive disability. Prior to this legislation, students with disabilities were often not served or not served appropriately in public school. In 1970, it is estimated that public schools educated only about 20% of children with disabilities (Office of Special Education and Rehabilitative Services, 2010). Services for students with disabilities provided under IDEA are now commonly referred to as SpEd services.

In order to qualify for SpEd under IDEA, students must fall within at least one of thirteen disability categories, which include autism, emotional disturbance, specific learning disability, other health impairment (which includes ADHD), and various physical disabilities (Reschly, 1996). To be evaluated for SpEd, a student is typically referred by a parent or teacher. After the initial referral, the student is evaluated via a series of tests to determine what, if any, disability he has and whether his disability adversely affects his educational performance. If a student is deemed eligible, an Individualized Education Plan (IEP) is written for them by a team of professionals, including both special educators and GE teachers, in addition to the student's parent(s) or legal guardian(s). The IEP states exactly 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 both the student's disability and the school they attend. This may include a teacher's aide in the classroom, direct instruction in small groups with a special educator or speech language pathologist, or direct services or consultation from other service providers such as occupational/physical therapists and social-emotional learning specialists. IEPs are reviewed at least once a year, and students are typically re-evaluated every three years to determine whether they still meet the eligibility requirements

for SpEd (Office of Special Education and Rehabilitative Services, 2000).

As previously noted, policy that monitors minority student representation in SpEd has been in place for many years due to the concern that some minority students are inappropriately placed in SpEd. The U.S. Department of Education began requiring that school districts monitor disproportionality of minority students in SpEd in its re-authorization of IDEA in 1997, and strengthened this requirement in 2004 by making it one of three priority areas (Strassfeld, 2017). Districts must also report whether such disproportionality is the result of inappropriate identification (Office of Special Education and Rehabilitative Services, 2009). The threshold for what constitutes "significant" disproportionality is left up to states to decide (Office of Special Education Programs, 2017). If a state decides a district has significant disproportionality, that district must allocate part of their federal SpEd funding to improving early intervention services for students with disabilities six years or younger (Office of Special Education and Rehabilitative Services, 2009).

2.2 Policy Environment in Texas

In the summer of 2004, the Texas Education Agency (TEA), introduced the Performance Based Monitoring Analysis System (PBMAS) (Texas Education Agency, 2016b). This system monitors three groups of students: Special Education, Bilingual/English as a Second Language, and Migrant students. For each group of students there is a set of outcomes that are monitored at the district level. Districts are assigned a *performance level* based on how they are performing relative to state standards for each of the monitored outcomes. If enough outcomes fall below a certain performance level, a district is *staged for intervention*, meaning they must develop a plan for improving their ability to meet adequate performance levels in subsequent school years (Texas Education Agency, 2016a). If districts are staged for intervention several years in a row and/or their performance levels are well below the compliance thresholds, the consequences can escalate from improvement plans, to on-site visits and third party consultations intended to provide feedback to districts on how to improve

in the future.[9]

One such monitored outcome for SpEd students under this policy was the requirement that districts have a disproportionality rate of 1 percent or less to be in compliance with state standards. The disproportionality rate is defined as the percent of black or Hispanic students in SpEd minus the overall district percent of black or Hispanic students.[10] 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. Hereafter, we refer to these thresholds as the black and Hispanic disproportionality caps. A second important outcome monitored under this policy was the district SpEd rate. This part of the policy required that districts have at most 8.5 percent of students in SpEd to be in compliance with state 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. We hereafter refer to this threshold as the SpEd enrollment cap.

The threshold limiting SpEd enrollment was not widely publicized until an article exposing the policy was published in the Houston Chronicle in 2016 (Rosenthal, 2016). Much public debate ensued after the publication of this article, and it sparked an investigation by the Federal Department of Education that took place in February 2017 (Office of Special Education and Rehabilitative Services, 2017). In May 2017, the Texas Legislature passed a bill banning the use of targets in SpEd enrollment, and the cap has since been removed from practice. In January 2018, the Federal DOE released the findings of its investigation and concluded that the TEA had failed to comply with the federal law IDEA (U.S. Department of Education, 2018). As a result, the TEA issued a corrective action plan in April 2018 to retroactively and proactively address its noncompliance with IDEA (Texas Education Agency, 2018). The thresholds monitoring disproportionality were not highlighted in the Houston Chronicle article, and to our knowledge were not investigated or deemed illegal by

[9]The PBMAS policy was created in the summer of 2004 and the 2004-2005 school year was the first year of the policy (Texas Education Agency, 2004).

[10]To our knowledge, prior to 2004 Texas did not have a systematic method for determining whether a district had "significant" disproportionality.

the federal government. As previously discussed in Section 2.1, the federal government itself requires public schools monitor disproportionality under IDEA.

To demonstrate the impact this policy had over time, we graph the percent of students in SpEd in Texas relative to the rest of the U.S. in Figure 1. Prior to the policy's implementation, in the 2003-2004 school year, the statewide average percent of students in SpEd was about 14 percent. As of the 2016-2017 school year, the statewide district-level average had fallen to 9 percent. This is in contrast to the national average percent of students in SpEd, which remained approximately steady at around 13.5% from 2004 to 2016. In Figure 2a, we show district-level averages of the percent of students in SpEd in Texas overall and by race. Again, we see a dramatic decrease in the rate of SpEd enrollment after 2004. In Figure 2b, 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 2005, the statewide average Hispanic disproportionality rate was already below 0.[11]

In addition to the outcomes described above, this policy monitors other outcomes related to performance of SpEd students. In Ballis and Heath (2021), we show that the majority of districts were already meeting, or nearly meeting, these other thresholds prior to policy implementation. We find no evidence that these indicators are biasing our results of the impacts of the disproportionality and SpEd caps on our sample of SpEd students. 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, districts were significantly less likely to respond to these thresholds.[12]

[11] In 2004, about 50% of districts were already meeting the threshold for black disproportionality, whereas 63% of districts were meeting the threshold for Hispanic disproportionality.

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

3 Data

Data for this paper come from the Texas Schools Project (TSP) housed at the Education Research Center at the University of Texas at Dallas. This restricted-access administrative data provide researchers the ability to link individual-level information from public school records from the Texas Education Agency to public post-secondary school information from the Texas Higher Education Coordinating Board. We merge these data together to obtain a panel data set from 1994 to 2017 containing a rich set of individual-level background characteristics. Our final sample includes 72,197 black students in SpEd at 5th grade and 153,098 Hispanic students in SpEd at 5th grade.[13] Table 1 presents descriptive statistics for all students, black students, and Hispanic students, as well as all SpEd students, Black SpEd students, and Hispanic SpEd students. About 10.3% of students are in SpEd over the full time period of our data and 45% of students are Hispanic. Black students have a higher SpEd rate at 12.9% relative to Hispanic students at 9.4%.[14]

In this paper, we do not estimate effects on math and reading exam performance for SpEd students. Students in SpEd are often exempt from the exams or take modified or accommodated versions of the exams.[15] Losing SpEd services is likely to reduce test scores mechanically as a result of no longer having access to modified or accommodated versions of the exam. In addition, modified and accommodated versions of the exams were not offered until 2001 and are not available in our data until 2008. Therefore, we do not expect the selected test scores of only those SpEd 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

[13]These data are particularly advantageous for a study focused on minority students in SpEd, as these tend to be very small samples in survey data. Our administrative data contain roughly 14.4 million unique individuals in public elementary and secondary school in Texas between 1994 and 2017.

[14]Appendix Table A.1 illustrates the proportions of individuals with each disability type by race. Among black, Hispanic, and white students, learning disabilities is the most common disability type, followed by speech impairments. Appendix Table A.2 shows the proportions of students with each disability type before and after policy implementation in 2005. After the policy was implemented, learning disabilities in particular declined significantly.

[15]Beginning in 2001, SpEd students are able to take accommodated and/or modified versions of the standardized exams. In certain cases, these exams cover lower than grade-content material, so that these exams are not equivalent to those offered to GE students.

students.

Instead, we focus on long-run outcomes, which include an indicator for whether an individual graduated from high school, attended a post-secondary institution in Texas, and obtained an associate's or bachelor's degree.[16] 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.[17] For post-secondary outcomes, we do not condition on high school graduation, and each outcome is censored such that individuals have 6 years after expected high school graduation to enroll in college and obtain an associate's or bachelor's degree.

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. In 2008 and 2009 only 3.7% of students attended college out of state (compared to 64.5% who attended in-state) (Mountjoy, 2021) and from 2008 to 2012 only 1.7 percent of SpEd students enrolled in college out of state within two years of their high school graduation (Ballis & Heath, 2021).[18] Finally, although earnings are available in the TSP data, the policy change occurs too close to the end of our earnings data to provide accurate estimates of changes in earnings. Ideally, we would like to estimate earnings 10 years after expected high school graduation in order to avoid changes in rates of college-going that could bias the results.[19] Thus, we leave for future work estimates of the impact of the policy on changes in earnings in the labor

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

[17]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.

[18]These estimates are limited to cohorts in the Texas Schools Project data that can be linked to the National Student Clearinghouse data, in order to estimate the fraction of Texas high schoolers who attend college out of state.

[19]Since the policy took place in 2004, this means that the most recent earnings estimate, measured in 2017, only covers up to high school cohort 2007. This cohort was only affected by the policy in 10th grade and later, which is particularly late given that we are considering changes in SpEd status which most often take place in grades K through 5.

market.

In Tables 2 and 3, we illustrate raw differences in the characteristics of districts that are above and below the 8.5% threshold for SpEd enrollment, and above and below the 1% thresholds for the disproportionality caps among black and Hispanic students in 2003-2004, one year prior to policy implementation. In Table 2, we see that districts above the 8.5% SpEd threshold have more white students, fewer black and Hispanic students, and have somewhat higher test scores on the math and reading exams, relative to districts below the SpEd enrollment threshold. In Table 3, we see that districts above the 1% threshold for black disproportionality had fewer Hispanic students and more black students. Most other 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 cap. There are more Hispanic students and fewer black students in districts above the Hispanic disproportionality cap.[20]

4 Empirical Strategy

We estimate the causal impact of reducing disproportionality by exploiting cross-district and cross-cohort variation in exposure to the black and Hispanic disproportionality caps. We employ a dose-response difference-in-differences estimation strategy to determine whether students in districts with higher rates of black or Hispanic disproportionality at baseline experience larger gains or losses in outcomes. We estimate effects separately for black and Hispanic students. In addition, we utilize cross-district and cross-cohort variation in exposure to the SpEd enrollment cap. This allows us to compare how minority students fare when access is limited through disproportionality remediation versus policies that limit SpEd access for all students. In each specification we include two treatment variables: (1) either the

[20]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 5.5 we demonstrate that our results are robust to controlling for district-level time-trends in the baseline levels of the demographic variables.

2003-2004 black or Hispanic disproportionality rate (depending on whether effects are being estimated for black or Hispanic students) interacted with the number of years an individual is in school under the policy; and (2) the 2003-2004 SpEd rate interacted with the number of years in school under the policy. We choose the 2003-2004 school year (hereafter referred to as 2004) since this is one year prior to policy implementation.[21]

Appendix Figures A.4, A.5, and A.6 further illustrate the intuition behind our treatment variables. In Appendix Figure A.4, we sort districts by their 2004 SpEd rate. The bottom series, denoted with circles, shows the average SpEd rate from 1994 to 2017 for districts already below the 8.5% threshold for SpEd enrollment in 2004. In the three top series, districts are split into terciles based on their 2004 SpEd rate, conditional on having a SpEd rate greater than 8.5%. Comparing the top most series, denoted with x's, to the bottom series illustrates that districts with the highest rates of SpEd made the largest reductions across the post-period in their SpEd rates, indicating that they are more treated by the policy relative to those already meeting or nearly meeting the threshold. In Appendix Figures A.5 and A.6 we present analogous figures for the district-level black and Hispanic disproportionality rates, respectively. In these figures, we see a somewhat similar pattern in the topmost series, illustrating that districts with the highest rates of disproportionality made the greatest reductions to their disproportionality rates in the post-period. However, we now find much less response in the first and second terciles above the 1% threshold.

We estimate effects for two groups of black and Hispanic students: (1) SpEd students and (2) GE students.[22] Given the nature of the policy change, we are not able to causally estimate the effect of the policy by simply comparing SpEd student outcomes before and after policy implementation. To reduce their SpEd rate, districts must decide which students

[21] Our results are robust to instead using the average disproportionality rate or SpEd rate in the pre-period, and these estimates are available upon request.

[22] In theory, we could also estimate results for an aggregate sample of SpEd and GE students. The goal of the aggregate sample would be to include an extended number of cohorts, since it does not need to be restricted to students in 5th grade prior to policy implementation. However, the youngest cohort that we could include was in 5th grade in 2004 (given that our data goes through 2017 and we need 6 years post-expected high school graduation to measure the long-run outcomes). Thus, our aggregate sample would use the same number of cohorts included in the SpEd and GE samples, and does not provide additional or new information.

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. Thus, we estimate the effect of limiting access to SpEd for students already identified for SpEd prior to policy implementation. To do so, we select students who were in SpEd as of 5th grade prior to policy implementation. 5th grade is a reasonable choice since most SpEd enrollment decisions take place prior to 5th grade.[23] To estimate spillover effects on GE students, we employ a similar strategy by estimating effects for students in GE as of 5th grade prior to policy implementation.[24]

To identify the impact of limiting disproportionality on minority SpEd and GE students, we estimate the following difference-in-differences specification:

$$Y_{idc} = \beta_0 + \beta_1 Disp_{2004,d} * Exposure_c + \beta_2 SpEd_{2004,d} * Exposure_c + \beta_3 X_{idc} + \eta_d + \theta_c + \varepsilon_{idc} \quad (1)$$

where Y_{idc} is the outcome of interest for individual i, who attended school in district d, in cohort c. We estimate the impact of the policy on the likelihood of persisting in SpEd by expected 9th grade and on the long-run outcomes of secondary and post-secondary attainment.[25] Our first treatment measure is the interaction of $Disp_{2004,d}$, the 2004 district black or Hispanic disproportionality rate and $Exposure_c$. Disproportionality is measured as the district SpEd black or Hispanic percentage minus the district overall black or Hispanic percentage. For the SpEd status outcome, $Exposure_c$ is the number of years the cohort was exposed to the policy between 5th and 9th grade.[26] For high school graduation and post-secondary outcomes, $Exposure_c$ is the number of years each cohort is expected to be in school between 5th and 12th grade under the policy. The second treatment measure is

[23] Appendix Figure A.7 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. However, when we use students in SpEd as of 4th or 6th grade prior to policy implementation instead, results remain qualitatively and quantitatively similar. These estimates are presented in Appendix Table A.3.

[24] These results are also robust to choosing 4th or 6th grade, and are presented in Appendix Table A.4.

[25] We chose 9th grade since this is prior to when most dropout decisions are made. This is measured as expected 9th grade, that is, 4 years after 5th grade in order to avoid endogenous changes in grade repeating.

[26] The number of years exposed varies across cohorts. Since cohort is defined net of endogenous changes in grade-repeating, exposure is based on the expected number of years in school under the policy, rather than actual years to avoid endogenous changes in exposure.

the interaction of $SpEd_{2004,d}$, the percent of students in SpEd in 2004 in each district and $Exposure_c$.[27] Models are run separately for black and Hispanic students, such that the black disproportionality rate is included in models run on black students and the Hispanic disproportionality rate is included in models run on 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.[28] When estimating results for the SpEd sample, we additionally include controls for baseline disability type and an indicator for whether the student spent greater than 50% of the day in a GE classroom. When estimating results for the GE sample, we additionally control for 5th grade math and reading standardized exam scores. We also include district fixed effects, η_d, and cohort fixed effects, θ_c.[29] Students are assigned the district in which they are observed in 2004, the year prior to policy implementation, and their cohort year corresponds to the year they were in kindergarten.[30] The coefficients of interest in these regressions are β_1 and β_2. These are the difference-in-differences estimates of the effect of reducing disproportionality and SpEd enrollment, respectively, among black or Hispanic students.

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 either black or Hispanic students) provide an accurate counterfactual for trends

[27]The treatment terms appear in the model additively since as Appendix Figures A.8a and A.8b show, there is no correlation between the two treatment variables. We graph each district's SpEd rate in 2004 on the x-axis, and each district's black or Hispanic disproportionality rate on the y-axis. The correlation coefficient between the two treatment variables in Figure A.8a is 0.0022 and in Figure A.8b is 0.0310.

[28]English as a Second Language is the term previously given to the program now referred to as English Language Learner. In this paper, we use the language that is consistent with the language in the Texas Schools Project data.

[29]Our specifications are robust to including school fixed effects, rather than district fixed effects, and these results are available upon request. Additionally, standard errors are clustered at the district level, since this is the level at which treatment varies.

[30]If students are not observed in the data 2004, they are assigned the district in which they are first observed. If the 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 also 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.

among districts with high disproportionality rates. Likewise, trends in outcomes among districts with a low SpEd rate provide an accurate counterfactual for trends among districts with a high SpEd rate. We test these assumptions directly by implementing an event study analysis. To do so, we create indicator variables for each cohort and interact each of these indicator variables with $Disp_{2004,d}$ and $SpEd_{2004,d}$. We then graph the coefficients of each of these interactions to present a visual of the trends in β_1 and β_2 over time. The results of the event study analysis are presented in Section 5.[31]

For our specifications to be identified it must also be the case that there are no contemporaneous shocks correlated with outcomes. The only 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 a statewide accountability system in place that had been implemented under President Bush when he was governor of Texas. The main difference between Texas' state accountability system and NCLB is that NCLB monitored the performance of SpEd students as their own subgroup on the standardized exams.[32] However, the vast majority of districts (97%) were already meeting the standardized 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 SpEd enrollment and disproportionality caps in our setting.

[31]We also have strong anecdotal evidence against pre-treatment trends, since the policy was not widely known to the public until 12 years after its implementation.

[32]While NCLB did not monitor black SpEd students or Hispanic SpEd students separately, it likely still contributed to an incentive to improve the performance of black and Hispanic students in SpEd.

5 Results

5.1 Direct Effects on SpEd Students

In Figures 3a and 3b we present the results of our event study analysis on SpEd status as of expected 9th grade (hereafter referred to simply as 9th grade) for black and Hispanic students who were enrolled in SpEd as of 5th grade prior to the policy. On the x-axis we plot the 9th grade cohort year, and on the y-axis we plot the coefficients of indicator variables for each cohort year interacted with either the 2004 district disproportionality rate or SpEd rate. The effect of the district disproportionality rate is depicted in orange, and the effect of the SpEd rate is depicted in blue. The coefficients in our models are measured in percentage points, such that the scale on the y-axis for these graphs range from -1.5 to 1 percentage point. The shaded regions denote the boundaries of the 95% confidence intervals.

Figures 3a and 3b demonstrate that there was no impact of the black or Hispanic disproportionality cap (shown in orange) on 9th grade SpEd status for black or Hispanic students throughout the pre-policy period. The trends in the black and Hispanic disproportionality caps remain flat and centered around 0 throughout the pre-period. Turning to the effect of the SpEd enrollment cap in each figure (shown in blue), we see that SpEd participation appears to trend upward in the pre-period until 2001, although this effect is, for the most part, not statistically distinguishable from 0 at conventional levels.[33] Additionally, a positive trend in the pre-period in the likelihood of being in SpEd at 9th grade would imply that more treated districts (with higher 2004 SpEd rates) were on a trend of increasing the likelihood of remaining in SpEd during the pre-policy period. This trend is in the opposite direction of the effects we estimate in the post-period. Reassuringly, this trend also flattens out for both black and Hispanic students before the policy was implemented, between 2001 and

[33]This upward trend is likely a result of the accountability and finance policies in place in Texas during this time period. Between 1994 and 1998 SpEd students did not count toward districts' accountability ratings, creating an incentive to increase SpEd enrollment. In addition, Texas' SpEd finance system between 1992 to 1997 created incentives that led to an increase in enrollment (Cullen, 2003). Our event-studies are robust to dropping years 1998 to 2001 (which correspond to students in kindergarten between 1990 and 1995), and are available upon request.

2003. After 2005 we see a distinct downward trend in the likelihood of being in SpEd due to the black disproportionality cap, and even larger declines in SpEd enrollment for black and Hispanic students due to the SpEd enrollment cap. However, we do not see a downward trend in the likelihood of being in SpEd as a result of the Hispanic disproportionality cap, as it remains relatively flat and centered around 0 in the post-period.

Figure 4 presents the results of the event study analysis for the long-run outcomes for the 5th grade SpEd sample. These graphs are organized in the same way as before. The effect sizes for the impact of the SpEd cap and the disproportionality caps in the years leading up to the policy are, for the most part, statistically indistinguishable from 0.[34] During the post-period we do not find statistically significant impacts of the policy in the first few cohorts after 2001 with the least amount of exposure. This aligns with the timing of Figures 3a and 3b. Students in the first 9th grade cohort after 2001 were only exposed to the policy for one year, during their last year of high school. In the event study figures showing high school completion and college enrollment we see the negative impacts of the policy beginning around 2004, which was the first cohort fully exposed to the policy between 9th and 12th grade.

In Table 4, we present estimates of the effect of the disproportionality caps and the SpEd cap on outcomes for black and Hispanic students who were in SpEd as of 5th grade prior to policy implementation. In columns (1) through (3) we present effects for black students. In column (1), we show estimates with district and cohort fixed effects only, in column (2) we add individual-level controls, and in column (3) we add both individual-level and district-cohort level controls. Columns (4) through (6) present the same set of estimates for Hispanic students. Our preferred specification for black students is in column (3) and our preferred specification for Hispanic students is in column (6).

We first turn to the estimates of the impact of the disproportionality caps on the likelihood of continuing in SpEd. In Table 4 column (3), we find that a 1 percentage point

[34]The one exception is in Figures 4e and 4f, where we see a small positive effect in the likelihood of obtaining an associates degree as a result of the SpEd cap in 2002 for both black and Hispanic students. However, this effect size is very small and in the opposite direction of the negative effects we estimate in the post-period.

increase in a district's 2004 black disproportionality rate led to a 0.09 percentage point decline in the likelihood of continuing in SpEd at 9th grade for black SpEd students. We scale our estimates to give an effect size for the fully exposed student at the average district by multiplying the coefficient by the average distance above the 1% disproportionality threshold in 2004 (which was 3.2) and by the number of years students fully exposed to the policy were in school between 5th and 9th grade. This implies, for fully exposed black SpEd students, the likelihood of continuing in SpEd at 9th grade fell by 1.21 percentage points as a result of the cap on black disproportionality. This is a somewhat small impact, representing a 1.6% decrease, given that 77.8% of black 5th grade SpEd students were still enrolled in SpEd as of 9th grade. For Hispanic SpEd students, we do not find a statistically significant impact of the Hispanic disproportionality cap on the likelihood of SpEd in 9th grade.[35]

In the long-run, we find that the black disproportionality cap improved black SpEd student's outcomes, although effect sizes are again quite small.[36] For black SpEd students enrolled in the average district who were fully exposed to the policy between 5th and 12th grade, the likelihood of completing high school increased by 1.2 percentage points (2.0%) and the likelihood of enrolling in college increased by 1.5 percentage points (4.6%), as a result of the cap on black disproportionality. In order 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, college enrollment, and college completion (Kling, Liebman, & Katz, 2007). The results using this summary measure, shown in Table 5, also indicate an improvement in the long-run outcomes of Black SpEd students due to the black disproportionality cap.[37]

[35]As previously noted, the statewide district-level average Hispanic disproportionality rate was already below the 1% threshold in 2004, at about -2.9 percentage points. This may account for why we do not find statistically significant impacts of this cap on SpEd participation for Hispanic students.

[36]Although the estimated effects are small, we would like to highlight that these results represent intent-to-treat (ITT) estimates of the impact of the black disproportionality cap (and SpEd cap) on black students in SpEd at 5th grade. It is highly likely that the impact of this policy is much larger on students whose SpEd status changes as a result of the policy. In fact, we show in Ballis and Heath (2021) that treatment-on-the-treated (TOT) estimates of the SpEd cap do produce much larger results for the impact of the SpEd cap on marginal SpEd students whose SpEd status changes as a result of the policy.

[37]Appendix Table A.5 presents results for the impact of the disproportionality caps non-parametrically, by splitting districts into categories based on their 2004 levels of black or Hispanic disproportionality. Overall,

For Hispanic SpEd students, on the whole we do not find a large significant impact of the Hispanic disproportionality cap on long-run outcomes. This can likely be attributed to the fact that the Hispanic disproportionality cap was much less binding, as the majority of districts were already meeting this threshold when it went into effect. The one exception we find is a small negative impact of the Hispanic disproportionality cap on high school completion by 0.36 percentage points (0.61%).[38] However, this result is not not robust to including baseline demographic trends (discussed in more detail in Section 5.5). Moreover, results shown in Table 5 demonstrate that there is no impact of the Hispanic disproportionality cap on a summary index of long-run outcomes for Hispanic SpEd students.[39] Thus, we view these results as only suggestive that reducing disproportionality for Hispanic students may harm long-run outcomes.

Turning to the cap on SpEd enrollment in Table 4, we find that a 1 percentage point increase in a district's 2004 SpEd rate is predicted to decrease the likelihood of remaining in SpEd during 9th grade by 0.27 (0.25) percentage points for each additional year a black (Hispanic) SpEd student is in school under the policy. For the SpEd cap, we scale our estimates to give an effect size for the fully exposed student in the average district, which was 3.2 percentage points above the 8.5% SpEd enrollment threshold in 2004. This implies the likelihood of continuing in SpEd until 9th grade for fully exposed black SpEd students fell by 3.4 percentage points (4.3%) as a result of the cap on SpEd enrollment. For fully exposed Hispanic SpEd students, the likelihood of continuing in SpEd decreased by 3.2 percentage points (4.1%) as a result of the cap on SpEd enrollment. In the long-run, we do not estimate

this table is consistent with the prior that districts furthest from the threshold set by the state are those whose students experience the largest impacts of the policy. Indeed, we find the largest positive impacts of the black disproportionality cap for students in districts in the second and third terciles above the disproportionality cap.

[38]Since the statewide district-level average Hispanic disproportionality rate was already below the 1% threshold in 2004, we scale the impact of the Hispanic disproportionality cap using the average distance above the black disproportionality cap (which as 3.2 percentage points). Given that there existed districts with Hispanic disproportionality rates as high as 3.2 percentage points above the threshold, this allows us to compare estimates across black and Hispanic students on a similar scale. Although we note that the impact of the Hispanic disproportionality cap for the average district would be much smaller.

[39]In addition, in Appendix Table A.5, demonstrating the impact of the disproportionality cap non-parametrically, the impact of the cap on high school completion is no longer statistically significant.

statistically significant impacts of the SpEd enrollment cap on black SpEd student's long-run outcomes in Table 4.[40] For Hispanic SpEd students, the SpEd enrollment cap reduced the likelihood of completing high school by 2.2 percentage points (3.7%), reduced the likelihood of enrolling in college by 1.6 percentage points (5.6%), and reduced the likelihood of obtaining an associate's degree by 0.32 percentage points (16%).

Overall, we conclude that the Hispanic disproportionality cap did not have a statistically significant impact on Hispanic students, and the SpEd cap worsened long-run outcomes for Hispanic students in SpEd prior to policy implementation (consistent with the estimated impacts of the SpEd cap that we find in Ballis and Heath (2021)). In contrast, we find strong evidence that the black disproportionality cap improved outcomes for black SpEd students, but that the SpEd enrollment had a negligible (or suggestively negative) impact on outcomes. We investigate further in Sections 5.3 and 5.4 the mechanisms at play that led to the different long-run impacts of these two caps on black students.

5.2 Effects on GE Students

Next, we turn to estimating the impact of the policy on GE students, defined as students in GE as of 5th grade prior to policy implementation. While these policies directly targeted SpEd students, disproportionality remediation and overall SpEd program reductions could have also impacted GE students. While the majority of the impact on GE students will likely be driven by spillover effects from having more unaccommodated SpEd peers in the GE classroom, a component of this effect could be a direct effect driven by reductions in the likelihood that GE students received SpEd services in later grades.[41] Figures 5 and 6 present event study estimates for the effect of the policy on black and Hispanic GE students' 9th grade SpEd participation and long-run outcomes, respectively. For each of the outcomes

[40] Although, we note that while not statistically significant, we see a negative impact of the SpEd enrollment cap on the likelihood of completing high school and enrolling in college among black SpEd students in the most recent fully exposed cohorts in event-study Figure 4.

[41] The SpEd students losing access due to these policies were, for the most part, already in the GE classroom for the majority of the school day. Thus, spillovers are likely to be driven by the the fact that their SpEd peers were no longer supported by additional services, such as teacher's aides within the GE classroom.

we do not find evidence of differences in pre-treatment trends across districts more or less treated by the disproportionality and SpEd enrollment caps.

In the top panel of Table 6 we present estimates of the effects of the policy on GE students' SpEd participation as of 9th grade. For the fully exposed black GE student in the average district, the likelihood of participating in SpEd in 9th grade fell by 0.30 percentage points (6.6%) as a consequence of the black disproportionality cap, and by 0.98 percentage points (22%) as a result of the SpEd enrollment cap. For the fully exposed Hispanic GE student, the Hispanic disproportionality cap led to a 0.14 percentage point (4.6%) decline in SpEd participation, and the SpEd enrollment cap led to a 0.56 percentage point (17%) decline in SpEd participation for Hispanic students. This implies that part of the impact on GE students will indeed be driven by the direct impact of being less likely to receive SpEd services later on.

Turning to the long-run outcomes in the remaining panels of Table 6, consistent with the effects we found for black SpEd students, we find improvements in long-run outcomes for black GE students resulting from the cap on black disproportionality. In particular, we find for the fully exposed black GE student that the black disproportionality cap increased the likelihood of completing high school by 0.74 percentage points (1.1%) and enrolling in college by 1.7 percentage points (2.5%). This result also holds up to concerns of multiple inference. In column (3) of Table 5 we find a statistically significant positive impact of the black disproportionality cap on a summary index of long-run outcomes for black GE students. For Hispanic GE students, we do not find a statistically significant impact of the Hispanic disproportionality cap on long-run outcomes. This result is reflected in column (4) of Table 5, where the impact of this cap on the summary index of long-run outcomes is also not statistically significant. For black students, the SpEd enrollment cap did not have a statistically significant impact on long-run outcomes, despite having a large negative impact on SpEd participation. For Hispanic students, the SpEd enrollment cap had a statistically significant negative impact on high school completion, college enrollment, and associate's degree attainment. This result is consistent with the negative impacts we find on Hispanic

SpEd students in Table 4. Given that we find a significant decrease in SpEd participation among Hispanic GE students, this result likely reflects a combination of spillover effects from SpEd students as well as a direct effect of a reduction in the likelihood of GE students receiving SpEd services later on in school.

Thus far, we have discussed the impacts for black and Hispanic GE students, but it is possible that an increase in the number of unsupported black and Hispanic students with disabilities in the GE classroom would have an impact on students of other races in the classroom as well. We therefore turn to estimating the effects for white students, proportionally the next largest racial group, in column (3) of Table 6. We now incorporate all three treatment variables additively into our model.[42] Despite the fact that we find improvements in high school completion and college enrollment for black SpEd and GE students as a result of the black disproportionality cap, we do not find statistically significant impacts of the black disproportionality cap on these outcomes for white GE students. In addition, we find negative impacts on the likelihood of receiving an associate's degree among white GE students as a result of the black disproportionality cap. Thus, if anything, white students may have been harmed by black disproportionality remediation. In contrast, the Hispanic disproportionality cap had negative impacts on long-run outcomes for white GE students' likelihood of completing high school, enrolling in college, and earning an associate's degree. Overall the effect sizes are small in magnitude, although precisely estimated.[43]

5.3 Heterogeneous Impacts

Given that we find positive impacts of the black disproportionality cap on black students, in contrast to the lack of a significant impact from the SpEd enrollment cap (and negative

[42] Appendix Figure A.9 illustrates the correlation in the treatment variation between the 2004 district-level black and Hispanic disproportionality rates. This figure illustrate a lack of correlation between the treatment variables, thus motivating why they are incorporated additively into our model.

[43] We note here that we draw similar conclusions in Ballis and Heath (2021), which demonstrate the impact of the 8.5% enrollment cap on all GE students. We reproduce those results here in column (4), but now additionally control for the impacts of the disproportionality caps. This allows us to both highlight the importance of these caps on all GE students and to provide a comparison of the effect of the disproportionality caps on all GE students to the effect on black and Hispanic GE students discussed above.

impact of the SpEd enrollment cap on Hispanic and white students), we focus on black students when investigating heterogeneity and mechanisms. Appendix Table A.6 presents impacts of the policy on black SpEd students by disability type. Column (1) replicates the effect of the policy on black SpEd students from column (3) of Table 4 for comparison purposes. Columns (2) through (5) represent disabilities that we categorize as "malleable", which we determined to have a relatively greater amount of subjectivity in their evaluation criteria compared to more severe or physical disability types. These malleable disabilities include specific learning disability (SLD), speech impairments, emotional disturbance (ED), and other health impairment (OHI) (which is a category that includes ADHD). In column (6) we present results for intellectual disability (ID). This is a more severe cognitive disability, which students may be evaluated for using an IQ test. Finally, column (7) presents results for students with physical disabilities such as deaf or blindness, hearing impairments, visual impairments, or orthopedic impairments.

In Appendix Table A.6 we see that the declines in SpEd participation as a result of the black disproportionality cap were driven by students with SLD (column 2). For these students, we also find positive long-run impacts on high school completion and college enrollment, resulting from the disproportionality cap. The cap on SpEd enrollment led to declines in SpEd participation for those with SLD and Speech Impairments (columns 2-3), but for the most part did not lead to significant changes in long-run outcomes. The one exception is a significant decline in the likelihood of obtaining an associate's degree for individuals with speech impairments. We do not find significant impacts on the likelihood of SpEd placement for black students with ID or physical impairments due to either cap. As we would expect, it is much more difficult to deny SpEd services to students with relatively more severe and objective disability types.

In Appendix Table A.7 we present results for black SpEd students estimated separately by gender and income. While the disproportionality cap led to declines in SpEd participation for both genders, we find that the positive effects of the black disproportionality cap in the long-run are driven by male students. The SpEd cap led to declines in SpEd participation for

both male and female students, but did not lead to statistically significant changes in long-run outcomes for either gender. By income, we find that the declines in SpEd participation are driven by low-income students for both the disproportionality and SpEd enrollment caps.[44] 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 an increase in college enrollment for non-FRL students as a result of the disproportionality cap.[45] The SpEd enrollment cap did not have a statistically significant impact on long-run outcomes for FRL students, although we do find a significant negative impact on high school completion for non-FRL students.

Turning to black GE students in Appendix Table A.8, we again find that the positive long-run impacts of the disproportionality cap are driven by males. Although the SpEd enrollment cap had a statistically significant negative impact on SpEd participation, it again did not have a significant impact on black GE students' long-run outcomes. By income, we find that non-FRL black GE students are slightly more positively affected by the cap on disproportionality, although the direction and magnitude of the coefficients are similar across FRL and non-FRL black GE students. The SpEd enrollment cap did not have an impact on FRL or non-FRL black GE students in the long-run.

Finally, we estimate effects by whether the district is rural or urban and by the district's tax-base wealth in Appendix Table A.9. We find that the positive impact on long-run outcomes for black SpEd students as a result of the black disproportionality cap appears to be driven by urban districts. This is in line with what we would expect, given that the fraction of black students is much higher in urban districts rather than rural (15.7% vs.

[44]Reductions in SpEd placement for lower-income students is consistent with less advantaged parents being less able to challenge SpEd removal decisions made by school personnel under pressure to reduce SpEd access, whereas higher income parents may have additional resources that allow them to more easily advocate to keep their SpEd placement (Koseki, 2017).

[45]Ballis and Heath (2021) find that FRL students (of all races) are significantly more likely to be harmed by SpEd removal compared to non-FRL students, as a result of the SpEd enrollment cap. The fact that there is relatively little difference in the impact of the disproportionality cap across income suggests that the marginal student is more likely to be affected by changes that are more difficult to compensate for financially, such as a reduction in stigma or racial bias in school. We discuss this mechanism in greater detail in Section 5.4.

8.9%). In addition, the black disproportionality rate in 2004 was higher in urban rather than rural districts (4.5 vs. 3.3 percentage points). In terms of wealth, we find positive effects of the black disproportionality cap for black students in districts both above and below the median level of district tax-based wealth per pupil.[46] We find a negative impact of the SpEd enrollment cap on college enrollment for black SpEd students in districts with below median wealth. This is consistent with the idea that higher wealth districts would be better able to compensate for SpEd removal by providing other resources (such as tutoring) to students, whereas students in lower wealth districts who lose SpEd as a result of the enrollment cap do worse in the long-run without other resources to help compensate.

5.4 Mechanisms

To this point, we have found that the cap on black disproportionality had positive impacts on long-run outcomes for black students, whereas the cap on SpEd enrollment overall did not have a statistically significant impact on long-run outcomes. This result is perhaps surprising and points to the fact that there are likely meaningful differences between marginal black students impacted by the disproportionality cap and marginal black students impacted by the SpEd cap. In this section we explore several potential explanations for this difference.

In prior literature, Elder et al. (2021) link birth certificate records to public school data in Florida to develop a model to predict the likelihood of SpEd placement based on a rich set of individual characteristics.[47] The authors find that minority students are more likely to be under-represented conditional on observables in heavily minority schools relative to white students, and are more likely to be conditionally over-represented in heavily white schools relative to white students. Building on this insight, we estimate the impact of the policy

[46]The fact that the positive effect of the disproportionality cap is similar by district wealth suggests that the marginal student impacted by the disproportionality cap was again more likely to be influenced by factors that are more difficult to change with additional resources (such as a possible reduction in racial bias).

[47]In particular, by linking in birth certificate records the authors are able to include a variety of health information such as infant birth weight, gestational age, APGAR scores, the mother's prior births, and diagnosis codes for congenital anomalies, abnormal conditions, complications during delivery, and the mother's pregnancy-related health diagnoses. The birth certificate data also include demographic and economic information such as mother's marital status, mother's educational attainment, mother's race, mother's immigration status, and the mother's zip code of residence when the child was born (Elder et al., 2021).

separately for districts we categorize as having a conditional over- or under-representation of black students in SpEd. In theory, districts should be placing students in SpEd until the marginal cost of providing SpEd services exceeds the marginal benefit to the student. If it is the case that districts with an over-representation of black students in SpEd have placed students in SpEd whose marginal cost exceeds the marginal benefit, this would imply that removal from SpEd would improve long-run outcomes. Similarly, if it is the case that districts with an under-representation of black students in SpEd have not yet reached the equilibrium level of SpEd placement, where the marginal benefit equals the marginal cost, this would be one reason why removal from SpEd is detrimental to long-run outcomes. If this theory is true, and if districts with high black disproportionality rates are also districts with a conditional over-representation of black students, this would help explain why black students removed from SpEd in these districts do better in the long-run.

We test this theory in our context by sorting districts into those with an under- or over-representation of black students in SpEd. To do so, we follow Elder et al. (2021)'s approach and implement a Blinder-Oaxaca decomposition. First, we use a logit model to predict the likelihood of SpEd placement for white students, based on pre-treatment characteristics.[48] Next, we apply the coefficients from this model to black students, to predict the likelihood of SpEd placement for black students as if they were white. Then, we subtract the prediction from an indicator for whether a student is actually in SpEd. This gives us a measure of whether the student is predicted to be over- or under-represented in SpEd relative to an observationally-equivalent white student. Finally, we aggregate these differences to the district-level, to obtain a prediction for whether each district has an over- or under-representation of black students in SpEd.

Our estimates for the impact of the policy separately by districts over- and under-

[48]The results of this logit model are presented in Appendix Table A.10. SpEd status is predicted as of 5th grade. The pre-treatment characteristics are measured as of 3rd grade and include age (as of September 1st), gender, FRL, Title I, Bilingual, Limited English Proficient (LEP), Gifted, At Risk, Migrant, Even Start Program participant, math score, and reading score at the individual. We also include gender, race, FRL, Bilingual, LEP, Title I, At Risk, and Gifted status at the grade and district levels. As discussed above, 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 0.0708.

representing black SpEd students are presented in Appendix Table A.11. The impact of the disproportionality cap on the likelihood of SpEd placement is negative, although not statistically significant for either type of district. In line with our prior, we find a statistically significant increase in the likelihood of college completion for black students in districts predicted to have an over-representation of black students in SpEd. Taken together, this provides suggestive evidence that black students benefit from SpEd removal in districts that over-identify black students for SpEd. However, we also explore other possible mechanisms to contribute to our understanding of the reasons for the differences in effects across the disproportionality and SpEd enrollment caps.

The second approach we take to better understand the differences in the impacts of the two caps is to estimate differences in the types of black students more likely to be impacted by each of the caps across districts. When the policy went into effect, districts exogenously increased the proportions of students removed from SpEd between 5th and 9th grade. We estimate changes in the district-level composition of students removed from SpEd to investigate whether students were removed based on observable characteristics.[49] If it is the case that districts with high rates of black disproportionality removed certain types of black students from SpEd and districts with high rates of SpEd enrollment removed some other types of black students from SpEd, this would help explain why we find different impacts of these two caps.

We present these results characterizing district-level changes in the composition of black students removed from SpEd in Table 7. We find that black students removed from SpEd by 9th grade were more likely to be higher performing on the reading exam and more likely to be in resource rooms for less than 50% of the day at baseline in districts relatively more impacted by the disproportionality cap. In contrast, black students removed from SpEd

[49]In particular, 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 follows: we take 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. This outcome is regressed on the district-level 2004 black disproportionality rate and SpEd rates interacted with exposure (the number of years a district-cohort was in school under the policy). We also include cohort fixed effects and robust standard errors clustered at the district level.

in districts more impacted by the SpEd cap were more likely to be lower performing on the reading exam and less likely to have a malleable disability type at baseline, relative to districts less impacted by the SpEd cap. We conclude from this that black students removed from SpEd as a result of the disproportionality cap are more likely to be higher achieving and have relatively more mild disability types. And, these students do better in the long-run as a result of this cap. However, black students removed from SpEd as a result of the SpEd cap are more likely to be lower performing and are less likely to have more mild disability types. And, these students potentially do worse in the long-run as a result of this cap. Taken together, these results are consistent with a misclassification of relatively higher performing black students in SpEd in districts with high rates of black disproportionality.

To provide further evidence in line with the potential misclassification of black students in districts with high rates of black disproportionality, we estimate impacts by district-level teacher experience and racial composition for SpEd students. Intuitively, it may be the case that districts with lower teacher experience may be more likely to misclassify low-performing black students for SpEd services, and thus, black students in these districts might benefit the most from SpEd removal as a result of the black disproportionality cap. In Appendix Table A.12 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 teaching experience of 11.7 years. In fact, we find that the black disproportionality cap has a somewhat larger positive impact on high school completion and college enrollment in districts with below-average teacher experience.

In Appendix Table A.12 columns (3-4), we present estimates for black SpEd students in districts with above or below the median proportion of black teachers. It may be the case that students with a same-race teacher may be less likely to be misidentified for SpEd services. Thus, we might expect black students in districts with below median proportions of black teachers to be more likely to benefit from the black disproportionality cap in the long run. Although the positive effects of the black disproportionality cap on high school completion and college enrollment are slightly larger in magnitude for districts with below

median proportions of black teachers, these results are not statistically significant and thus provide only suggestive evidence of a difference across teacher racial composition.[50]

Finally, we investigate intermediate outcomes to determine whether part of the differences we find across the caps can be explained by changes in outcomes before high school completion and college enrollment decisions are made. In particular, we estimate effects on individual-level outcomes for SpEd students including absences, suspensions, and expulsions in Appendix Table A.13.[51] We find an increase in the percent of days absent for black SpEd students as a result of the disproportionality cap, although we do not find statistically significant impacts of either cap on the likelihood of being truant (i.e. having unexcused absences from school). In terms of the disciplinary outcomes, we find that the cap on black disproportionality increased black students' likelihood of being suspended and having multiple suspensions. It is perhaps somewhat surprising that we find increases in absences and disciplinary outcomes for black students as a result of the disproportionality cap, given the positive long-run outcomes we find. This may be suggestive that improvements in the long-run outcomes of black SpEd students affected by the disproportionality cap could potentially be even higher if absences and disciplinary actions had not increased.

Next, we investigate mechanisms behind the positive long-run impacts of the black disproportionality cap on black GE students. Although we are not able to directly test the mechanisms behind this effect due to data constraints, we propose three potential mechanisms. First, since we find a statistically significant decrease in SpEd participation among black GE students in later grades, this could be a direct effect on students, possibly in-line with a similar story of reduced misclassification for SpEd, as we argue for SpEd students. Second, if this effect is not driven solely by the direct impact on reduced SpEd participation in later grades, it may be the case that all black students benefited from a perceived reduction in racial bias when the disproportionality policy went into place. If black students were

[50]One caveat of comparing the estimates of students in districts with above or below median proportions of black teachers in particular is that there are significantly fewer black students in districts with below median proportions of black teachers.

[51]For this table the specification uses SpEd students (identified as being in SpEd as of 5th grade prior to policy implementation) in all grades from K-12. We replace exposure with "post," an indicator variable equal to 1 in the post-period, and include district, grade, and year fixed effects.

placed in SpEd for racially motivated reasons, SpEd and GE students may both benefit from a reduction in racially biased policies. Indeed we find that black GE students experienced the largest positive impacts due to black disproportionality remediation relative to other races, suggesting that the effects we document are likely driven by changes that are specific to black students. Third, if there is a reduction in the rate at which students are receiving SpEd services, GE teachers may change their practices and techniques to compensate for the difficulties potentially associated with fewer SpEd supports in the classroom (such as teacher's aides). In this case, GE teachers may be improving instruction in a way that benefits both SpEd and GE students.[52]

5.5 Robustness

First, we estimate whether districts facing greater pressure under the policy were on differential trends in terms of background characteristics. We do so by including interaction terms of our district-level control variables and indicator variables for each cohort year. These estimates are presented for SpEd and GE students in Tables 8 and 9, respectively. In each table, columns (2) and (5) include the district-level composition of black, Hispanic, other races, FRL, ESL, and gifted students. In columns (3) and (6) we additionally include tends in whether a school was urban or rural and the district size at baseline. For black students, these results are similar to our original estimates. For black SpEd students in Table 8, while the estimate on high school completion is not statistically significant, it is fairly similar in magnitude and we do still find a statistically significant increase in the likelihood of college enrollment. For black GE students in Table 9, we find that the positive impacts on high school completion and college enrollment are robust to the inclusion of both the demographic characteristics in column (2) as well as school location and district size in column (3).

For Hispanic SpEd students in Table 8, the impact of the Hispanic disproportionality

[52]In our view, the pattern of heterogeneity that we document, with impacts of the black disproportionality cap on GE students being driven by black students, suggests that a combination of direct and spillover effects driven by a reduction in misclassification (potentially resulting from racial bias) is the most likely explanation for our findings for GE students.

cap on high school completion is no longer statistically significant once we add demographic controls in column (5). Although we estimate a statistically significant decrease in college enrollment in column (5), the effect is not statistically significant in column (6) once we control for school location and district size. For Hispanic GE students in Table 9, in our original estimates we did not find significant impacts of the Hispanic disproportionality cap in the long-run. However, once we add demographic controls in column (5) we find very small increases in college enrollment and bachelor's degree attainment and a small decrease in associate's degree attainment. Therefore, we maintain that at best we have only suggestive evidence of a negative impact of the Hispanic disproportionality cap on Hispanic students.

One potential concern for our results is the possibility that students move out of Texas public schools or to a different district to obtain SpEd services upon being denied in their current district. In Table 10, we directly test whether students are systematically moving out of public schools or to a different district as a result of the policy. To do so, we estimate the effect of the policy on the likelihood of switching districts or leaving the data between 5th and 9th grade for our sample of students in SpEd as of 5th grade.[53] For black students, we find increases in the likelihood of being enrolled at 9th grade for FRL students, but do not a statistically significant impact on the likelihood that non-FRL students were enrolled by 9th grade as a result of the disproportionality cap.[54] The positive 9th grade enrollment effect for FRL students is in the opposite direction of what we may have expected if the policy induced students to leave public schools to acquire SpEd services elsewhere. We conclude that this positive effect 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 disproportionality cap.[55] In addition, increases in lower-income and potentially lower-

[53]We note that our data can only follow students who remain in the public school system in Texas. Therefore, we are only able to observe exits from public school and cannot look at where students go after exiting.

[54]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 or in home school), as non-FRL students are more likely to have families that have the resources to move them in response to the policy.

[55]In Appendix Table A.14, 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

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. In the bottom panel of Table 10, we estimate changes in the likelihood of switching districts between 5th and 9th grade. We find increases in the likelihood that black SpEd students switched districts between 5th and 9th grade as a result of the disproportionality cap. However, we note that excessive district switching does not pose a threat to identification since we assign treatment based on each student's pre-policy district. Thus, we view this district switching as only serving to potentially attenuate our estimated effects.[56]

Turning to Hispanic students, we find statistically significant decreases in the likelihood of being enrolled in school at 9th grade and switching districts between 5th and 9th grade as a result of the Hispanic disproportionality cap. As discussed previously, we do not find evidence that the effect of the Hispanic disproportionality cap on Hispanic SpEd students' long-run outcomes is very robust. However, this reduction in enrollment may be contributing to the suggestive evidence we find of a decrease in high school completion and college enrollment for Hispanic SpEd students as a result of the Hispanic disproportionality cap. In terms of the SpEd enrollment cap, we do not find statistically significant changes in the likelihood of being enrolled at 9th grade or in the likelihood of switching districts between 5th and 9th grade. Thus, we conclude that the negative impacts of the SpEd cap on Hispanic students' outcomes were not biased by students moving out of public schools or switching districts in order to game the system to prevent loosing SpEd services in their current district.

Finally, we investigate the extent to which districts altered their spending on SpEd and GE students as a result of the policy. In Appendix Table A.15 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 we do not find changes in the level of SpEd spending per SpEd students. We also do not

impact on being enrolled at 7th grade for FRL students as a result of the black disproportionality cap. We view this as additional evidence in favor of an enrollment effect, mitigating dropout, rather than changes in the composition of the sample for other reasons.

[56]Overall, for black students we do not find evidence of systematic changes in exiting or district switching as a result of the cap on SpEd enrollment.

find evidence of changes in GE spending per GE students.

6 Conclusion

We have presented estimates of the impact of limiting minority student access to SpEd on the likelihood of continuing in SpEd, high school completion, and post-secondary attainment for black and Hispanic students in SpEd and GE. Three district-level caps in particular allow us to quantify causal estimates of the effect of reducing disproportionality and SpEd enrollment on long-run outcomes. Under the Performance Based Monitoring Analysis System (PBMAS) introduced in 2004, Texas capped the percent of black and Hispanic students in SpEd relative to the percent of black and Hispanic students in the district, known as disproportionality. Texas also capped the overall SpEd rate at 8.5% at the district level. We exploit 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 2004-2005 school year, it impacted districts differentially based on their pre-treatment disproportionality rates and their pre-treatment percent of students in SpEd. We show that districts with the highest 2004 proportions of black and Hispanic disproportionality were impacted the most by the caps on district-level disproportionality. These districts made the largest reductions in their disproportionality rates to meet the standard set by the state. We show the analogous relationship for districts with the highest rates of SpEd enrollment.

We estimate the impact of these caps separately for students in SpEd as of 5th grade prior to policy implementation and students in GE as of 5th grade prior to policy implementation. Overall, we find that the black disproportionality and the SpEd enrollment 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. And we find suggestive evidence of a negative impact of the SpEd enrollment cap on black students' long-run outcomes. The

Hispanic disproportionality cap did not have a significant impact on Hispanic SpEd students. Although we find negative impacts of the SpEd cap on Hispanic students' high school completion and college enrollment.

We explore several potential mechanisms behind the positive effect of the black disproportionality cap on black students. We point to a variety of evidence which we view as consistent with a story of misclassification of black students for SpEd in districts with high rates of disproportionality, leading these students to benefit from SpEd removal in the long-run. In particular, we find that black students removed from SpEd by 9th grade are relatively higher performing with more mild disability types at baseline. In contrast, black students removed from SpEd by 9th grade as a result of the SpEd enrollment cap are relatively lower performing with more severe disability types at baseline. Given that these relatively higher performers with less severe disability types do better in the long-run after SpEd removal, we conclude that districts with high rates of black disproportionality are likely misclassifying students for SpEd services.

The impacts we find for GE students are consistent with the impacts we find for SpEd students. As a result of the cap on black disproportionality, we find improvements in long-run outcomes for black GE students. To the extent that the black disproportionality cap alleviates racial bias in schools, this could positively impact both SpEd and GE student outcomes. As a result of the SpEd enrollment cap, we find that GE students do worse in the long-run, regardless of race. This is likely a result of increases in the number of unsupported and unaccommodated students who were previously in SpEd (who were worse off after SpEd removal). This could be driven by several potential mechanisms. Although SpEd students may have already been in the GE classroom for the majority of the day prior to SpEd removal, they may have had a teacher's aide or other accommodations that would have been aimed at boosting academic achievement. In the absence of these services, previously enrolled SpEd students may require additional attention from the teacher, which would leave less attention for the GE students in the classroom. Additionally, a teacher's aide in the classroom could have benefited both SpEd and GE students if they were able to

answer questions and provide help to all students in the room.

Overall, the estimates we find have meaningful implications for SpEd policy in public schools. We find different treatment effects for black SpEd students as a result of different policies limiting SpEd enrollment. Students who require SpEd services greatly benefit from them in the long-run. However, those who are misclassified for SpEd can be significantly harmed in the long-run. 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. Finally, whether SpEd students are appropriately served has important implications for all students in the classroom. This highlights the fact that both SpEd and GE students' outcomes should be considered when making SpEd policy-related decisions.

References

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

Ballis, B., & Heath, K. (2021). The Long-Run Impacts of Special Education. *Conditionally accepted at American Economic Journal: Economic Policy*.

Bear, G., Clever, A., & Proctor, W. (1991). Self-Perceptions of Nonhandicapped Children and Children with Learnings Disabilities in Integrated Classes. *Journal of Special Education*, *24*(4), 409–426.

Blanchett, W. J. (2006). Disproportionate Representation of African American Students in Special Education: Acknowledging the Role of White Privilege and Racism. *Educational Researcher*, *35*(6), 24–28.

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

Cullen, J. (2003). The Impact of Fiscal Incentives on Student Disability Rates. *Journal of Public Economics*, *87*, 1557–1589.

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

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

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

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (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.

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.

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

Koseki, M. H. (2017). Meeting the needs of all students: Amending the idea to support special education students from low-income households. *Fordham Urb. LJ*, *44*, 793.

Lackaye, T., & Margalit, M. (2006). Comparisons of Achievement, Effort, and Self-Perceptions Among Students with Learning Disabilities and their Peers from Different Achievement Groups. *Journal of Learning Disabilities*, *39*(5), 432–446.

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

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

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

Mountjoy, J. (2021). Community Colleges and Upward Mobility. *Revise and resubmit at American Economic Review*.

Office of Special Education and Rehabilitative Services. (2000). A Guide to the Individualized Education Program.

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

Office of Special Education and Rehabilitative Services. (2010). Thirty-five Years of Progress in Education Children with Disabilities Through IDEA.

Office of Special Education and Rehabilitative Services. (2017). Texas Part B 2017 Monitoring Visit Letter. Retrieved from https://www2.ed.gov/fund/data/report/idea/partbdmsrpts/dms-tx-b-2017-enclosure.pdf

Office of Special Education Programs. (2017). Significant Disproportionality (Equity in IDEA). Retrieved from https://sites.ed.gov/idea/files/significant-disproportionality-qa-03-08-17-1.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.

Prenovitz, S. (2017). Accountability Incentives and Special Education.

Reschly, D. (1996). Identification and Assessment of Students with Disabilities. *The Future of Children: Special Education for Students with Disabilities*, *6*(1), 40–53.

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

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.

Strassfeld, N. (2017). The Future of IDEA: Monitoring Disproportionate Representation of Minority Students in Special Education and Intentional Discrimination Claims. *Case Western Reserve Law Review*, *67*(4), 1121–1151.

Texas Education Agency. (2004). Performance-Based Monitoring Analysis System 2004-2005 Manual. Retrieved from https://tea.texas.gov/sites/default/files/0405PBMASManual.pdf

Texas Education Agency. (2016a). 2016-2017 PBMAS Staging Framework.

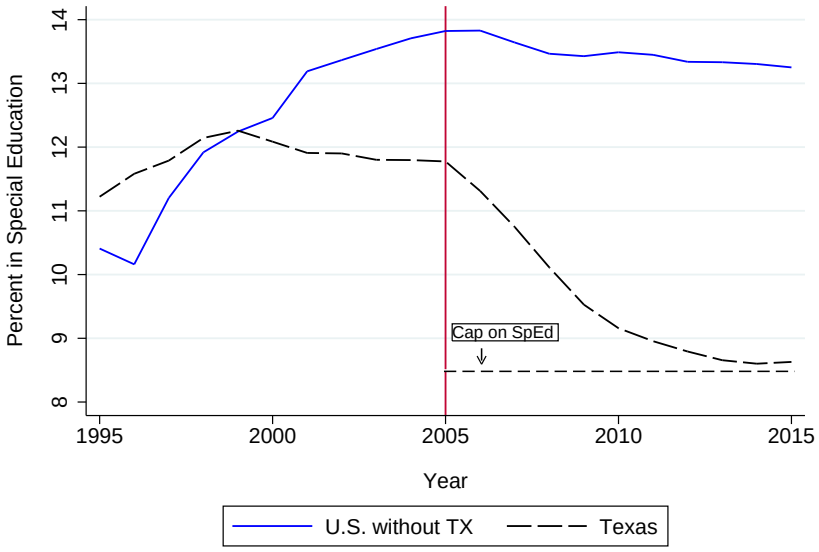
Texas Education Agency. (2016b). Performance-Based Monitoring Analysis System (PBMAS) Overivew. Retrieved from http://tea.texas.gov/Student_Testing_and_Accountability/PBMAS/

Texas Education Agency. (2018). Improving Special Education in Texas. Retrieved from https://tea.texas.gov/TexasSPED/

U.S. Department of Education. (2018). U.S. Department of Education Issues Findings in Texas Individuals with Disabilities Education Act Monitoring. Retrieved from https://www.ed.gov/news/press-releases/us-department-education-issues-findings-texas-individuals-disabilities-education-act-monitoring

Figures and Tables

Figure 1 Percent of Students in Special Education



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

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 2

(a) Special Education Rate by Race

(b) Disproportionality Rate for Black and Hispanic Students

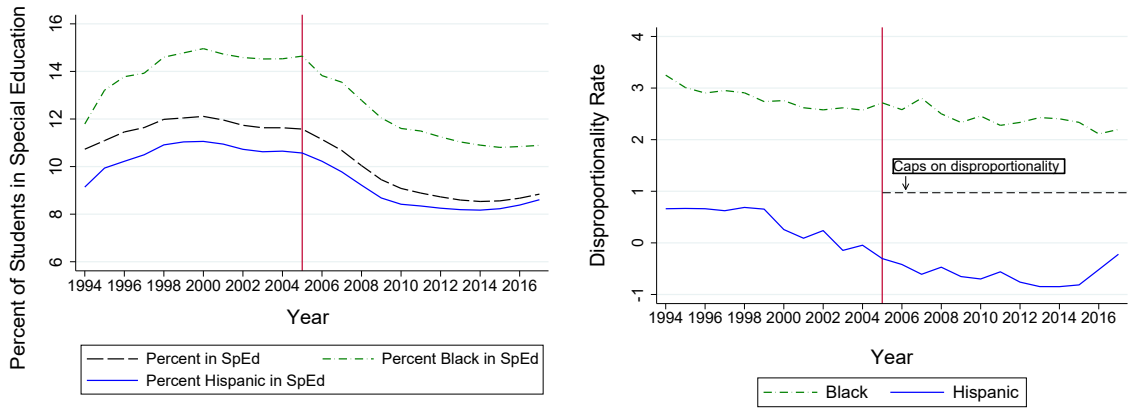
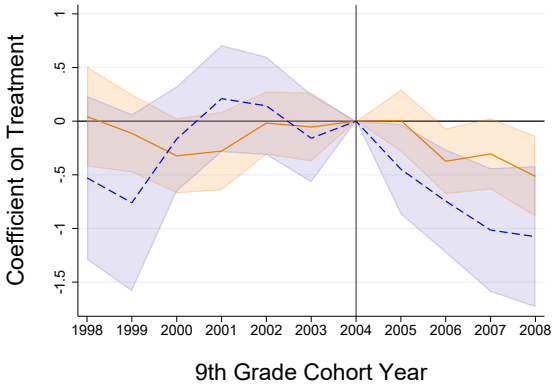


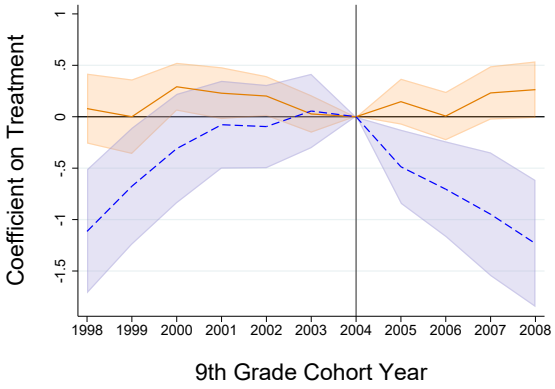
Figure (a) plots the percent of students in special education in Texas by race. Figure (b) plots the average disproportionality rate for black and Hispanic students. The disproportionality rate is measured as the percent of black or Hispanic students in special education minus the percent of black or Hispanic students in a given district.

Figure 3 Event Study for Special Education Students

(a) Grade 9 SpEd Status for Black SpEd Students



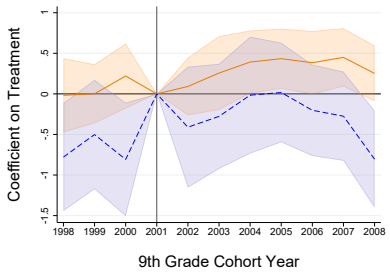
(b) Grade 9 SpEd Status for Hispanic SpEd Students



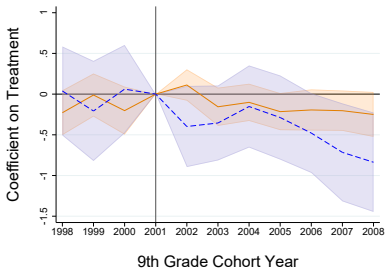
In each graph, the series in blue denotes the coefficient on the average district-level black SpEd rate in 2004 interacted with indicators for each 9th grade cohort year. The series in orange denotes the average district-level black or Hispanic disproportionality rate in 2004 interacted with indicators for each 9th grade cohort year. 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.

Figure 4 Event Study for Special Education Students

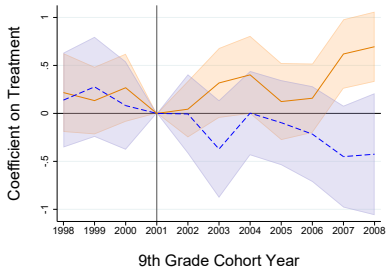
(a) High School Completion for Black SpEd Students



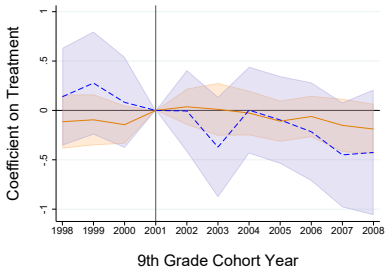
(b) High School Completion for Hispanic SpEd Students



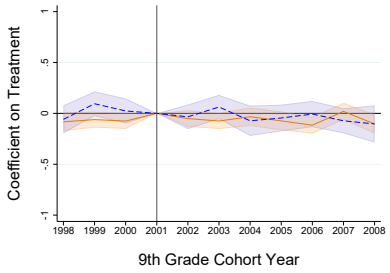
(c) College Enrollment for Black SpEd Students



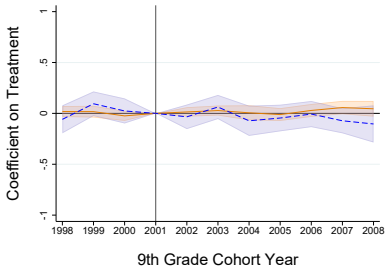
(d) College Enrollment for Hispanic SpEd Students



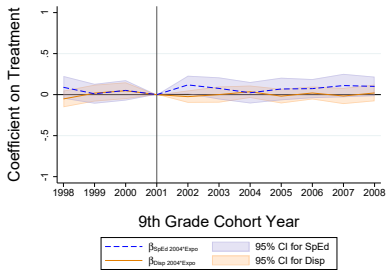
(e) Associate's Degree for Black SpEd Students



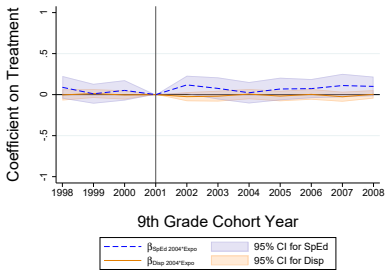
(f) Associate's Degree for Hispanic SpEd Students



(g) Bachelor's Degree for Black SpEd Students



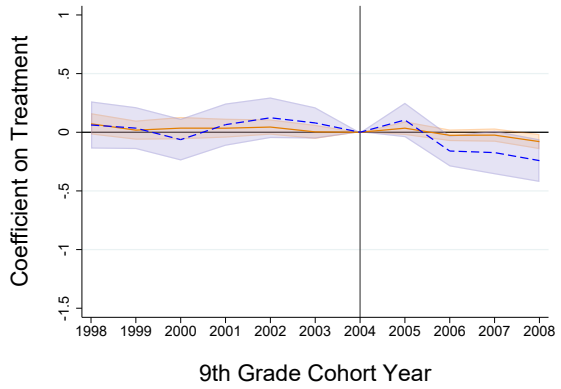
(h) Bachelor's Degree for Hispanic SpEd Students



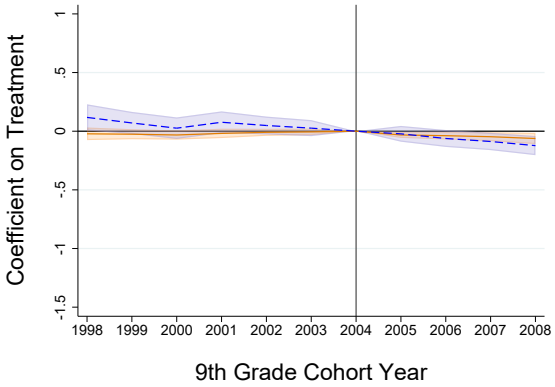
Each figure plots coefficients from regressions interacting the SpEd enrollment and disproportionality caps with 9th grade cohort years. Outcomes are labeled in the title of each figure. See Figure A.8 for a description of the full set of regression controls. The vertical line is placed at 2001 since individuals in 9th grade in 2001 would have been in 12th grade in 2004.

Figure 5 Event Study for General Education Students

(a) Grade 9 SpEd Status for Black GE Students



(b) Grade 9 SpEd Status for Hispanic GE Students

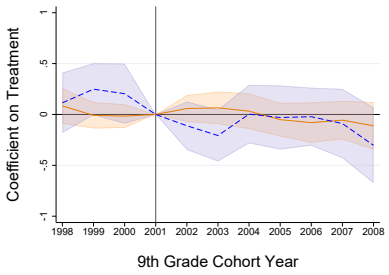


In each graph, the series in blue is the coefficient on the district-level black SpEd rate in 2004 interacted with indicators for each 9th grade cohort year. The series in orange is the coefficient on the district-level black or Hispanic disproportionality rate in 2004 interacted with indicators for each 9th grade cohort year. Regressions include controls for individual gender, FRL, ESL, gifted, and Title I status, as well as district-cohort level gender, race, ESL, FRL, gifted, and Title I composition. We additionally include district and cohort fixed effects, and robust standard errors are clustered at the district level.

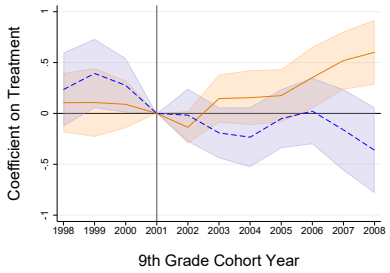
Figure 6 Event Study for General Education Students

(a) High School Completion for Black GE Students

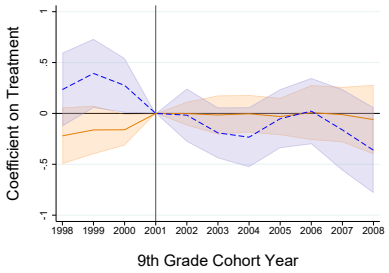
(b) High School Completion for Hispanic GE Students



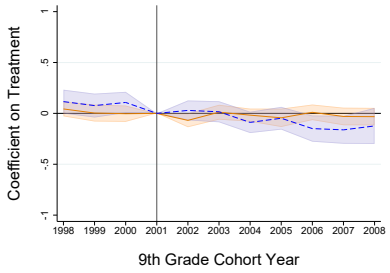
(c) College Enrollment for Black GE Students



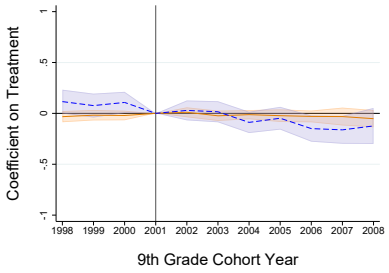
(d) College Enrollment for Hispanic GE Students



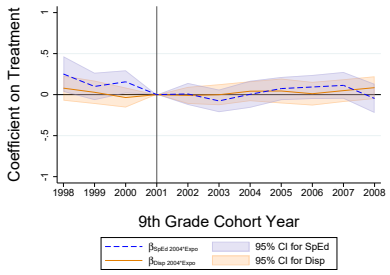
(e) Associate's Degree for Black GE Students



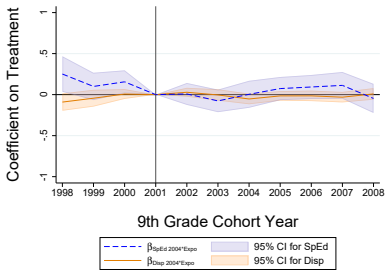
(f) Associate's Degree for Hispanic GE Students



(g) Bachelor's Degree for Black GE Students



(h) Bachelor's Degree for Hispanic GE Students



Each figure plots coefficients from regressions interacting the SpEd enrollment and disproportionality caps with 9th grade cohort years. Outcomes are labeled in the title of each figure. See Figure 5 for a description of the full set of controls. The vertical line is placed at 2001 since individuals in 9th grade in 2001 would have been in 12th grade in 2004.

Table 1 Descriptive Statistics for Grades K through 12

	All Students			SpEd Students		
	All Races	Black Students	Hispanic Students	All Races	Black Students	Hispanic Students
	(1)	(2)	(3)	(4)	(5)	(6)
Covariates						
Male	0.513	0.510	0.512	0.670	0.670	0.668
FRL	0.545	0.680	0.769	0.607	0.746	0.787
ESL	0.031	0.003	0.059	0.036	0.002	0.080
Title I	0.564	0.608	0.752	0.548	0.569	0.716
Gifted	0.078	0.047	0.057	0.010	0.004	0.006
White	0.374	·	·	0.397	·	·
Black	0.136	·	·	0.171	·	·
Hispanic	0.454	·	·	0.415	·	·
Other	0.036	·	·	0.018	·	·
SpEd Rate	0.103	0.129	0.094	·	·	·
N	103,454,758	14,082,052	46,976,377	10,648,571	1,816,730	4,415,654
Long-run Outcomes						
High School Diploma	0.652	0.595	0.590	0.611	0.581	0.575
Enroll College	0.513	0.465	0.418	0.341	0.316	0.284
Obtain Associate's	0.043	0.025	0.041	0.024	0.013	0.020
Obtain Bachelor's	0.126	0.073	0.061	0.034	0.017	0.014
N	4,777,246	692,037	1,840,448	417,366	76,681	160,865

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 stands for English as a Second Language program (now commonly referred to as English Language Learners). Gifted is a separately defined category from Special Education in Texas, and is a program for high achieving students. High School diploma is an indicator for graduating from high school within 2 years of expected high school graduation, and conditional on being observed in the data in grade 9. Indicators for enrolling in college and obtaining an Associate's or Bachelor's degree are censored to 6 years after expected high school graduation, but not conditional on high school diploma. SpEd status is designated as of 5th grade for the long-run outcomes.

Table 2 Difference in Means Between Districts Above and Below SpEd Enrollment Cap in 2004

	Less than 8.5%	Greater than 8.5%	Difference
Male	0.501	0.516	-0.014***
White	0.256	0.581	-0.325***
Black	0.247	0.093	0.154***
Hispanic	0.465	0.314	0.151***
Other	0.032	0.012	-0.019***
FRL	0.559	0.542	0.018
ESL	0.019	0.022	-0.002
Title I	0.624	0.681	-0.058*
Gifted	0.056	0.079	-0.023***
Standardized Math	-0.313	-0.051	-0.261***
Standardized Reading	-0.200	-0.009	-0.191***
N	106	1,018	

*** p<0.01, ** p<0.05, * p<0.1 This table provides differences in characteristics across all students in districts in 2004 for grades 3 through 8. We compare districts with greater than 8.5% of students in SpEd in 2004 to those with fewer than 8.5% of students in SpEd in 2004. N represents the number of districts.

Table 3 Difference in Means Between Districts Above and Below Disproportionality Caps in 2004

	Black Disprop. Rate			**Hispanic Disprop. Rate**		
	Less than 1pp	Greater than 1pp	Difference	Less than 1pp	Greater than 1pp	Difference
Male	0.512	0.514	-0.001	0.514	0.511	0.003
White	0.511	0.522	-0.011	0.506	0.536	-0.030
Black	0.074	0.170	-0.096***	0.152	0.072	0.080***
Hispanic	0.401	0.290	0.111***	0.328	0.371	-0.043***
Other	0.014	0.018	-0.004***	0.014	0.020	-0.007**
Econ Disadvantage	0.550	0.540	0.010	0.576	0.491	0.085***
ESL	0.023	0.019	0.004*	0.022	0.020	0.0002
Title I	0.677	0.664	0.013	0.685	0.644	0.041*
Gifted	0.092	0.100	-0.008***	0.071	0.079	-0.007**
Special Ed	0.120	0.125	-0.005	0.124	0.120	0.004
Standardized Math	-0.117	-0.089	-0.028	-0.155	-0.007	-0.148***
Standardized Reading	-0.057	0.035	-0.028	-0.097	0.046	-0.144***
N	557	567		703	421	

*** p<0.01, ** p<0.05, * p<0.1 This table provides descriptive statistics on all students in 2004 for grades 3 to 8 in districts with less than a 1% disproportionality rate for black/Hispanic students and districts with greater than a 1% disproportionality rate for black/Hispanic students. N represents the number of districts.

Table 4 Effect of Policy on Students in SpEd as of 5th Grade Prior to Policy Implementation

	Black Students			Hispanic Students		
SpEd Status G9	(1)	(2)	(3)	(4)	(5)	(6)
$Disp_{d,2004} \times Exposure$	-0.0366	-0.0745*	-0.0946**	0.0027	0.0193	0.0264
	(0.038)	(0.039)	(0.039)	(0.029)	(0.028)	(0.030)
$SpEd_{d,2004} \times Exposure$	-0.2938***	-0.2984***	-0.2657***	-0.2855***	-0.2423***	-0.2471***
	(0.068)	(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						
$Disp_{d,2004} \times Exposure$	0.0472**	0.0609***	0.0523**	-0.0208*	-0.0236*	-0.0281**
	(0.024)	(0.023)	(0.022)	(0.012)	(0.012)	(0.014)
$SpEd_{d,2004} \times Exposure$	-0.0176	-0.0050	-0.0047	-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***	0.0702***	0.0659***	-0.0059	-0.0158	-0.0155
	(0.017)	(0.017)	(0.018)	(0.012)	(0.014)	(0.016)
$SpEd_{d,2004} \times Exposure$	0.0285	0.0433	0.0259	-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
Associate's Degree						
$Disp_{d,2004} \times Exposure$	-0.0076**	-0.0052	-0.0022	0.0066*	0.0055	0.0058
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
$SpEd_{d,2004} \times Exposure$	-0.0017	-0.0001	-0.0025	-0.0140*	-0.0149*	-0.0143*
	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)
Mean Dept Var	0.014	0.014	0.014	0.020	0.020	0.020
Bachelor's Degree						
$Disp_{d,2004} \times Exposure$	-0.0032	0.0009	0.0007	0.0029	0.0002	-0.0004
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
$SpEd_{d,2004} \times Exposure$	-0.0028	-0.0006	0.0008	0.0069	0.0065	0.0074
	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)
Mean Dept Var	0.018	0.018	0.018	0.015	0.015	0.015
Observations	72,197	72,197	72,197	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 and obtain a degree. Regressions include 5th grade cohorts from 1994 (when the data begins) to 2004 (the year prior to policy implementation).

Table 5 Effect of Policy on Index Measure of Long-Run Outcomes

	SpEd Sample		**GE Sample**	
	Black	Hispanic	Black	Hispanic
	(1)	(2)	(3)	(4)
$Disp_{d,2004} \times Exposure$	0.0586***	-0.0160	0.0590***	-0.0040
	(0.019)	(0.014)	(0.020)	(0.013)
$SpEd_{d,2004} \times Exposure$	0.0081	-0.0994***	0.0041	-0.0372
	(0.033)	(0.035)	(0.035)	(0.026)
Mean Dept Var	-0.261	-0.276	0.027	0.023
Observations	72,197	153,098	280,464	770,520

*** p<0.01, ** p<0.05, * p<0.1 The outcome variable is a summary measure of all the long-run outcome variables. We standardize each outcome to have mean 0 and standard deviation 1, including indicators for high school graduation, college enrollment, associate's degree attainment, and bachelor's degree attainment. Then, we create one summary index of all the long-run outcomes by averaging across the standardized long-run outcomes for each individual. Specifications are run as before, with the summary index as the outcome variable. See Table 4 for full set of controls for the SpEd sample and Table 6 for the full set of controls for the GE sample.

Table 6 Effect of Policy on Students in GE as of 5th Grade Prior to Policy Implementation

	Black	Hispanic	White	All Races
SpEd Status G9	(1)	(2)	(3)	(4)
$DispBlack_{d,2004} \times Exposure$	-0.0232***		0.0022	-0.0034
	(0.008)		(0.005)	(0.005)
$DispHispanic_{d,2004} \times Exposure$		-0.0116**	0.0022	-0.0048
		(0.005)	(0.005)	(0.005)
$SpEd_{d,2004} \times Exposure$	-0.0766***	-0.0437***	-0.0399***	-0.0385***
	(0.020)	(0.009)	(0.006)	(0.006)
Mean Dept Var	0.045	0.032	0.030	0.032
High School Completion				
$DispBlack_{d,2004} \times Exposure$	0.0332**		-0.0067	-0.0070
	(0.014)		(0.008)	(0.010)
$DispHispanic_{d,2004} \times Exposure$		-0.0186	-0.0323***	-0.0273***
		(0.013)	(0.007)	(0.009)
$SpEd_{d,2004} \times Exposure$	-0.0036	-0.0402*	-0.0061	-0.0197*
	(0.033)	(0.021)	(0.010)	(0.011)
Mean Dept Var	0.680	0.667	0.788	0.728
College Enrollment				
$DispBlack_{d,2004} \times Exposure$	0.0772***		-0.0177	-0.0130
	(0.018)		(0.013)	(0.016)
$DispHispanic_{d,2004} \times Exposure$		0.0111	-0.0304**	-0.0446***
		(0.018)	(0.012)	(0.015)
$SpEd_{d,2004} \times Exposure$	-0.0500	-0.0591*	0.0437**	-0.0014
	(0.040)	(0.030)	(0.019)	(0.021)
Mean Dept Var	0.571	0.507	0.687	0.603
Associate's Degree				
$DispBlack_{d,2004} \times Exposure$	-0.0038		-0.0107**	-0.0118**
	(0.005)		(0.005)	(0.005)
$DispHispanic_{d,2004} \times Exposure$		-0.0046	-0.0136***	-0.0125***
		(0.005)	(0.004)	(0.004)
$SpEd_{d,2004} \times Exposure$	0.0022	-0.0336***	-0.0035	-0.0163***
	(0.010)	(0.010)	(0.005)	(0.006)
Mean Dept Var	0.033	0.054	0.062	0.055
Bachelor's Degree				
$DispBlack_{d,2004} \times Exposure$	0.0070		-0.0100	-0.0160***
	(0.007)		(0.008)	(0.006)
$DispHispanic_{d,2004} \times Exposure$		0.0008	-0.0020	-0.0168***
		(0.004)	(0.007)	(0.006)
$SpEd_{d,2004} \times Exposure$	-0.0211	-0.0074	0.0287***	0.0101
	(0.017)	(0.010)	(0.010)	(0.010)
Mean Dept Var	0.098	0.079	0.233	0.158
Observations	280,460	770,533	991,241	2,100,670

*** 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 4, except that we omit controls for disability type and classroom setting, and include controls for 5th grade math and reading standardized exam scores.

Table 7 District-Level Changes in the Composition of Black Students Who Lose SpEd at Expected 9th Grade, Given SpEd at 5th Grade

	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 8 Robustness Checks for Effect of Policy on SpEd students

	Black Students			Hispanic Students		
	Original	Dist-level Trends	Dist-level Trends	Original	Dist-level Trends	Dist-level Trends
SpEd Status G9	(1)	(2)	(3)	(4)	(5)	(6)
$Disp_{d,2004} \times Exposure$	-0.0946**	-0.1501***	-0.1484***	0.0064	0.0373	0.0163
	(0.039)	(0.041)	(0.040)	(0.030)	(0.037)	(0.033)
$SpEd_{d,2004} \times Exposure$	-0.2657***	-0.2764***	-0.2585***	-0.2471***	-0.2525***	-0.1657*
	(0.067)	(0.074)	(0.077)	(0.081)	(0.093)	(0.095)
High School Completion						
$Disp_{d,2004} \times Exposure$	0.0523**	0.0307	0.0350	-0.0281**	-0.0213	-0.0272*
	(0.0222)	(0.024)	(0.025)	(0.014)	(0.021)	(0.016)
$SpEd_{d,2004} \times Exposure$	-0.0047	-0.0132	0.0012	-0.0973***	-0.1071***	-0.1535***
	(0.037)	(0.044)	(0.046)	(0.034)	(0.037)	(0.035)
College Enrollment						
$Disp_{d,2004} \times Exposure$	0.0659***	0.0458**	0.0372*	-0.0155	-0.0396**	-0.0033
	(0.018)	(0.021)	(0.020)	(0.016)	(0.019)	(0.017)
$SpEd_{d,2004} \times Exposure$	0.0259	0.0308	0.0245	-0.0720**	-0.0846**	-0.1089***
	(0.038)	(0.041)	(0.041)	(0.035)	(0.035)	(0.041)
Associate's Degree						
$Disp_{d,2004} \times Exposure$	-0.0022	-0.0036	-0.0062	0.0058	-0.0023	0.0087*
	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)
$SpEd_{d,2004} \times Exposure$	-0.0025	-0.0021	-0.0023	-0.0143*	-0.0046	-0.0200**
	(0.009)	(0.014)	(0.014)	(0.008)	(0.008)	(0.010)
Bachelor's Degree						
$Disp_{d,2004} \times Exposure$	0.0007	-0.0031	-0.0014	-0.0004	-0.0016	-0.0023
	(0.004)	(0.004)	(0.005)	(0.003)	(0.004)	(0.004)
$SpEd_{d,2004} \times Exposure$	0.0008	0.0060	0.0060	0.0074	0.0150**	0.0075
	(0.006)	(0.008)	(0.008)	(0.006)	(0.007)	(0.006)
Demographic Chars		X	X		X	X
Urban/Rural and District Size			X			X

*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors are clustered at the district level. See Table 4 for full list of controls. In columns (1) and (4) we have our original main results from columns (3) and (6) of Table 4 for black and Hispanic students, respectively. In columns (2) and (5), we include indicator variables for each cohort year interacted with baseline demographics, including gender, ESL, FRL, Title I, gifted, and race. In columns (3) and (6), we include the same set of trends in demographic characteristics, and additionally include tends indicating whether the school is urban or rural and the district size at baseline.

Table 9 Robustness Checks for Effect of Policy on General Education students

	Black Students			Hispanic Students		
	Original	Dist-level Trends	Dist-level Trends	Original	Dist-level Trends	Dist-level Trends
SpEd Status G9	(1)	(2)	(3)	(4)	(5)	(6)
$Disp_{d,2004} \times Exposure$	-0.0232***	-0.0316***	-0.0335***	-0.0116**	-0.0126**	-0.0122**
	(0.008)	(0.010)	(0.011)	(0.005)	(0.006)	(0.006)
$SpEd_{d,2004} \times Exposure$	-0.0766***	-0.0638***	-0.0552***	-0.0437***	-0.0507***	-0.0490***
	(0.020)	(0.020)	(0.021)	(0.009)	(0.009)	(0.010)
High School Completion						
$Disp_{d,2004} \times Exposure$	0.0332**	0.0395**	0.0323**	-0.0186	-0.0024	0.0003
	(0.014)	(0.017)	(0.014)	(0.013)	(0.012)	(0.012)
$SpEd_{d,2004} \times Exposure$	-0.0036	-0.0387	-0.0125	-0.0402*	-0.0691***	-0.0667***
	(0.033)	(0.031)	(0.032)	(0.021)	(0.019)	(0.035)
College Enrollment						
$Disp_{d,2004} \times Exposure$	0.0772***	0.0488***	0.0384**	0.0111	0.0245*	0.0212
	(0.018)	(0.019)	(0.016)	(0.018)	(0.014)	(0.015)
$SpEd_{d,2004} \times Exposure$	-0.0500	-0.0299	-0.0115	-0.0591*	-0.0822***	-0.0870***
	(0.040)	(0.041)	(0.041)	(0.030)	(0.030)	(0.029)
Associate's Degree						
$Disp_{d,2004} \times Exposure$	-0.0038	0.0002	-0.0026	-0.0046	-0.0142**	-0.0123**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
$SpEd_{d,2004} \times Exposure$	0.0022	0.0001	0.0029	-0.0336***	-0.0338***	-0.0320***
	(0.010)	(0.011)	(0.012)	(0.010)	(0.010)	(0.010)
Bachelor's Degree						
$Disp_{d,2004} \times Exposure$	0.0070	0.0058	0.0050	0.0008	0.0118**	0.0111**
	(0.007)	(0.009)	(0.009)	(0.004)	(0.005)	(0.005)
$SpEd_{d,2004} \times Exposure$	-0.0211	-0.0234	-0.0181	-0.0074	-0.0118**	-0.0052
	(0.017)	(0.020)	(0.020)	(0.010)	(0.005)	(0.010)
Demographic Chars		X	X		X	X
Urban/Rural and District Size			X			X

*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors are clustered at the district level. See Table 6 for full list of controls and variable definitions. In columns (1) and (4) we have our original main results from columns (3) and (6) of Table 6 for black and Hispanic students, respectively. Remaining columns are as defined in Table 8.

Table 10 District Switching or Leaving for SpEd Sample

	Black Students			Hispanic Students		
	All	FRL	Non-FRL	All	FRL	Non-FRL
Enrolled G9	(1)	(2)	(3)	(4)	(5)	(6)
$Disp_{d,2004} \times Exposure$	0.0720**	0.0801**	0.0440	-0.0548***	-0.0530***	-0.0537
	(0.033)	(0.032)	(0.053)	(0.015)	(0.019)	(0.035)
$SpEd_{d,2004} \times Exposure$	-0.0708	-0.0571	-0.1806*	-0.0302	-0.0320	-0.0640
	(0.049)	(0.049)	(0.101)	(0.032)	(0.035)	(0.057)
Mean Dept Var	0.843	0.840	0.852	0.865	0.862	0.882
Observations	86,489	69,533	16,956	179,150	150,652	28,498
Switch Districts						
$Disp_{d,2004} \times Exposure$	0.1378***	0.1230**	0.2561***	-0.1105**	-0.1034**	-0.1755***
	(0.049)	(0.052)	(0.075)	(0.043)	(0.049)	(0.051)
$SpEd_{d,2004} \times Exposure$	-0.0927	-0.0868	-0.0459	-0.0450	-0.0548	0.0378
	(0.064)	(0.070)	(0.141)	(0.036)	(0.039)	(0.075)
Mean Dept Var	0.248	0.253	0.226	0.180	0.181	0.171
Observations	72,197	57,825	14,372	179,469	128,125	24,973

*** 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-level and cohort-district controls. See Table 4 for full set of controls. In the top panel we estimate the likelihood of being enrolled as of expected 9th for black and Hispanic SpEd students, respectively (given that they were enrolled in 5th grade). In the bottom panel we estimate the likelihood of switching districts between 5th and expected 9th grade. In columns (2) and (5) we condition on students being free or reduced-price lunch eligible (FRL) as of 5th grade, and in columns (3) and (6) we condition on non-FRL eligibility as of 5th grade.

Appendix

Figure A.1 PBMAS Manual 2004 Criteria for District-Level Black Disproportionality Rates

District Performance Level Criterion: District SPED African American Representation				
Performance Level (PL) Assignments				
Performance Level = Special Analysis	**Performance Level = 0 (met standard)**	**Performance Level = 1**	**Performance Level = 2**	**Performance Level = 3**
Fewer than 30 African American students or fewer than 30 students served in special education in the district in 2003-2004 and PL not equal to 0.	The district percent of special education students who are African American is no more than 1.0 percentage point higher than the percent of all district students who are African American. Minimum size requirements not applicable if PL = 0.	The district percent of special education students who are African American is between 1.1 and 2.0 percentage points higher than the percent of all district students who are African American.	The district percent of special education students who are African American is between 2.1 and 5.0 percentage points higher than the percent of all district students who are African American.	The district percent of special education students who are African American is at least 5.1 percentage points higher than the percent of all district students who are African American.

Source: Texas Performance Based Monitoring Analysis System Manual 2004.

Figure A.2 PBMAS Manual 2004 Criteria for District-Level Hispanic Disproportionality Rates

District Performance Level Criterion: District SPED Hispanic Representation				
Performance Level (PL) Assignments				
Performance Level = Special Analysis	**Performance Level = 0 (met standard)**	**Performance Level = 1**	**Performance Level = 2**	**Performance Level = 3**
Fewer than 30 Hispanic students or fewer than 30 students served in special education in the district in 2003-2004 and PL not equal to 0.	The district percent of special education students who are Hispanic is no more than 1.0 percentage point higher than the percent of all district students who are Hispanic. Minimum size requirements not applicable if PL = 0.	The district percent of special education students who are Hispanic is between 1.1 and 2.0 percentage points higher than the percent of all district students who are Hispanic.	The district percent of special education students who are Hispanic is between 2.1 and 5.0 percentage points higher than the percent of all district students who are Hispanic.	The district percent of special education students who are Hispanic is at least 5.1 percentage points higher than the percent of all district students who are Hispanic.

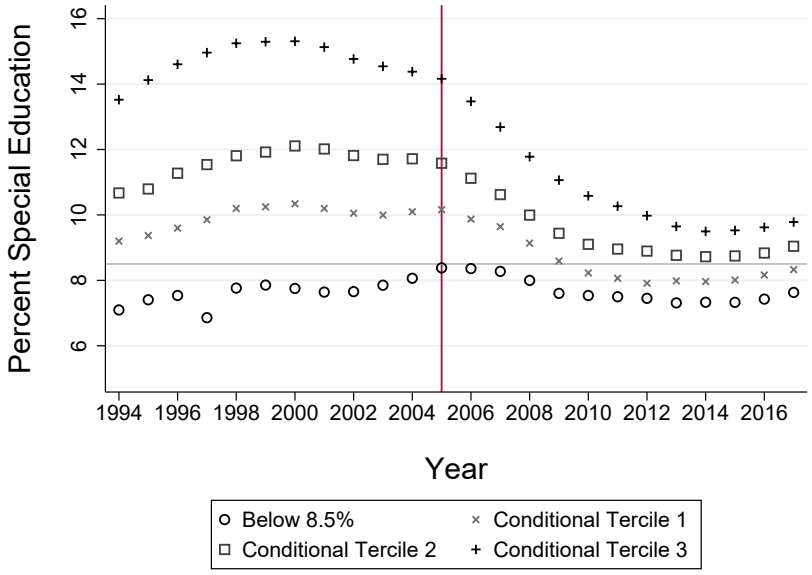
Source: Texas Performance Based Monitoring Analysis System Manual 2004.

Figure A.3 PBMAS Manual 2004 Criteria for District-Level Special Education Rates

District Performance Level Criterion: District Percentage of Students Receiving SPED Services				
Performance Level (PL) Assignments				
Performance Level = Special Analysis	**Performance Level = 0 (met standard)**	**Performance Level = 1**	**Performance Level = 2**	**Performance Level = 3**
Fewer than 30 students in special education in the district in 2003-2004 and PL not equal to 0.	The district identification of students to receive special education services is 8.5% or lower. Minimum size requirements not applicable if PL = 0.	The district identification of students to receive special education services is between 8.6% and 11.0%.	The district identification of students to receive special education services is between 11.1% and 16.0%.	The district identification of students to receive special education services is 16.1% or higher.

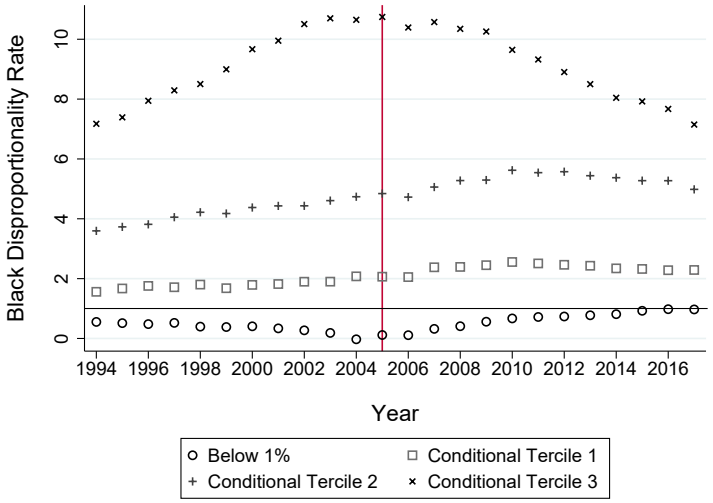
Source: Texas Performance Based Monitoring Analysis System Manual 2004.

Figure A.4 Percent of Students in SpEd by District SpEd Rate at Baseline



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

Figure A.5 Black Disproportionality Rate by District Black Disproportionality Rate at Baseline

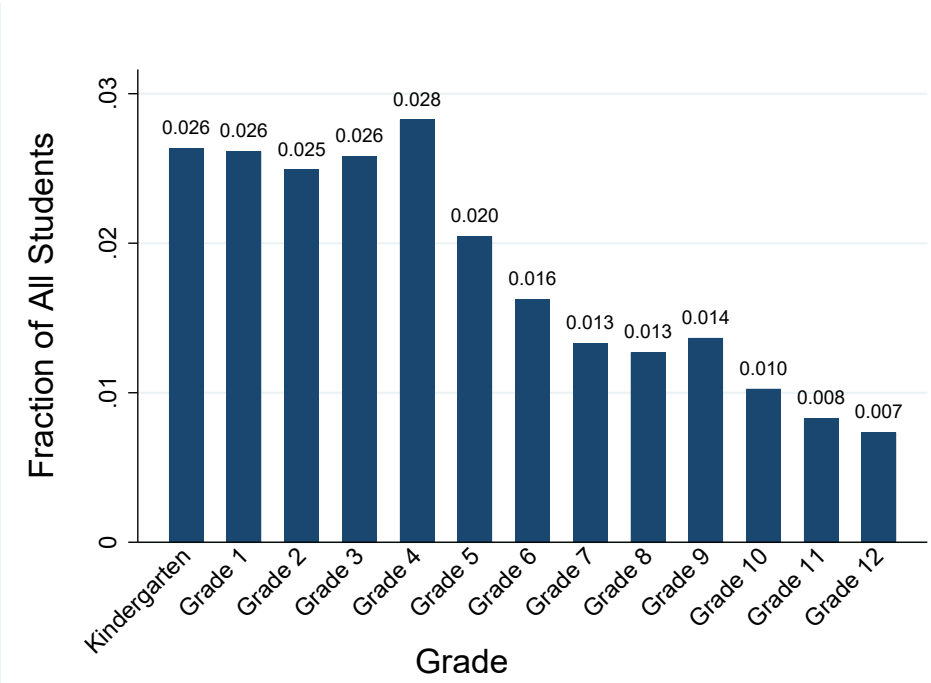


This figure plots the black disproportionality rate from 1994 to 2017. The bottom series consists of districts with black disproportionality rates less than 1% prior to 2004. The top three series split the remaining districts above the 1% black disproportionality threshold into terciles based on the pre-period black disproportionality rate.

Figure A.6 Hispanic Disproportionality by District Hispanic Disproportionality Rate at Baseline

This figure plots the Hispanic disproportionality rate from 1994 to 2017. The bottom series consists of districts with Hispanic disproportionality rate less than 1% prior to 2004. The top three series split the remaining districts above the 1% Hispanic disproportionality threshold into terciles based on the pre-period Hispanic disproportionality rate.

Figure A.7 Fraction of All Students Entering SpEd in Each Grade

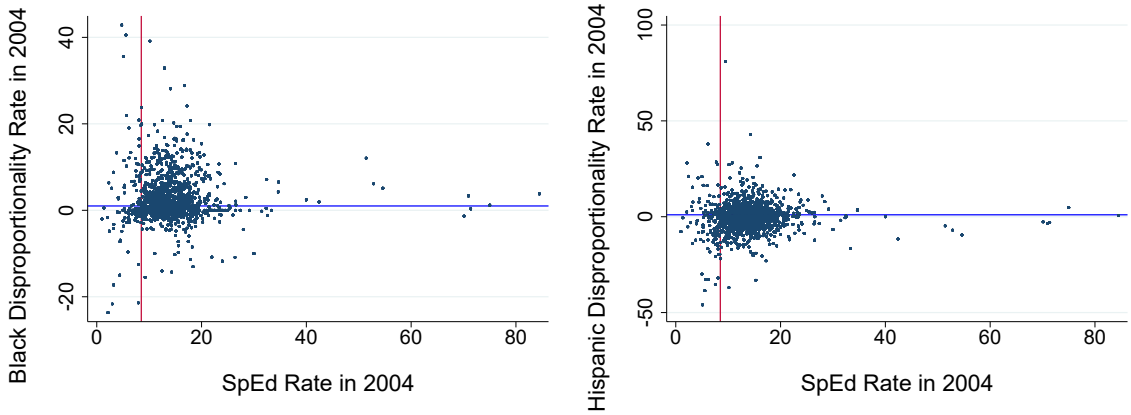


Each bar represents the fraction of students entering SpEd in each grade, out of the total number of students in each grade. This figure includes data from 1994 to 2017.

Figure A.8

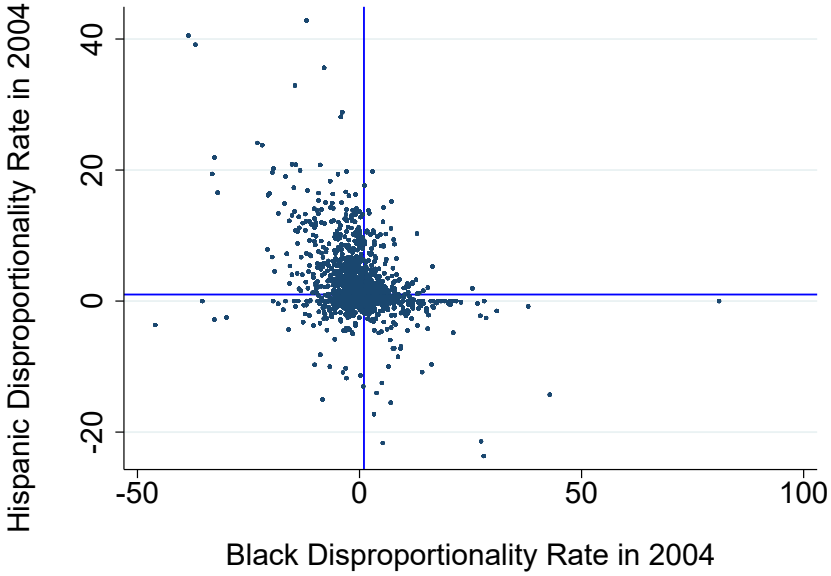
(a) District-Level Treatment Variation for Black Students

(b) District-Level Treatment Variation for Hispanic Students



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

Figure A.9 District-Level Treatment Variation in Black and Hispanic Disproportionality



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

Table A.1 Disability Type by Race

Disability Type	Black	Hispanic	White
Learning Disability	54.921	59.867	49.347
Speech Impairment	15.299	18.945	23.323
Intellectual Disability	10.107	6.138	4.221
Emotional Disturbance	8.508	5.140	7.980
Other Health Impairment	6.372	4.891	9.721
Autism	1.450	0.808	1.594
Orthopedic Impairment	1.021	1.488	1.279
Auditory Impairment	1.045	1.464	1.147
Visual Impairment	0.575	0.613	0.704
Noncategorical Early Childhood	0.479	0.441	0.367
Traumatic Brain Injury	0.156	0.141	0.196
Developmental Delay	0.035	0.036	0.096
Deaf/Blind	0.031	0.028	0.025

We present the percent of black, Hispanic, and white SpEd students with each disability type by race, for individuals in school in grades K to 12 prior to policy implementation.

Table A.2 Disability Type Before and After Policy Implementation

Disability Type	Pre-2005	Post-2005
Learning Disability	53.497	40.449
Speech Impairment	20.438	20.250
Other Health Impairment	7.729	12.715
Orthopedic Impairment	1.596	0.992
Emotional Disturbance	6.993	5.997
Intellectual Disability	5.940	7.874
Autism	1.479	7.967
Auditory Impairment	1.267	1.528
Visual Impairment	0.658	0.851
Deaf/Blind	0.028	0.043
Developmental Delay	0.064	0.055
Traumatic Brain Injury	0.180	0.300
Early Childhood Disability	0.432	0.979
Total	5,260,007	5,388,563

We present the percent of special education students with each disability type, for individuals in school in grades K to 12 before and after policy implementation.

Table A.3 Direct Effect on SpEd Students by Grade

	Black Students			Hispanic Students		
	4th	5th	6th	4th	5th	6th
SpEd Status	(1)	(2)	(3)	(4)	(5)	(6)
$Disp_{d,2004} \times Expo$	-0.0850**	-0.0946**	-0.1636**	0.0360	0.0264	0.0355
	(0.034)	(0.039)	(0.075)	(0.030)	(0.030)	(0.057)
$SpEd_{d,2004} \times Expo$	-0.1810***	-0.2657***	-0.4095***	-0.2536***	-0.2471***	-0.3866***
	(0.067)	(0.067)	(0.124)	(0.081)	(0.081)	(0.133)
Mean Dept Var	0.744	0.778	0.824	0.714	0.763	0.819
High School Completion						
$Disp_{d,2004} \times Expo$	0.0261	0.0523**	0.0774***	-0.0235*	-0.0281**	-0.0289*
	(0.021)	(0.022)	(0.027)	(0.014)	(0.014)	(0.016)
$SpEd_{d,2004} \times Expo$	-0.0258	-0.0047	0.0168	-0.1013***	-0.0973***	-0.0881**
	(0.032)	(0.037)	(0.047)	(0.033)	(0.034)	(0.042)
Mean Dept Var	0.614	0.597	0.590	0.609	0.589	0.575
College Enrollment						
$Disp_{d,2004} \times Expo$	0.0611***	0.0659***	0.0511**	-0.0196	-0.0155	-0.0179
	(0.017)	(0.018)	(0.024)	(0.015)	(0.016)	(0.019)
$SpEd_{d,2004} \times Expo$	0.0110	0.0259	0.0464	-0.0832**	-0.0720**	-0.0582
	(0.038)	(0.038)	(0.048)	(0.034)	(0.035)	(0.036)
Mean Dept Var	0.336	0.323	0.310	0.311	0.254	0.270
Associate's Degree						
$Disp_{d,2004} \times Expo$	-0.0008	-0.0022	0.0028	0.0054	0.0058	0.0042
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)
$SpEd_{d,2004} \times Expo$	-0.0021	-0.0025	-0.0099	-0.0018	-0.0143*	-0.0111
	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
Mean Dept Var	0.015	0.014	0.012	0.025	0.020	0.016
Bachelor's Degree						
$Disp_{d,2004} \times Expo$	0.0049	0.0007	0.0028	-0.0019	-0.0004	-0.0039
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)
$SpEd_{d,2004} \times Expo$	-0.0028	0.0008	-0.0035	0.0067	0.0074	0.0045
	(0.007)	(0.006)	(0.008)	(0.006)	(0.006)	(0.007)
Mean Dept Var	0.022	0.018	0.014	0.020	0.015	0.012
Observations	57,982	72,197	57,390	126,769	153,098	120,356

*** 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 4 for full set of controls. Each sample contains estimates for students in SpEd as of 4th,5th, or 6th grade prior to policy implementation.

Table A.4 Spillover Effect on General Education Students by Grade

	Black Students			Hispanic Students		
	4th	5th	6th	4th	5th	6th
SpEd Status G9	(1)	(2)	(3)	(4)	(5)	(6)
$Disp_{d,2004} \times Expo$	-0.0272***	-0.0232***	-0.0198*	-0.0189***	-0.0116**	-0.0187***
	(0.010)	(0.008)	(0.012)	(0.006)	(0.005)	(0.006)
$SpEd_{d,2004} \times Expo$	-0.1007***	-0.0766***	-0.0957***	-0.0536***	-0.0437***	-0.0548***
	(0.025)	(0.020)	(0.024)	(0.011)	(0.009)	(0.013)
Mean Dept Var	0.066	0.045	0.032	0.047	0.032	0.022
High School Completion	(1)	(2)	(3)	(4)	(5)	(6)
$Disp_{d,2004} \times Expo$	0.0309**	0.0332**	0.0371**	-0.0286***	-0.0186	0.0083
	(0.014)	(0.014)	(0.018)	(0.009)	(0.013)	(0.016)
$SpEd_{d,2004} \times Expo$	-0.0130	-0.0036	0.0207	-0.0255	-0.0402*	-0.0450
	(0.024)	(0.033)	(0.034)	(0.019)	(0.021)	(0.028)
Mean Dept Var	0.690	0.680	0.670	0.681	0.667	0.653
College Enrollment						
$Disp_{d,2004} \times Expo$	0.0805***	0.0772***	0.0734***	-0.0076	0.0111	0.0349*
	(0.018)	(0.018)	(0.024)	(0.016)	(0.018)	(0.021)
$SpEd_{d,2004} \times Expo$	-0.0360	-0.0500	-0.0336	-0.0548**	-0.0591*	-0.0629*
	(0.031)	(0.040)	(0.043)	(0.025)	(0.030)	(0.035)
Mean Dept Var	0.578	0.571	0.561	0.515	0.507	0.496
Associate's Degree						
$Disp_{d,2004} \times Expo$	0.0031	-0.0038	-0.0025	-0.0113**	-0.0046	-0.0004
	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)
$SpEd_{d,2004} \times Expo$	0.0056	0.0022	0.0015	-0.0273***	-0.0336***	-0.0374***
	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)
Mean Dept Var	0.034	0.033	0.031	0.058	0.054	0.051
Bachelor's Degree						
$Disp_{d,2004} \times Expo$	0.0058	0.0070	0.0091	-0.0061	0.0008	0.0085*
	(0.006)	(0.007)	(0.009)	(0.004)	(0.004)	(0.005)
$SpEd_{d,2004} \times Expo$	-0.0218*	-0.0211	-0.0277	0.0052	-0.0074	-0.0030
	(0.011)	(0.017)	(0.017)	(0.010)	(0.010)	(0.012)
Mean Dept Var	0.098	0.098	0.097	0.081	0.079	0.076
Observations	278,997	280,460	272,363	735,141	770,533	776,565

*** 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. Each sample contains estimates for students in SpEd as of 4th, 5th, or 6th grade prior to policy implementation. See Table 6 for full set of controls. However, for students in SpEd as of 4th grade, we control for 4th grade math and reading scores (and likewise for 5th and 6th grade SpEd samples).

Table A.5 Direct Effect on SpEd Students using Tercile Cutoffs of Disproportionality Cap

	Black Students				
	SpEd Status	High School Completion	College Enrollment	Associate's Degree	Bachelor's Degree
	(1)	(2)	(3)	(4)	(5)
$Disp\,1_{d,2004} \times Expo$	0.0070	0.0109***	0.0389	0.0005	0.0002
	(0.006)	(0.004)	(0.036)	(0.001)	(0.001)
$Disp\,2_{d,2004} \times Expo$	0.0007	0.0082**	0.0128***	0.0008	0.0011
	(0.005)	(0.003)	(0.003)	(0.001)	(0.001)
$Disp\,3_{d,2004} \times Expo$	-0.0027	0.0120***	0.0144***	0.0004	0.0006
	(0.005)	(0.003)	(0.003)	(0.001)	(0.001)
$SpEd_{d,2004} \times Expo$	-0.2656***	0.0029	0.0389	-0.0019	0.0015
	(0.068)	(0.033)	(0.036)	(0.010)	(0.006)
Mean Dept Var	0.778	0.597	0.323	0.014	0.018
Observations	72,199	72,199	72,199	72,199	72,199

	Hispanic Students				
	SpEd Status	High School Completion	College Enrollment	Associate's Degree	Bachelor's Degree
	(1)	(2)	(3)	(4)	(5)
$Disp\,1_{d,2004} \times Expo$	-0.0016	-0.0007	-0.0045	0.0011*	-0.0004
	(0.007)	(0.004)	(0.004)	(0.001)	(0.001)
$Disp\,2_{d,2004} \times Expo$	0.0089**	-0.0050	-0.0012	-0.0005	-0.0006
	(0.004)	(0.004)	(0.004)	(0.001)	(0.001)
$Disp\,3_{d,2004} \times Expo$	-0.0018	-0.0026	0.0006	0.0006	0.0001
	(0.004)	(0.002)	(0.002)	(0.001)	(0.000)
$SpEd_{d,2004} \times Expo$	-0.2362***	-0.1048***	-0.0702**	-0.0127	0.0090*
	(0.072)	(0.031)	(0.031)	(0.008)	(0.005)
Mean Dept Var	0.763	0.589	0.290	0.020	0.015
Observations	153,098	153,098	153,098	153,098	153,098

*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors clustered at the district level. Regressions include district and cohort fixed effects. See Table 4 for full set of controls. The district-level disproportionality rate in 2004 is split into four indicator variables. The first is an indicator for whether a district had below 1% disproportionality (i.e. was in compliance). Then districts are split into terciles above 1%. The indicator for being below 1% is excluded from the regression. $Disp1_{d,2004}$ corresponds to the first tercile above 1% (and so on for Disp 2, and Disp 3).

Table A.6 Effect of the Policy on 5th Grade Black SpEd Students by Disability Type

	Black Students						
	All	SLD	Speech	ED	OHI	ID	Physical
SpEd Status	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$Disp_{d,2004} \times Expo$	-0.0946**	-0.1054**	-0.0208	0.0178	-0.1114	-0.0461	0.0988
	(0.039)	(0.047)	(0.127)	(0.098)	(0.086)	(0.035)	(0.160)
$SpEd_{d,2004} \times Expo$	-0.2657***	-0.3083***	-0.5356**	-0.2444	0.0131	-0.0416	-0.0542
	(0.067)	(0.073)	(0.245)	(0.159)	(0.142)	(0.066)	(0.081)
Mean Dept Var	0.778	0.813	0.326	0.800	0.855	0.958	0.921
High School Completion							
$Disp_{d,2004} \times Expo$	0.0523**	0.0588**	-0.0182	-0.0071	0.0610	0.1748*	-0.0522
	(0.022)	(0.025)	(0.056)	(0.069)	(0.074)	(0.096)	(0.239)
$SpEd_{d,2004} \times Expo$	-0.0047	0.0146	0.0292	-0.0208	0.0765	0.3392**	-0.0014
	(0.037)	(0.047)	(0.118)	(0.142)	(0.149)	(0.156)	(0.105)
Mean Dept Var	0.597	0.601	0.645	0.450	0.631	0.602	0.676
College Enrollment							
$Disp_{d,2004} \times Expo$	0.0659***	0.0605***	0.0566	0.0006	0.1210*	-0.0704	0.2671
	(0.018)	(0.023)	(0.066)	(0.059)	(0.069)	(0.091)	(0.265)
$SpEd_{d,2004} \times Expo$	0.0259	0.0223	-0.0205	0.1162	-0.1565	0.4576***	-0.0268
	(0.038)	(0.041)	(0.127)	(0.121)	(0.143)	(0.143)	(0.095)
Mean Dept Var	0.323	0.321	0.484	0.278	0.385	0.128	0.410
Associate's Degree							
$Disp_{d,2004} \times Expo$	-0.0022	-0.0009	-0.0099	-0.0242	-0.0185	0.0119	-0.0185
	(0.004)	(0.005)	(0.018)	(0.015)	(0.018)	(0.013)	(0.093)
$SpEd_{d,2004} \times Expo$	-0.0025	0.0023	-0.1176***	-0.0042	-0.0229	-0.0255	0.1239***
	(0.009)	(0.010)	(0.032)	(0.025)	(0.033)	(0.017)	(0.042)
Mean Dept Var	0.014	0.012	0.028	0.009	0.018	0.002	0.029
Bachelor's Degree							
$Disp_{d,2004} \times Expo$	0.0007	-0.0078*	-0.0102	0.0130	0.0049	0.0030	0.1406*
	(0.004)	(0.005)	(0.029)	(0.011)	(0.017)	(0.004)	(0.082)
$SpEd_{d,2004} \times Expo$	0.0008	0.0058	-0.0121	-0.0147	-0.0572**	0.0054	0.0595
	(0.006)	(0.006)	(0.059)	(0.026)	(0.029)	(0.014)	(0.039)
Mean Dept Var	0.018	0.013	0.064	0.009	0.014	0.002	0.029
Observations	72,197	44,251	8,008	5,905	5,197	6,408	1,463

*** 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 4 for full set of controls. Disability type is measured as of 5th grade.

Table A.7 Direct Effect on Black SpEd Students by Gender and Economic Disadvantage

	Male	Female	FRL	Non-FRL
SpEd Status	(1)	(2)	(3)	(4)
$Disp_{d,2004} \times Expo$	-0.0799**	-0.1377**	-0.0983**	-0.0866
	(0.038)	(0.055)	(0.039)	(0.066)
$SpEd_{d,2004} \times Expo$	-0.2604***	-0.2828***	-0.3058***	-0.1508
	(0.070)	(0.102)	(0.067)	(0.112)
Mean Dept Var	0.792	0.752	0.797	0.703
High School Completion				
$Disp_{d,2004} \times Expo$	0.0831***	-0.0136	0.0557**	0.0002
	(0.025)	(0.030)	(0.023)	(0.040)
$SpEd_{d,2004} \times Expo$	0.0157	-0.0383	0.0098	-0.1380*
	(0.049)	(0.042)	(0.042)	(0.071)
Mean Dept Var	0.584	0.623	0.572	0.697
College Enrollment				
$Disp_{d,2004} \times Expo$	0.0812***	0.0264	0.0464**	0.1361***
	(0.021)	(0.028)	(0.020)	(0.047)
$SpEd_{d,2004} \times Expo$	0.0337	0.0075	0.0249	0.0528
	(0.041)	(0.048)	(0.036)	(0.085)
Mean Dept Var	0.304	0.360	0.286	0.472
Associate's Degree				
$Disp_{d,2004} \times Expo$	-0.0043	0.0004	-0.0038	0.0017
	(0.005)	(0.008)	(0.004)	(0.015)
$SpEd_{d,2004} \times Expo$	-0.0058	0.0002	-0.0038	-0.0111
	(0.007)	(0.020)	(0.009)	(0.024)
Mean Dept Var	0.012	0.017	0.010	0.030
Bachelor's Degree				
$Disp_{d,2004} \times Expo$	0.0006	0.0024	-0.0035	0.0294*
	(0.005)	(0.007)	(0.004)	(0.016)
$SpEd_{d,2004} \times Expo$	-0.0005	0.0008	0.0003	0.0032
	(0.008)	(0.011)	(0.006)	(0.026)
Mean Dept Var	0.016	0.022	0.011	0.047
Observations	47,249	24,948	57,825	14,372

*** 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 4 for full set of controls. Gender and FRL status are measured as of 5th grade.

Table A.8 Spillover Effect on Black General Education Students by Gender and Economic Disadvantage

	Male	Female	FRL	Non-FRL
SpEd Status G9	(1)	(2)	(3)	(4)
$Disp_{d,2004} \times Expo$	-0.0342***	-0.0144*	-0.0326***	-0.0031
	(0.011)	(0.009)	(0.010)	(0.009)
$SpEd_{d,2004} \times Expo$	-0.0964***	-0.0598***	-0.0961***	-0.0291
	(0.029)	(0.017)	(0.023)	(0.020)
Mean Dept Var	0.041	0.020	0.034	0.021
High School Completion				
$Disp_{d,2004} \times Expo$	0.0505***	0.0189	0.0280*	0.0350**
	(0.018)	(0.015)	(0.017)	(0.015)
$SpEd_{d,2004} \times Expo$	-0.0320	0.0186	-0.0129	0.0016
	(0.035)	(0.036)	(0.039)	(0.032)
Mean Dept Var	0.679	0.736	0.652	0.823
College Enrollment				
$Disp_{d,2004} \times Expo$	0.0858***	0.0679***	0.0576***	0.1032***
	(0.022)	(0.019)	(0.019)	(0.020)
$SpEd_{d,2004} \times Expo$	-0.0357	-0.0635	-0.0595	-0.0209
	(0.038)	(0.046)	(0.042)	(0.040)
Mean Dept Var	0.540	0.645	0.537	0.714
Associate's Degree				
$Disp_{d,2004} \times Expo$	0.0023	-0.0099	-0.0059	0.0037
	(0.005)	(0.006)	(0.004)	(0.010)
$SpEd_{d,2004} \times Expo$	0.0052	-0.0001	0.0144	-0.0047
	(0.009)	(0.014)	(0.009)	(0.019)
Mean Dept Var	0.025	0.043	0.028	0.048
Bachelor's Degree				
$Disp_{d,2004} \times Expo$	-0.0004	0.0117	-0.0044	0.0409***
	(0.008)	(0.010)	(0.006)	(0.013)
$SpEd_{d,2004} \times Expo$	-0.0054	-0.0340	-0.0141	-0.0163
	(0.019)	(0.021)	(0.015)	(0.027)
Mean Dept Var	0.079	0.131	0.069	0.182
Observations	130,128	150,332	186,297	94,163

*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors clustered at the district level. Regressions include district and cohort fixed effects. See Table 6 for full set of controls. Gender and FRL status are measured as of 5th grade.

Table A.9 Direct Effect on Black SpEd Students by District-Level Location and Tax-Based Wealth

	Location		Wealth	
	Rural	Urban	Above Median	Below Median
SpEd Status	(1)	(2)	(3)	(4)
$Disp_{d,2004} \times Expo$	-0.0178	-0.1477***	-0.0134	-0.1504***
	(0.062)	(0.053)	(0.057)	(0.057)
$SpEd_{d,2004} \times Expo$	-0.1995**	-0.2909***	-0.2739**	-0.3267***
	(0.085)	(0.109)	(0.109)	(0.079)
Mean Dept Var	0.841	0.757	0.770	0.787
High School Completion				
$Disp_{d,2004} \times Expo$	0.0237	0.0537*	0.0558	0.0422
	(0.027)	(0.031)	(0.035)	(0.028)
$SpEd_{d,2004} \times Expo$	0.0547	0.0280	-0.0289	0.0453
	(0.057)	(0.060)	(0.045)	(0.055)
Mean Dept Var	0.695	0.568	0.579	0.619
College Enrollment				
$Disp_{d,2004} \times Expo$	0.0128	0.0919***	0.0721**	0.0656***
	(0.030)	(0.025)	(0.029)	(0.022)
$SpEd_{d,2004} \times Expo$	0.0895	0.0371	0.0103	-0.1514***
	(0.056)	(0.053)	(0.053)	(0.046)
Mean Dept Var	0.313	0.329	0.323	0.323
Associate's Degree				
$Disp_{d,2004} \times Expo$	-0.0119*	0.0024	-0.0005	-0.0073
	(0.006)	(0.006)	(0.006)	(0.005)
$SpEd_{d,2004} \times Expo$	-0.0023	-0.0021	0.0163	-0.0122
	(0.017)	(0.016)	(0.015)	(0.012)
Mean Dept Var	0.012	0.014	0.014	0.013
Bachelor's Degree				
$Disp_{d,2004} \times Expo$	-0.0048	0.0059	0.0022	-0.0020
	(0.005)	(0.006)	(0.006)	(0.005)
$SpEd_{d,2004} \times Expo$	0.0187	-0.0027	-0.0014	0.0137
	(0.014)	(0.009)	(0.009)	(0.009)
Mean Dept Var	0.011	0.021	0.020	0.015
Observations	16,264	53,257	39,567	32,630

*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors clustered at the district level. Regressions include district and cohort fixed effects. See Table 4 for full set of controls. Schools are deemed rural if they are categorized as town-fringe, town-distant, town-remote, rural-fringe, rural-distant, or rural-remote. Schools are deemed urban if they are categorized as suburban-small, suburban-midsize, suburban-large, city-small, city-midsize, or city-large. Districts with above median wealth have tax-based wealth per pupil above $291,434.

Table A.10 Prediction of SpEd Status for White Students

Variable	
Age	0.3429***
	(0.018)
Male	0.5232***
	(0.016)
FRL	0.0549***
	(0.020)
ESL	-0.2168
	(0.219)
Bilingual	0.2735
	(0.274)
Title I	0.0311
	(0.038)
At Risk	0.5850***
	(0.033)
Gifted	-0.5352***
	(0.045)
LEP	-0.7360***
	(0.185)
Migrant	-0.2815
	(0.459)
Math	-0.4843***
	(0.012)
Reading	-0.6610***
	(0.009)
District Controls	X
Grade Controls	X
Constant	-6.2820***
	(0.651)
Observations	978,804

*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors are clustered at the district level. We use a logit model to predict 5th grade SpEd status for white students only, using covariates measured as of 3rd grade. Regressions include district and year fixed effects. Age is measured as of September 1st of the current year. Limited English Proficiency (LEP) and Migrant are imputed such that their value is set to 0 if missing in the original data. At Risk indicates that a student did not perform satisfactorily on a readiness test or assessment during the current school year. District and grade-level measures include the same set of variables as used at the individual level.

Table A.11 Effect of Policy by Prediction of District Over- or Under-Representation Relative to White Students

	Black Students	
	Over	Under
SpEd Status	(1)	(2)
$Disp_{d,2004} \times Expo$	-0.0713	-0.0766
	(0.060)	(0.048)
$SpEd_{d,2004} \times Expo$	-0.2184**	-0.2565***
	(0.086)	(0.084)
Mean Dept Var	0.785	0.775
High School Completion		
$Disp_{d,2004} \times Expo$	0.0201	0.0453
	(0.031)	(0.030)
$SpEd_{d,2004} \times Expo$	-0.0001	0.0038
	(0.060)	(0.050)
Mean Dept Var	0.647	0.573
College Enrollment		
$Disp_{d,2004} \times Expo$	0.0570**	0.0274
	(0.027)	(0.025)
$SpEd_{d,2004} \times Expo$	-0.0055	0.0613
	(0.049)	(0.046)
Mean Dept Var	0.352	0.309
Associate's Degree		
$Disp_{d,2004} \times Expo$	-0.0050	-0.0033
	(0.007)	(0.005)
$SpEd_{d,2004} \times Expo$	0.0019	-0.0029
	(0.013)	(0.014)
Mean Dept Var	0.017	0.012
Bachelor's Degree		
$Disp_{d,2004} \times Expo$	-0.0046	-0.0005
	(0.005)	(0.006)
$SpEd_{d,2004} \times Expo$	0.0051	-0.0011
	(0.010)	(0.007)
Mean Dept Var	0.020	0.017
Observations	23,228	49,130

*** 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 4 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.

Table A.12 Direct Effect on Black SpEd Students by Teacher Experience and Racial Composition

	Experience		Racial Composition	
	Above Average	Below Average	Above Median	Below Median
SpEd Status	(1)	(2)	(3)	(4)
$Disp_{d,2004} \times Expo$	-0.1379***	-0.0749	-0.1314***	-0.0496
	(0.041)	(0.070)	(0.044)	(0.072)
$SpEd_{d,2004} \times Expo$	-0.3920***	-0.1266	-0.2823***	-0.1780**
	(0.075)	(0.087)	(0.084)	(0.084)
Mean Dept Var	0.808	0.741	0.774	0.798
High School Completion				
$Disp_{d,2004} \times Expo$	0.0178	0.1092***	0.0461*	0.0594
	(0.026)	(0.033)	(0.026)	(0.052)
$SpEd_{d,2004} \times Expo$	-0.0176	-0.0484	-0.0163	0.0668
	(0.052)	(0.051)	(0.043)	(0.081)
Mean Dept Var	0.590	0.606	0.587	0.648
College Enrollment				
$Disp_{d,2004} \times Expo$	0.0726***	0.0864***	0.0500**	0.0679
	(0.023)	(0.032)	(0.021)	(0.049)
$SpEd_{d,2004} \times Expo$	0.0926**	-0.0571	0.0364	0.0203
	(0.049)	(0.047)	(0.046)	(0.069)
Mean Dept Var	0.285	0.371	0.318	0.350
Associate's Degree				
$Disp_{d,2004} \times Expo$	-0.0021	0.0049	-0.0062	0.0158
	(0.004)	(0.008)	(0.004)	(0.012)
$SpEd_{d,2004} \times Expo$	-0.0030	0.0104	-0.0014	0.0109
	(0.009)	(0.019)	(0.012)	(0.014)
Mean Dept Var	0.011	0.016	0.013	0.017
Bachelor's Degree				
$Disp_{d,2004} \times Expo$	-0.0008	0.0034	-0.0001	0.0158
	(0.004)	(0.009)	(0.005)	(0.015)
$SpEd_{d,2004} \times Expo$	0.0001	-0.0015	0.0024	0.0025
	(0.008)	(0.009)	(0.007)	(0.015)
Mean Dept Var	0.013	0.024	0.018	0.021
Observations	40,177	32,020	59,728	12,469

*** p<0.01, ** p<0.05, * p<0.1 Robust standard errors clustered at the district level. Regressions include district and cohort fixed effects. See Table 4 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 median level (i.e. 3.6%) in columns (3) and (4).

Table A.13 Effect of Policy on SpEd Students' Intermediate Outcomes

	Black Students					
	Absences			**Disciplinary Actions**		
	(1)	(2)	(3)	(4)	(5)	(6)
	% Days Absent	3+ Truant	10+ Truant	Suspended	Mult. Suspended	Expulsion
$Disp_{d,2004} \times Post$	0.0554***	0.0029	0.0134	0.3423***	0.3314***	0.0308
	(0.018)	(0.016)	(0.016)	(0.118)	(0.110)	(0.021)
$SpEd_{d,2004} \times Post$	-0.0313	-0.0828	-0.0451	0.0451	0.1502	0.1460***
	(0.023)	(0.057)	(0.029)	(0.158)	(0.160)	(0.035)
Mean Dept Var	0.063	0.002	0.023	0.259	0.170	0.013
Observations	752,071	910,729	910,729	910,729	910,729	910,729
	Hispanic Students					
	Absences			**Disciplinary Actions**		
	(1)	(2)	(3)	(4)	(5)	(6)
	% Days Absent	3+ Truant	10+ Truant	Suspended	Mult. Suspended	Expulsion
$Disp_{d,2004} \times Post$	0.0189	-0.0444**	-0.0118	-0.0518	-0.0515	0.0302
	(0.017)	(0.020)	(0.022)	(0.078)	(0.058)	(0.020)
$SpEd_{d,2004} \times Post$	-0.0501*	0.0029	-0.0064	-0.0174	0.0254	0.0351
	(0.030)	(0.029)	(0.036)	(0.164)	(0.144)	(0.029)
Mean Dept Var	0.064	0.002	0.003	0.189	0.116	0.011
Observations	1,624,095	1,950,302	1,950,302	1,950,302	1,950,302	1,950,302

*** p<0.01, ** p<0.05, * p<0.1 Regressions are run on individuals who were in SpEd as of 5th grade prior to policy implementation. Outcomes are now measured in each grade. Controls include individual-level gender, age, ESL, FRL, title I, and gifted status. District and grade-level controls include gender, ESL, FRL, Title I, gifted, and racial composition. Regressions also include year, district, and grade fixed effects. Robust standard errors are clustered at the district level. $Disp_{d,2004} \times Post$ is the district-level 2004 black or Hispanic disproportionality rate interacted with an indicator for the post-policy period. $SpEd_{d,2004} \times Post$ is the district-level 2004 SpEd rate interacted with an indicator for the post-policy period. % Days Absent is the percent of school days an individual was absent in a given school year. 3 Truant indicates that a student had 3 unexcused absences. 10 Truant indicates that a student had 10 unexcused absences. Suspended is an indicator for whether a student was ever suspended, including both in-school and out-of-school suspensions. Mult. Suspended is an indicator for being suspended multiple times. Expulsion is an indicator for being expelled or otherwise displaced from school (e.g. this includes placement in a juvenile justice setting).

Table A.14 District Leaving by Grade for SpEd Sample

	Black Students			Hispanic Students		
	All	FRL	Non-FRL	All	FRL	Non-FRL
Enrolled G6	(1)	(2)	(3)	(4)	(5)	(6)
$Disp_{d,2004} \times Expo$	0.0166	0.0201	0.0023	-0.0009	0.0022	-0.0133
	(0.016)	(0.017)	(0.034)	(0.011)	(0.011)	(0.022)
$SpEd_{d,2004} \times Expo$	-0.0806**	-0.0796**	-0.0499	-0.0205	-0.0187	-0.0538
	(0.032)	(0.034)	(0.071)	(0.023)	(0.023)	(0.051)
Mean Dept Var	0.934	0.937	0.923	0.950	0.951	0.944
Observations	86,489	69,533	16,956	179,150	150,652	28,498
Enrolled G7						
$Disp_{d,2004} \times Expo$	0.0245	0.0319*	0.0096	-0.0165	-0.0147	-0.0121
	(0.018)	(0.017)	(0.052)	(0.013)	(0.014)	(0.026)
$SpEd_{d,2004} \times Expo$	-0.0661*	-0.0562	-0.0750	-0.0013	0.0024	-0.0582
	(0.034)	(0.038)	(0.080)	(0.029)	(0.032)	(0.044)
Mean Dept Var	0.905	0.908	0.895	0.924	0.925	0.921
Observations	86,489	69,533	16,956	179,150	150,652	28,498
Enrolled G8						
$Disp_{d,2004} \times Expo$	0.0624**	0.0615**	0.0665	-0.0386***	-0.0403***	-0.0211
	(0.025)	(0.026)	(0.043)	(0.014)	(0.014)	(0.030)
$SpEd_{d,2004} \times Expo$	-0.0846*	-0.0697	-0.2152**	0.0040	0.0088	-0.0535
	(0.044)	(0.048)	(0.084)	(0.031)	(0.032)	(0.049)
Mean Dept Var	0.876	0.877	0.872	0.899	0.898	0.904
Observations	86,489	69,533	16,956	179,150	150,652	28,498

*** 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 4 for full set of controls. Enrolled G6 indicates than an individual was enrolled in the data at expected 6th grade (given that they were enrolled in 5th grade). FRL status is measured as of 5th grade.

Table A.15 Effect of Policy on General and Special Education Spending

	Special Education Spending			
	(1) SpEd Spending Per All	(2) SpEd Spending Per SpEd	(3) Instr. SpEd Spending Per All	(4) Instr. SpEd Spending Per SpEd
$Disp\,Black_{d,2004} \times Post$	-391	3,575	-294	2,635
	(331)	(3,283)	(249)	(2,728)
$Disp\,Hispanic_{d,2004} \times Post$	51	6,211**	92	4,835**
	(324)	(2,960)	(228)	(2,221)
$SpEd_{d,2004} \times Post$	-1,341*	1,379	-1,067**	-1,167
	(752)	(4,132)	(490)	(3,009)
Mean Dept Var	773	10,172	773	7,594

	General Education Spending			
	(1) GE Spending Per All	(2) GE Spending Per GE	(3) Instr. GE Spending Per All	(4) Instr. GE Spending Per GE
$Disp\,Black_{d,2004} \times Post$	783	785	42	-89
	(799)	(940)	(550)	(634)
$Disp\,Hispanic_{d,2004} \times Post$	-543	-717	-258	-504
	(724)	(827)	(500)	(552)
$SpEd_{d,2004} \times Post$	1,112	-1,385	738	-827
	(1,019)	(1,166)	(703)	(794)
Mean Dept Var	4,272	4,781	3,480	3,892

*** p<0.01, ** p<0.05, * p<0.1 Regressions are run at the district-level and include controls for district-level gender, ESL, FRL, title I, gifted, and racial composition. Regressions include district and year fixed effects. Robust standard errors are clustered at the district level. Instr. stands for instructional expenditures.