The Long-Run Impacts of Special Education

Briana Ballis† and Katelyn Heath*

December 18, 2020

Latest version available at: `https://brianaballis.weebly.com`

Abstract

Over 13 percent of US students participate in Special Education (SE) programs annually, at a cost of $40 billion. However, due to selection issues the effect of SE placements remains unclear. This paper uses administrative data from Texas to examine the long-run effect of reducing SE access. Our research design exploits variation in SE placement driven by a unique state policy that required school districts to reduce SE caseloads to 8.5 percent. This policy led to sharp reductions in SE enrollment. These reductions in SE access generated significant reductions in educational attainment, suggesting that marginal participants experience long-run benefits from SE services.

---

† University of California, Merced. Department of Economics. Email: bballis@ucmerced.edu. * Rice University. Texas Policy Lab. Email: kmh8@rice.edu.

We thank Scott Carrell, Paco Martorell, Marianne Page, Marianne Bitler, Maria Fitzpatrick, Michael Lovenheim, Douglas Miller, Jordan Matsudaira, Michal Kurlaender, and Jacob Hibel for useful discussions and suggestions. We would like to thank participants at the UC Davis Applied Microeconomics Brown Bag, the Association for Education Finance and Policy 43rd Annual Conference, the UC Davis Alumni conference, the Western Economic Association International 93rd Annual Conference, the 40th annual Association for Public Policy Analysis and Management Conference, and Claremont McKenna - Southern California Conference in Applied Microeconomics (SoCCAM) for valuable comments. We also thank everyone at the UT Dallas Research Center who have helped us get acquainted with the administrative data. 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, the Texas Workforce Commission, or the state of Texas. This material is based upon work supported by the National Science Foundation under Grant Number (#1824547) and a grant from the American Educational Research Association which receives funds for its "AERA Grants Program" from the National Science Foundation under NSF award NSF-DRL (#1749275). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or AERA. This project received support from the Social Security Administration (SSA) funded as part of the Disability Research Consortium. The opinions and conclusions expressed are solely those of the authors and do not represent the opinions or policy of SSA or any agency of the Federal Government. Briana Ballis acknowledges financial support from the Disability Research Consortium Fellowship from SSA via Mathematica Policy, the National Science Foundation, and the UC Davis Economics department. Katelyn Heath acknowledges financial support from the National Academy of Education and the National Academy of Education/Spencer Dissertation Fellowship Program, the Bronfenbrenner Center for Translational Research, Christopher Wildeman, Cornell University Graduate School, Cornell Policy Analysis and Management Department, and the Cornell Economics department Labor Grant. All errors are our own.

1 Introduction

In this paper, we provide evidence on the long-run causal effects of Special Education (SE) by exploiting a unique policy change that introduced exogenous variation in SE participation. In 2005, the Texas Department of Education implemented a district-level SE enrollment target of 8.5 percent. Over the next ten years, statewide SE enrollment declined by 4.5 percentage points, from 13 percent to 8.5 percent. By 2018, roughly 225,000 fewer students were enrolled in SE programs annually across the state.[1] To our knowledge, this is the first major policy change that caused such a large and sudden change in SE participation for a large representative sample of students. In 2016, more than 10 years after the policy was implemented, an investigative Houston Chronicle article was published that alerted the greater public, for the first time, to the existence of this policy, sparking significant public outcry and debate (Rosenthal, 2016). Subsequently, the federal government conducted their own investigation and in 2018 determined that the 8.5 percent district SE target was illegal and in violation of federal disability law (U.S. Department of Education, 2018). The fact that this policy was illegal highlights why policies such as this one are so rare. We use this policy change in combination with administrative data from Texas that follows the universe of public school students into adulthood, to provide the first long-run causal estimates of losing access to SE programs.

In the United States, SE program participation grew by over 40 percent between 1975 and 2018. Currently, over 13 percent of public school students participate in SE programs annually, at a cost of $40 billion (National Center for Education Statistics, 2015; Elder, Figlio, Imberman, & Persico, in press). A-priori the net-benefits of SE participation are ambiguous. Students are likely to benefit from the *individualized* educational support (such as one-on-one tutoring, a classroom aide, therapy, or standardized testing modifications) that SE offers. However, for students with less severe conditions there are several reasons why SE participation could be harmful: being placed in segregated learning environments or held to relatively lower expectations regarding achievement may inhibit long-run success. While the purpose of SE is to ameliorate the challenges students with

[1]This is computed by multiplying total enrollment in Texas public schools in 2018 (roughly 5 million) by the 4.5 percentage point reduction in total SE enrollment that occurred post-policy.

disabilities may face throughout schooling and later in life, considerable uncertainty surrounds the effectiveness of SE spending.

Despite significant increases in SE participation, there is little causal evidence on how SE placement (or lack thereof) affects long-run trajectories of marginal participants. The main difficulty in evaluating the effectiveness of SE programs is identifying a plausible counterfactual. Students are selected into SE because they are at risk of low achievement or poor behavioral outcomes. In addition, SE inclusion criteria are neither straightforward nor standardized for students with less severe conditions. Thus, it is not possible to exploit discontinuities in SE diagnostic criteria to identify the causal impacts of SE participation.[2] Instead, exogenous changes in SE participation are required for causal identification. However, SE eligibility rules were determined federally in 1975 (with minor changes since), making it difficult to identify variation in SE placement across locations or over time that is plausibly exogenous.

Our research design exploits pre-policy variation in SE rates across districts in Texas, which led to significant differences in policy pressure to reduce SE enrollment. To identify the direct impacts of SE programs for students with disabilities, we focus on students who were already enrolled in SE prior to the policy change and estimate the long-run effect of a reduction in SE access using two different identification strategies.[3] First, we use a difference-in-differences (DiD) strategy that compares changes in SE removal and educational attainment in districts with lower versus higher pre-policy SE rates, across cohorts who had varying levels of time exposed to the policy. This strategy produces estimates of the average effect of reducing overall *access* to SE for students with disabilities. Second, we use exposure to the policy as an instrument for SE removal in an instrumental variables (IV) framework. This second strategy allows us to identify the long-run

[2]For the vast majority of SE students with learning or behavioral impairments, the most common symptoms are poor academic performance or classroom behaviors. Since many students exhibit these symptoms occasionally, there are inconsistencies in SE placement based on how teachers, parents, or diagnosticians perceive these symptoms.

[3]The policy pressure to reduce SE enrollment significantly changed the incentives to classify marginal students, which in turn changed the underlying conditions of SE students in the post-policy period. This sample restriction ensures that our results will not be negatively biased as a result of the underlying changes in the ability of students in SE after the policy change. As we will justify in Section 3.2, we focus on 5th grade SE cohorts, but demonstrate that our results are not sensitive to this grade cohort restriction.

impacts of SE *removal* for students on the margin of SE placement decisions, precisely the group for whom the net benefits of SE are most unclear.

Our results suggest that SE students who lose access to SE services experience significant declines in educational attainment. Our DiD estimates imply that SE students enrolled in the average district experienced a 3.7 percentage points (or 13 percent) increase in the likelihood of SE removal, a 2.0 percentage points (or 2.7 percent) decrease in the likelihood of high school completion, and a 1.2 percentage points (or 3.6 percent) decrease in the likelihood of college enrollment after the policy's introduction.[4] We show that these declines in SE participation and educational attainment are mostly driven by SE students who already spent the majority of their day in general education (GE) classrooms and who had relatively mild disability types.[5] For students on the margin of SE placement decisions, our IV estimates imply that SE removal decreased high school completion by 51.9 percentage points and college enrollment by 37.9 percentage points. These large reductions in high school completion and college enrollment are suggestive that later life labor market outcomes may also be likely to decline.[6] Lower-income and minority students experience larger increases in SE removal, and the negative impact of SE removal on educational attainment is concentrated among these students.

Why do marginal SE students (i.e. those with relatively mild conditions) experience such large reductions in educational attainment after SE removal? There are at least two possible explanations. First, SE students may have the option to satisfy modified high school graduation requirements, which may make educational attainment easier. That is, students enrolled in SE may be able to graduate from high school without passing an exit exam, which is a graduation requirement for all GE students. Thus, students may mechanically be less likely to graduate from high school if they no longer have the option of being exempt from the exit exam. Second,

[4] These effect sizes are computed for SE students fully exposed to the policy after 5th grade and enrolled in the average school district that served 13 percent of their students in SE at baseline.

[5] Therefore, throughout this paper when considering students who are removed from SE, we view this as a removal of services within the GE classroom, rather than a compositional shift of students from SE classrooms to GE classrooms, for the majority of SE students.

[6] However, we leave these results for future work, given that the limited number of post-policy years has not yet allowed these outcomes to fully realize.

SE students likely benefit from the additional resources and more focused attention they receive. Therefore, when SE services are removed, students may experience a reduction in learning, which makes it more difficult for them to complete high school and attain post-secondary schooling.

We find that students who lose access to SE are significantly more likely to take the high school exit exam, suggesting that this mechanical effect could have played a role. However, we show that it is unlikely that changes in high school graduation requirements alone are driving our results. The largest declines in high school completion are driven by students who would have been very likely to take the exit exam regardless of the policy, likely ruling out a purely mechanical effect.[7] Moreover, the long-run negative impacts of SE removal are larger for students in lower-resourced districts. This highlights the potential importance of additional SE resources, especially in districts with fewer available resources to prepare students with special needs for adulthood. Finally, it is important to highlight that we are estimating SE program effects based on SE removal. Long-run responses for students who never participate in SE, but would have in the absence of the policy, may not mirror the impacts of SE removal. For instance, those never enrolled in SE do not incur any potential stigma associated with a disability label and do not become accustomed to additional supports during school. However, we are not able to identify students in the post-policy period who would have been in SE in the absence of the policy, but are now not.

Even though the policy was targeted towards SE students, limiting access to SE programs could have affected all students in Texas public schools. For instance, GE students could have experienced spillover effects resulting from their SE peers losing access to services. Additionally, GE students could have been directly impacted by the policy as a result of being less likely to gain access to SE in later grades themselves. Therefore, we also estimate the effects of this policy on a sample of GE students using a similar DiD model.[8] We find smaller and less precise declines in

[7] We do not find changes in the likelihood of taking or passing the exit exam for students predicted to be very likely to take the exit exam regardless of the policy. Although we caveat that SE students who took and failed the exit exam in the absence of the policy still could have been eligible to graduate. We expand in Section 6.1 however, that this caveat is likely not large enough to fully account for declines in high school completion as a result of the policy.

[8] We compare changes in educational attainment for GE students in districts with lower versus higher pre-policy SE rates across cohorts, who had varying levels of time exposed to the policy. Here, the sample is students who are in GE

educational attainment among GE students due to the policy pressure to reduce SE enrollment. Our DiD estimates imply that GE students enrolled in the average district experienced a 1 percentage point (or 1.6 percent) decline in the likelihood of enrolling in college. We do not find that policy exposure led to significant declines in high school completion among GE students, although the point estimate is negative. We do not expect these effects on GE students are a result of increases in classroom size, since SE removal was driven by those who were already educated in GE classrooms for the majority of the school day. Instead, we interpret these results as suggestive evidence that the additional resources SE programs bring to GE classrooms may have positive spillover effects on GE students.[9] In addition, greater access to SE programs is likely to directly benefit all students by ensuring that SE services are appropriately allocated throughout schooling.

Credibly estimating the long-run impacts of SE programs is difficult due to data limitations and the empirical challenges previously noted. The few studies that have examined SE placement have largely focused on short-run outcomes and mostly find positive effects. Various identification strategies have been used in an attempt to account for the endogenous placement of students into SE. For example, using within student changes, Hanushek, Kain, and Rivkin (2002) find that SE participation improves math performance for students with mild learning and behavioral conditions. Using strategic placement in SE due to an accountability policy that placed pressure on schools to improve overall student performance, Cohen (2007) finds that SE participation reduces absenteeism for marginal low-achieving students.[10] Only one paper, Prenovitz (2017), finds that SE participation harms student achievement. However, this deviation from prior studies is likely driven by the context that she focuses on, which resulted in strategic SE placements for students most unlikely to benefit from SE.[11] An important caveat for these studies is that estimating the effects of SE participation

as of 5th grade prior to policy implementation.

[9]SE resources could directly or indirectly benefit GE students. It is likely that additional resources in the GE classroom, such as teacher's aides, small group work, and extra time on tests improves SE students' focus and attention, thereby reducing behavioral challenges that could have negative spillover effects on GE students. However, these resources may also directly benefit GE students if for example, teacher's aides are are able to help both GE and SE students in the classroom.

[10]Cohen (2007) also finds suggestive evidence that SE placement reduces the probability of dropping out and improves GPA, but these results are not significant at conventional levels.

[11]Specifically, Prenovitz (2017) infers SE program effects based on the introduction of No Child Left Behind (NCLB),

on standardized exam performance could produce mechanically negative results, since SE students often take modified or accommodated versions of the exams. Moreover, these previous studies offer little insight into the role of SE participation on adult outcomes. To date, the only evidence on the long-run impacts of SE has been descriptive and focused on small samples (Newman et al., 2011).

We contribute to the literature in several important ways. First, we utilize linked administrative data to offer, to our knowledge, the first long-run causal impacts of SE participation for marginal students. Second, our focus on a unique policy change that led to the largest exogenous reduction in SE participation to-date allows us to isolate changes in SE access without having to make strong identification assumptions. Third, since we use population data from Texas, a large and diverse state, we are able to estimate differential responses to SE access across many subgroups. We find that less advantaged students and those in lower-resourced and lower-performing districts are more negatively impacted by reduced SE access, suggesting that less access to SE programs may serve to expand pre-existing gaps in later life outcomes among these groups.

More broadly, our results speak to central questions of how to raise human capital for vulnerable groups. First, we add to the literature that investigates the best way to allocate school resources. In particular, are targeted resources (such as SE) or broader improvements (that affect all students) more effective at improving long-run trajectories for at-risk groups? The closest related work by Setren (in press) finds that students with mild disabilities experience achievement gains when they transition to Boston charter schools from traditional public schools. This reduces individualized instructional support (by removing students from SE), but offers higher quality instruction than Boston public schools. However, whether effective charter schools can be replicated is unclear. Our results suggest large returns to investing in specialized educational support when overall improvements in school quality are not possible. A rough comparison suggests that targeting additional educational resources to students with less severe disabilities offer returns that are significantly larger than reducing classroom sizes or increasing school spending, but similar to

which held schools accountable for SE subgroup performance. In this setting, schools faced incentives to assign SE to higher-achieving students and remove SE services for lower-achieving students.

early childhood programs such as Head Start or Perry-Preschool, which are commonly viewed as highly effective interventions (Levin, Belfield, Muennig, & Rouse, 2007; Dynarski, Hyman, & Schanzenbach, 2013; Jackson, Johnson, & Persico, 2015).[12]

Second, we provide new evidence on the timing of human capital investments. While a large amount of evidence points to early childhood (i.e. before age 5) as the critical period to invest resources in vulnerable youth (Garces, Thomas, & Currie, 2002; Deming, 2009; Schweinhart et al., 2005), significantly less is known about the efficacy of interventions later during childhood. Our findings suggest that investing additional resources for vulnerable groups later during childhood can offer similar returns to those of early childhood investments.

2 Background

2.1 Special Education Programs

The Individuals with Disabilities Education Act (IDEA) requires public schools to provide all students with a "free and appropriate" education. Under IDEA, students with disabilities receive SE services to facilitate success in school and later in life. In Texas, as well as in other states, SE program eligibility depends on having a qualifying disability that adversely affects learning, as determined by teachers and specialists. The SE process begins when a parent, teacher, or school administrator requests that a student be evaluated for SE services. Once referred, a psychologist or special education teacher evaluates whether a student qualifies for SE services. SE students are re-evaluated once every three years (or sooner if a parent or teacher requests it). Typically students are first referred to SE during elementary school and continue to qualify for SE throughout their entire schooling. However, some students transition out of SE if they no longer require additional educational support to succeed in school.[13]

[12]For this cost/benefit analysis, we use the social cost of a high school drop-out suggested by (Levin et al., 2007). See Section 6 for more detail on the methodology used to compare the cost/benefit across programs.

[13]In our sample, over 70 percent of SE students who are placed in SE during elementary school (as of 5th grade) continue to participate in SE into high school (as of 9th grade).

Participating students receive *individualized* services and accommodations aimed at ameliorating the challenges they are likely to face throughout schooling and later in life. Because of this individualization, what SE offers is wide-ranging. Students may receive instruction in GE classrooms accompanied by a classroom aide, in resource rooms for part of the school day, or in separate classrooms or schools entirely.[14] Additionally, SE students may be eligible for extra time on standardized exams or take modified exams, which test content below grade level. In terms of grade promotion and high school graduation, SE students may be held to different passing standards or be exempt from test-taking to meet these requirements. Another important component of SE is the close tracking of goals in annual meetings with parents and teachers. Initially, yearly academic or behavioral goals are developed and tracked, and as students approach high school graduation the focus turns towards adulthood goals of either college enrollment or employment.[15]

SE participation has grown significantly nationwide since 1975 (from 8 to 13 percent). These increases in SE participation have been driven by large increases in learning disabilities, speech impairments, other health impairments (including ADHD), and emotional disturbance. Altogether, these conditions, hereafter referred to as "malleable disabilities", represent over 90 percent of total SE enrollment in Texas. Unlike conditions that are physical or more cognitively severe, SE eligibility for these conditions often involve discretion on the part of diagnosticians, teachers, and parents. First, because the most common symptoms for these disabilities are poor academic performance and classroom behaviors, which many students exhibit occasionally, there are inconsistencies in SE referrals (Kauffman, Hallahan, & Pullen, 2017). Moreover, even after being referred, determining whether these conditions adversely affect learning without additional support (the main SE inclusion criteria) can be subjective, as can determining whether students should remain in SE over time (American Psychiatric Association, 2013). This subjectivity underscores

[14]Nationwide, about 63.4% of students with disabilities spend 80% or more of their day in the GE classroom (National Center for Education Statistics, 2019). In our sample in Texas, 88% of SE students spend greater than 50% of their day in the GE classroom and 24% are educated exclusively in GE classrooms.

[15]This preparation for adulthood is called transition planning. Students who aim to enroll in college typically receive guidance on which colleges they should apply to and which courses would best prepare them for college. Those focused on employment typically receive guidance on apprenticeships or other career/technical courses that may be beneficial once they enter the labor market. A specific example of a transition plan is included in Appendix Figure A.1.

the empirical challenges involved in estimating the causal impact of SE participation.

2.2 Policy Background

Prior to the implementation of the SE enrollment target in Texas, concerns over rising SE costs in the U.S. more broadly began to receive attention (Parrish & Mahitivanichcha, 2005). In Texas, in particular, concerns grew stronger with an unexpected state budget cut in 2003, which was set to impact the 2004-05 biennial budget (Hill, 2004). Soon after, in 2004, the Texas state legislature met to consider possible SE cost containment measures. One such method that was considered was a cap on SE enrollment, that legislatures argued "could be used to control SE costs" (Texas House of Representatives, 2004). In the following academic year (2004-05), the Texas Education Agency set a district-level SE enrollment target of 8.5 percent as one part of the new Performance-Based Monitoring Analysis System (PBMAS), used to monitor SE programs annually.

Any district serving more than 8.5 percent of their students in SE was considered out of compliance, and faced state interventions (which ranged in severity based on a district's distance above the target) if they did not reduce SE enrollment to meet this new target.[16] Districts closer to the target were subject to developing monitoring improvement plans, while those further away were subject to third party on-site monitoring visits (Texas Education Agency, 2016). Despite minimal sanctions for districts closer to the target, all districts responded strongly.[17] The first PBMAS report was received in December of the 2004-05 academic year, and was met by a sharp decline in SE enrollment.[18] Figure 1 demonstrates that while the fraction of students enrolled in SE programs was constant during the five years prior to the SE enrollment target (2000-2005), there was a sharp decline during the five years afterwards (2005-2010). The average district experienced a 4.5 percentage point drop, with the largest reductions for districts furthest from the target.

[16] Appendix Figure A.2 shows the performance level rating that each district was assigned based on their SE rate.

[17] Based on a series of interviews featured in a Houston Chronicle investigation of this policy, school administrators report taking this target seriously. For instance, one special education director noted, "We live and die by compliance. You can ask any special ed director; they'll say the same thing: We do what the Texas Education Agency (TEA) tells us" (Rosenthal, 2016).

[18] Because the first PBMAS report wasn't received until the middle of the 2004-05 school year (i.e. December 2004), in what follows, we consider the 2005-06 academic year as the first post-policy year. This was the first academic year where districts would have responded to the policy pressure to reduce SE enrollment.

Given the swift implementation of the SE enrollment target, it was likely unanticipated by districts. There is strong anecdotal evidence in support of this (Rosenthal, 2016), and Figure 1 shows little indication of pre-trends in SE enrollment in the period leading up to the policy's introduction. According to interviews conducted by an investigative Houston Chronicle article in 2016, very few parents knew of the SE enrollment policy's existence and few teachers and administrators understood it (as some thought the policy was implemented nationally and backed by scientific research) (Rosenthal, 2016). Shortly after the greater public was first alerted to the existence of the policy via these articles, a federal investigation of the policy began. In 2017, the TEA ended the enrollment target, and in 2018, the US Department of Education determined the SE enrollment target violated disability law. While the SE enrollment target in Texas has since been reversed, many continue to be concerned about access to SE programs in Texas (Webb, 2019; Hawkins, 2019).

Finally, it is important to note that in addition to the SE enrollment target, the PBMAS introduced other thresholds to improve SE students' academic and behavioral outcomes, reduce the amount of services and accommodations being provided to SE students, and limit the disproportionate representation of minority students in SE.[19] However, beyond introducing strong downward pressure on SE enrollment, the PBMAS did not introduce significant policy pressure on districts to make other instructional changes for SE students. At the time the policy was introduced, roughly 99 percent of districts met or nearly met policy thresholds relating to behavioral and academic outcomes. In addition, 80 percent met or nearly met thresholds relating to the services and accommodations offered to SE students. In contrast, only 9 percent of districts met the thresholds relating to SE enrollment.[20] Throughout the paper we demonstrate that the pressure to make other instructional changes for SE students under the PBMAS is unlikely to be driving our results. We address the additional policy levers that monitored disproportionality in a companion paper (Ballis & Heath, 2020).[21]

[19]Disproportionality is defined as the relative proportion of minority students in SE compared to a district overall.

[20]Panel A of Appendix Table A.1 provides detail on the fraction of districts that met, nearly met, did not meet, or were far out of compliance in each of the SE monitoring areas.

[21]In Appendix Table A.2 we show that our results are not sensitive to the inclusion of controls for black and Hispanic disproportionality. And in Ballis and Heath (2020) we show that there is little correlation between the disproportionality

3 Data and Summary Statistics

3.1 Data Sources

We leverage restricted-access administrative data from the Texas Schools Project (TSP). These data follow the universe of Texas public school students into adulthood, tracking key education and labor market outcomes. Specifically, we start with student-level records from the Texas Education Agency (TEA). These data contain records for all Texas public school students in grades kindergarten through 12, including yearly information on demographics, academic, and behavioral outcomes.[22] Importantly, these data include information on annual SE program participation, as well as disability type, whether students took the unmodified version of standardized exams, and the amount of time spent in resource rooms (i.e. receiving instruction in separate classrooms).[23] We link these student-level school records from the TEA to post-secondary enrollment data from the Texas Higher Education Coordinating Board (THECB). The THECB data include enrollment and degree attainment information for all Texas universities.[24]

These administrative data are advantageous both in terms of the number of long-run outcomes and the large sample size. One drawback of using administrative data from a single state is that we cannot track people who leave Texas. However, outmigration from Texas is quite low. Most people born in Texas remain in the state (Aisch, Gebeloff, & Quealy, 2014) and only 1.7 percent of Texas residents leave the state each year (White et al., 2016). In addition, we are able to link a subset of our sample to the National Student Clearinghouse (NCS) data to determine how often students attend college out of state. Only 1.7 percent of SE students enroll in college outside

targets and SE enrollment target within districts.

[22] SE students may take modified versions of standardized exams if deemed necessary. However, our data does not include performance on modified versions of standardized exams. Because the policy significantly reduced SE enrollment, the fraction of students observed in the achievement data will be increasing endogenously over time, since fewer students are enrolled in SE. This underscores why we do not focus on the impact of SE removal on achievement as a primary outcome in this paper.

[23] Specifically, we observe whether students spent all day in GE classrooms (i.e. are mainstreamed), less than 50 percent of the day in separate classrooms, or greater than 50 percent of the day in separate classrooms.

[24] Although college completion and earnings are available in our data, we leave for future work the impacts of losing access to SE on these outcomes. Given that the policy was implemented in 2005, not enough time has passed for these outcomes to have fully realized.

of Texas within two years of their high school graduation.[25]

3.2 Sample Construction

To identify the direct impact of SE programs on student outcomes, we focus on students enrolled in SE prior to the implementation of the target and infer SE participation effects from policy-driven SE removals. In particular, we focus on students enrolled in SE programs as of 5th grade prior to policy implementation. We focus on 5th grade SE cohorts for several reasons. First, they capture a stable sample of SE students: as Appendix Figure A.3 makes clear, SE enrollment typically grows rapidly throughout elementary school and levels off by 5th grade (with very little new enrollment afterwards). Moreover, 5th grade cohorts have many remaining years in school making them more susceptible to the policy change than older cohorts would have been.[26]

Our main analysis sample consists of 5th grade SE cohorts enrolled between 1999-00 and 2004-05. The 2004-05 cohort was the last cohort placed in SE before the SE enrollment target was enforced.[27] Since the policy significantly changed the composition of students identified with disabilities, this restriction ensures that students in our sample were diagnosed under a similar policy environment.[28] Unless otherwise specified, we restrict our earliest cohort to 1999-00 (rather than the 1995-96 cohort when our data begins). We make this restriction to avoid including cohorts affected by the introduction of school finance equalization in Texas, which affected SE classification incentives (Cullen, 2003).[29] Finally, we limit our sample to students in districts that served between

[25]We demonstrate in Section 5.4 that our results are not sensitive to the inclusion of out of state college enrollment. Focusing on the subset of students for whom we observe NSC data and including out of state college enrollment provides nearly identical estimates to our main results, which only include college enrollment within Texas.

[26]However, our results are not sensitive to this grade cohort restriction. In Appendix Table A.3 we demonstrate that the impact of SE removal on 4th or 6th grade cohorts provides similar estimates to 5th grade cohorts.

[27]Our data reports SE participation as of October. Thus the 2004-05 cohort was enrolled in SE as of October 2004, prior to when districts received the first PBMAS report in December 2004.

[28]Appendix Table A.4 illustrates differences in the composition of each disability type in SE before and after policy implementation. The fraction of SE students with more severe disability types was higher in the post-period, underscoring why we restrict our cohorts to those diagnosed before the policy's implementation.

[29]In particular, Cullen (2003) demonstrates that school finance equalization increased fiscal incentives to enroll marginal students in SE in higher-wealth districts. By 1999-00, SE enrollment rates had leveled off. While school finance equalization changed classification incentives, it led to relatively small changes in SE access. As such, in Appendix Table A.3 we demonstrate that our results are largely unchanged if we use the extended number of cohorts (i.e. 1995-96 to 2004-05). Thus, it is sometimes helpful to extend the number of 5th grade cohorts back to 1995-96. For

6.6 and 21.5 percent of their students in SE in 2004-05 to focus on districts with typical rates of SE.[30] The final sample consists of roughly 38,000 SE students from each cohort, for a total of 227,555 students.[31]

To examine a particularly vulnerable subgroup, we use information about one's disability (as of 5th grade) to identify students whose diagnoses may have been easier to manipulate under the policy. We classify students as being more vulnerable to the policy pressure of reducing SE enrollment if they had a malleable disability type, which we define as learning disabilities, speech impairments, other health impairments (includes ADHD), or emotional disturbance, and if they received more than 50 percent of their instruction in GE classrooms at baseline.[32] In what follows, we refer to this subgroup as our "high-impact" sample consisting of 189,042 students.

3.3 Summary Statistics

Table 1 presents summary statistics for 5th grade cohorts enrolled between 1999-00 and 2004-05. Columns 1 vs. 2 compare students not enrolled in SE to those who are as of 5th grade. Students in SE are more likely to receive Free and Reduced-Price Lunch (FRL), are slightly more likely to be enrolled in the English Language Learner (ELL) program, have lower performance on standardized exams (conditional on taking the unmodified version of the tests), and have lower long-run outcomes (that is, less likely to graduate, enroll in college, and have lower early labor market outcomes). These differences help to highlight the fact that raw comparisons between those who receive SE services and those who do not will be biased due to negative selection into SE programs.

Column 2 of Table 1 demonstrates that 91 percent of students in SE by 5th grade have malleable disabilities, the most common of which is learning disability at 60 percent. The majority

instance, in our event-study analysis extending the number of cohorts back to 1995-96 allows us to provide more visual evidence of pre-trends.

[30]This drops roughly 1% of the overall sample since district outliers with respect to SE rates are small. We demonstrate in Appendix Table A.3 that our results are nearly identical if these districts are included.

[31]We note that charter schools are included in this analysis, as they are also public schools, and therefore also subjected to the PBMAS policy changes (Texas Education Agency, 2004). However, our results are robust to dropping charter schools. These are available upon request.

[32]The rationale for this restriction is that if students are receiving most of their instruction outside of GE classrooms then they are likely to have more severe conditions which may make it more difficult to justify SE removal.

of SE students, 87 percent, receive over 50 percent of their instruction in GE classrooms and 29 percent take unmodified exams. As previously noted, SE students may transition out of SE programs if it is no longer appropriate. Columns 3 vs. 4 of Table 1 compare SE students who remain in SE to those removed from SE by 9th grade. 5th grade SE students who lose SE by 9th grade are less likely to receive FRL, less likely to participate in ELL, have higher achievement on standardized exams, and have better long-run outcomes.[33] Nearly all students who lose SE by 9th grade have malleable disabilities (98 percent) and require fewer modifications to the regular curriculum; 97 percent receive over 50 percent of their instruction in GE classrooms and 58 percent take unmodified standardized tests. These differences highlight the positive selection into SE removal. Without exogenous changes in SE participation, comparisons between those who continue in SE vs. those who lose SE will be biased due to positive selection into SE removal.

Appendix Table A.5 presents summary statistics across districts' initial SE enrollment rates. More treated districts have smaller proportions of Hispanic students, slightly more students receiving FRL, and fewer students participating in ELL. The average district size was also smaller for more treated districts and they were more likely to be located in rural areas. While our identification strategy does not require that the distance above the SE enrollment target be uncorrelated with district characteristics, it does require that the distance above the SE enrollment target is uncorrelated with changes in outcomes that occur for any reason other than the introduction of the SE enrollment target. Reassuringly, we demonstrate in Section 5.4 that our results are robust to the inclusion of time trends interacted with district demographics (measured at baseline in 2004-05). In addition, we account for differences in baseline characteristics in our empirical strategy by including controls for demographic variables at the individual and cohort-district level, described further in Section 4.

[33]High school graduation is the one exception to this pattern which can likely be explained by accommodated graduation requirements available only to SE students.

4 Empirical Strategy

4.1 DiD Estimates of the Impact of the SE Enrollment Target on Outcomes

We first estimate the causal impact of the policy pressure to reduce SE enrollment on student outcomes. The SE enrollment target was introduced in all districts at the same time, so it is not possible to use cross-district variation in implementation date. Instead, we use differences in treatment intensity, which varies across students in two ways. First, districts with higher pre-policy rates of SE enrollment faced stronger policy pressure to reduce SE enrollment. Thus, any effect of the policy should be increasing with a district's pre-policy SE enrollment rate.[34] Second, 5th grade cohorts were differentially treated under the policy based on the number of years (after 5th grade) that they were expected to be enrolled in school after the policy's introduction in 2004-05.

Our DiD estimating equation thus takes the form:

$$Y_{icd} = \delta_0 + \delta_1(\text{SERate}_d^{Pre} \times \text{FracExposed}_c) + \lambda_1 X_{icd} + \lambda_2 Z_{dc} + \gamma_d + \phi_c + \varepsilon_{icd} \tag{1}$$

where Y_{icd} is either an indicator for SE removal or a long-run outcome for student i in 5th grade cohort c in district d. We control for 5th grade district fixed effects γ_d and 5th grade cohort fixed effects ϕ_c. The term SERate_d^{Pre} is the percent of SE students above the 8.5 percent target in a student's 5th grade district in the 2004-05 school year (the year prior to policy implementation), and is set to 0 if a district is already below 8.5 percent.[35] This term is interacted with FracExposed_c, a continuous measure of policy exposure, defined as the fraction of years a student spent in school under the policy between 5th and expected 9th grade.[36] We choose the relevant period of policy exposure to end in 9th grade, right before high school drop-out decisions are typically made (Texas

[34] While this district-level treatment is continuous, it may be helpful to think of districts under more policy pressure as forming the "treated" group, whereas, those under less pressure form the "control" group.

[35] We assign treatment intensity based on a student's 5th grade district (which was determined pre-policy). This ensures that our estimates will be free of bias from selection into districts facing less policy pressure.

[36] We use *expected* 9th grade, i.e. 4 years after 5th grade, in order to ensure that students within a cohort are assigned the same amount of policy exposure. This prevents more years of treatment being assigned to grade repeaters. To illustrate the cross-cohort variation we utilize, Appendix Table A.6 shows policy exposure by each 5th grade cohort.

Education Agency, 2018).[37] The vector X_{icd} includes an indicator for gender, race, FRL status, ELL classification, gender-race interactions, primary disability type, unmodified exam indicator, and level of classroom inclusion. These variables are all measured at baseline, in order to absorb differences by student demographics and disability type. Further, to control for changes in district-level demographics, Z_{dc} includes the district percent of students by racial group, FRL, ELL, and gender for the full student population and for the SE student population, all defined at baseline.[38]

The main variable of interest, δ_1, represents the average impact of the policy pressure to reduce SE enrollment on student outcomes. The key identification assumption is that districts under more policy pressure to reduce SE enrollment had similar counterfactual trends relative to districts facing less pressure. We present event-study estimates by replacing FracExposed_c in Equation 1 with 5th grade cohort indicators. This approach allows us to visualize any differences in outcomes in more versus less treated districts before and after the policy's enactment as a test of this assumption.

To further assess the plausibility of this assumption, we examine observed trends in SE participation, high school completion, and college enrollment across districts with high initial SE rates relative to those with low initial SE rates. In Panel A of Appendix Figure A.4, we split districts into four categories based on their SE rate in 2005. The bottom series contains districts already in compliance with the enrollment target in 2005. The top three series split districts into terciles based on their 2005 SE rate, given that it is above 8.5%. This figure illustrates that although the levels of SE were different across districts prior to policy implementation, the trend was fairly parallel over time. After 2005, districts with the highest SE rates in 2005 made the largest reductions to their SE rates post-policy in order to comply with the enrollment target, relative to districts with lower SE rates. This also provides additional intuition for using the distance to the enrollment target as the measure of treatment intensity in our empirical strategy. In Panels B and C of Appendix Figure A.4, we illustrate similar trends in the raw data for long-run outcomes. Each of these figures demonstrate

[37] Our results are not sensitive to accounting for policy exposure through high school (i.e. up to expected 12th grade). Additionally, results from event-study specifications presented in Section 5.1 support the relevance of this margin.

[38] We control for average district characteristics to account for overall changes in district demographics. Controlling for district averages using SE students only accounts for compositional changes in the students in our sample.

patterns that provide support for common trends in the pre-policy period.[39]

To additionally check whether parallel trends were likely to continue in the absence of the policy, we show that there were no trends in demographics by initial district SE rates in Columns 1, 2, and 3 of Appendix Table A.7. In addition, we use each of our covariates to generate predicted outcomes based on students' characteristics during the pre-policy period. Columns 4, 5, and 6 of Appendix Table A.7 show that conditional on 5th grade cohort and district fixed effects, there is generally little association between treatment and these predicted outcomes. In some instances the estimated effects are positive and significant, but they are not economically meaningful.[40] Moreover, the positive direction of these effects suggest, if anything, that students in more treated districts were becoming positively selected over time, which would lead us to underestimate the negative impact we find on long-run outcomes.

We also investigate whether the policy led to differential attrition, perhaps with students in high initial SE rate districts being more likely to drop out *before* expected 9th grade after the policy was implemented. As this type of attrition is more likely to occur for students with more family resources (who would be expected to have better long-run outcomes), this could have changed the underlying composition of students in a way such that parallel trends would not have been likely to continue. Reassuringly, we show that attrition from our sample is uncorrelated with initial SE rates in Column 7 of Appendix Table A.7.[41] We also investigate whether there was greater district switching in more treated districts, relative to less treated districts. While district switching would not pose a threat to our identification strategy since we assign treatment based on each student's 5th

[39] The 5th grade cohort year 1998 corresponds to individuals who were in 12th grade in 2005, and thus completely unexposed to the policy. 5th grade cohorts before 1998 were expected to graduate before the policy was implemented, while cohorts after 1998 had increasing numbers of years policy exposure. Beginning with the 2002 5th grade cohort (who was first exposed to the policy in 9th grade), we start to see a divergence in trends, with more treated districts experiencing smaller increases in high school graduation and college enrollment relative to untreated districts.

[40] For the full sample presented in Panel A, there are positive effects (significant at the 5 percent level) on predicted college enrollment and SE removal, but both are small and correspond to a 1 percentage point (or roughly 1 percent) change for both outcomes. For the high-impact sample presented in Panel B, we do not find that there is a significant relationship between predicted outcomes and treatment.

[41] Specifically, to look at attrition our outcome variable is an indicator for whether a student was enrolled in Texas public schools in expected 9th grade (conditional on being enrolled in 5th grade).

grade district, excessive district switching could attenuate our estimates. We find no evidence of district switching, as shown in Column 8 of Appendix Table A.7.

Finally, our identification strategy requires that there were no contemporaneous shocks that differentially impacted districts by initial SE enrollment rates. We assess the plausibility of this assumption in Section 5.4, where we consider other education and economic policies during this period. Overall, we do not find evidence that there are any other contemporaneous policy shocks that could have significantly biased our results.

4.2 IV Estimates of SE Removal on Long-Run Outcomes

Next, we use changes in SE access as an instrument for changes in SE participation. Since our setting focuses on students already enrolled in SE programs, our first stage outcome is SE removal and our instrument is our measure of policy exposure (i.e. $\text{SERate}_d^{Pre} \times \text{FracExposed}_c$). With this approach, we identify the local average treatment effect (LATE) of SE removal on long-run outcomes for students on the margin of SE placement decisions, precisely the group for whom the net benefits of SE are most unclear.

This IV approach hinges on two identifying assumptions. First, the policy must generate variation in SE removal. As we will demonstrate in Section 5.1, the policy significantly increased the likelihood of SE removal. Second, we must assume that the exclusion restriction holds. That is, policy exposure only impacted students through changes in SE removal. Thus, a potential concern is that the policy lead to changes that could affect student outcomes through other channels. For instance, if more treated districts re-allocated district resources or made other instructional changes for SE students, then we would not be able to attribute the reduced form effect on SE student outcomes to SE removal alone.

To rule out other channels, we first consider whether more treated districts changed resources for SE students. Given these districts were reducing the number of students enrolled in SE, they may have shifted resources from SE programs to GE, to the detriment of SE students' outcomes. Alternatively, if districts kept resources constant, students who continued in SE after the policy could have benefited from more resources per SE pupil. As shown in Appendix Table A.8, we find no significant impact of the enrollment target on district level SE or GE per-pupil spending or on student-teacher ratios during the five years after policy introduction, suggesting that changes in school-based resources for SE students are unlikely to be driving our results.

Next we investigate whether SE students in more treated districts experienced other instructional changes. This is important to consider since at the time the SE enrollment target was introduced, other aspects of SE instruction also began to be monitored under the PBMAS. As previously argued in Section 2.2, due to the minimal policy pressure that the PBMAS placed on districts (except for the strong pressure to reduce SE access), we believe it is unlikely students would have experienced other instructional changes. Nonetheless, it is important to rule out this possibility empirically. Appendix Table A.9 provides estimates of δ_1 from Equation 1 for time spent in resource rooms and unmodifed test-taking as of expected 9th grade.[42] We find little evidence that the policy pressure to reduce SE enrollment introduced other instructional changes in services or accommodations beyond SE removal. Importantly, for students who continued to be enrolled in SE during 9th grade (Panel C) there is no impact of the policy on the likelihood of spending the majority of the day in a GE classroom or taking unmodified exams. For the overall and high impact samples, we find that the policy led to significant increases in the likelihood of taking the unmodified exams. However, it is important to note that students no longer enrolled in SE will have to take unmodified exams, making it plausible that this positive effect is driven by SE removal as opposed to changes in how test-taking decisions for SE students are made. The magnitude of the coefficient for taking the unmodified math exam is nearly identical to the magnitude of the coefficient on SE removal (both corresponding to roughly a 3.5 percentage point increase), providing suggestive evidence in

[42]These outcomes were chosen based on the specific indicators monitored under the PBMAS. The only indicator that we cannot directly test is whether districts were making efforts to improve the academic achievement of SE students. Since we only observe scores for the unmodified version of the standardized exam, it is hard to address whether the academic performance of SE students was improving. However, as illustrated in Appendix Table A.1, 92% of all school districts were already meeting the academic standards outlined prior to policy implementation, suggesting very minimal policy pressure along this dimension. Furthermore, any pressure to improve academic outcomes would underestimate the negative effect of SE removal on long-run outcomes that we find.

support of this conjecture.[43] In Section 5.4 we provide further evidence that it is unlikely that the other aspects of the PBMAS are influencing our results.

These checks suggest that the only district-level response to the SE enrollment target was to remove students from SE programs. However, in addition to the removal of specific services (e.g. one-on-one aide or the option to take a modified exam), SE students could have been affected by other changes (such as class-size effects or spillover effects) driven by the large district-level reductions in SE participation occurring because of the policy. While we do not view changes in class size as likely,[44] and document small negative spillover effects on GE peers (see Section 5.5 for more detail), it is important to acknowledge that we may be confounding SE participation effects with these other possible changes. While we do not view this as a violation of the exclusion restriction, it could limit the generalizability of our results if losing SE when many other students in a district are also losing SE is different than when one student loses access.

5 Results

5.1 Difference-in Differences Results

We begin by establishing that the policy pressure to reduce SE enrollment increased the likelihood of SE removal for our sample. First, we examine the relationship between the 2004-05 district SE rate and the likelihood of SE removal for each 5th grade cohort separately with an event-study analysis. While our main sample includes SE students from 1999-00 through 2004-05 (as justified in Section 3.2), we extend the number of cohorts back to 1995-96 for this event-study to provide additional visual evidence of pre-treatment trends.[45] Panel A of Figure 2 presents event-study

[43]Furthermore, this increase in unmodified test taking would only introduce bias if the type of exams a SE student takes has a direct influence on long-run outcomes, which is a-priori unclear given the flexibility available to SE students regarding high school graduation requirements. For instance, even if SE students take the unmodified exit exams and fail them, high school graduation may still be deemed appropriate.

[44]We show in Section 5.1 that the declines in SE participation are driven by students who already spend the majority of the school day in GE classrooms, likely ruling out the possibility of compositional changes in GE or SE classrooms.

[45]We also present event study results for our main sample from 1999-00 to 2004-05 in Appendix Figure A.5. These figures demonstrate similar patterns to event-study plots that include an expanded number of 5th grade SE cohorts (i.e. between 1995-96 and 2004-05).

estimates where the outcome is an indicator for SE removal in the year a student was expected to be in 9th grade. The figure demonstrates that cohorts expected to graduate high school before the policy and cohorts with late exposure (after 9th grade) did not experience increases in SE removal. This pattern suggests that pre-trends in SE removal are unlikely to be driving our results. Cohorts exposed to the policy between 5th and 9th grade experienced significant increases in SE removal by expected 9th grade, with the largest increases for cohorts with more years of policy exposure.[46]

The DiD estimates for 5th grade SE cohorts between 1999-00 and 2004-05 are presented in Table 2. We show results for the full sample in Panel A and our high-impact sample (those with mild malleable disabilities) in Panel B. Starting with a model that only includes 5th grade cohort indicators and district fixed effects, we successively add controls. In line with our event-study results, we find that the policy significantly increased the likelihood of SE removal for students in districts with higher pre-policy SE rates. For both samples, our estimated effects are largely stable to choice of specification, especially once we condition on individual disability type.

The results for the full sample suggest that SE students at the average district (that was 4.5 percentage points above the SE enrollment target in 2004-05) who were fully exposed to the policy after 5th grade experienced a 3.7 percentage points (0.0083*4.5) or 13 percent increase in the likelihood of SE removal. We observe larger effects for our high-impact sample, implying that the policy had a larger impact on SE removal for students whose SE placement decisions may have been easier to manipulate. In the high-impact sample, SE students at the average district who were fully exposed to the policy after 5th grade experienced a 4.3 percentage points (0.0096*4.5) or 14 percent increase in the likelihood of SE removal. In addition, the policy had no impact on students whose SE removal would have been more difficult to justify. Appendix Table A.10 presents estimates for students with more severe malleable disabilities (who required separate instruction for more than 50 percent of the day) and those with non-malleable disabilities.[47] For both groups, the estimates are

[46]Appendix Figure A.6 shows an event study that uses an indicator for ever losing SE as the outcome variable. This figure shows a very similar pattern to the one presented in Figure 2. Again, 5th grade cohorts exposed in later grades (i.e. after 9th grade) are not more likely to lose SE despite being enrolled in school after the policy went into effect.

[47]Non-malleable disabilities include autism, deafness, blindness, developmental delay, hearing impairments, intellec-

statistically indistinguishable from zero, implying that these students were not more likely to lose SE under the policy.

5.1.1 Educational Attainment

Next, we estimate whether reduced access to SE due to the policy impacted educational attainment decisions. Again, we start with event-study figures for an extended number of cohorts.[48] Figure 2 presents event-study estimates where the outcome is an indicator for whether a student graduated from high school (Panel B) or enrolled in college within four years of their expected high school graduation (Panel C). Both figures demonstrate similar patterns. Cohorts expected to graduate high school before the policy was implemented or with late exposure did not experience significant declines in educational attainment. These patterns provide compelling evidence that differential trends in educational attainment are not driving our results. Moreover, the impacts of the policy are increasing across cohorts with the number of years that they were exposed to the policy after 5th grade and before 9th grade. These results demonstrate the relevance of our treatment margin, which defines treatment between 5th and expected 9th grade. Despite older cohorts being partially exposed to the policy later in high school, the effects on educational attainment are driven by 5th grade cohorts who were exposed to the policy before they were expected to be in 9th grade.

Our DiD estimates for students enrolled in 5th grade SE cohorts between 1999-00 and 2004-05 are presented in Table 3. This table provides estimates of δ_1 from Equation 1, where the outcome is either an indicator for whether a student graduated from high school (Panels A and B) or whether a student enrolled in college within 4 years of their expected high school graduation (Panels C and D). We show the results separately for the full sample (Panels A and C) and the high-impact sample (Panels B and D). Importantly, these estimates are very stable once individual disability type is controlled for, demonstrating that once we condition on a student's underlying condition exposure to the SE enrollment target is independent of these outcomes. These results illustrate that

tual disabilities, orthopedic impairments and traumatic brain injury.

[48] We also present event study results for our main sample in Appendix Figure A.5. Event-studies that include 5th grade cohorts between 1999-00 and 2004-05 demonstrate similar patterns to event-studies that include an expanded number of 5th grade SE cohorts (i.e. between 1996-97 and 2004-05).

the policy significantly reduced the likelihood of high school completion and college enrollment.

The results for the full sample suggest that at the average district (that was 4.5 percentage points above the SE enrollment target in 2004-05) full exposure to the policy after 5th grade decreased the likelihood of high school graduation by 2 percentage points (or 2.7 percent) and decreased the likelihood of college enrollment by 1.2 percentage points (or 3.6 percent). Moreover, the effects are stronger for students in our high-impact sample, who experienced a 2.2 percentage point (3.2 percent) decrease in the likelihood of high school graduation and a 1.6 percentage points (4.6 percent) decrease in the likelihood of college enrollment. The results for those with physical or more cognitively severe disabilities are presented in Appendix Table A.10. These groups, who were less likely to be impacted, did not experience reductions in educational attainment due to the policy. Thus, the negative impacts on educational attainment are driven by the students who were most likely to lose SE services.[49] This is reassuring for our IV approach, which assumes the reduced form effects are solely driven by SE removal.

5.2 IV Estimates

Having demonstrated that the SE enrollment target significantly increased the likelihood of SE removal, we apply an IV approach to identify the causal impact of SE *removal* on long-run outcomes. The results of this IV analysis are presented in Table 4. We provide estimates for 5th grade SE cohorts between 1999-00 and 2004-05 in our high-impact sample.[50] For reference, Columns 1 and 2 of Table 4 show the first stage effect (i.e. the impact of the policy on SE removal by 9th grade) and the reduced form effect (i.e. the impact of the policy on educational attainment outcomes), respectively. We produce OLS estimates of SE removal on educational outcomes in Column 3.

[49] Appendix Table A.11 shows results for each disability separately. The effects are largely driven by students with learning disabilities (LD). Looking at the other disability types separately, we do not estimate effects that are statistically significant, except for speech impairments. Although not statistically significant, for the other disabilities that we classify as malleable (other health impairments and emotional disturbance) we document similar patterns. We attribute the loss in significance to the much smaller sample sizes for students with these disabilities. Additionally, Appendix Table A.12 shows results for each classroom setting separately. The increases in SE removal and declines in educational attainment are largest for students who were in GE classrooms for the majority of the day, at baseline.

[50] We focus on the high-impact sample, since these students experience the largest declines in SE removal. Our results are largely unchanged if we estimate this for the full sample.

Using OLS models, we find that SE removal is associated with small decreases in high school completion and small increases in college enrollment.[51] However, OLS estimates will be biased upwards since students who typically exit SE do so because they experience improvements in their learning or behavioral outcomes. Our IV estimates presented in Column 4 illustrate the extent to which OLS estimates of the impact of SE removal are biased upwards. Students in our high-impact sample on the margin of SE placement were 51.9 percentage points less likely to graduate high school and 37.9 percentage points less likely to enroll in college, as a consequence of SE removal.[52]

While these are large effects, given that SE removal is accompanied with a significant change in a student's instructional supports (e.g. teacher's aides, ability to work in smaller groups, additional time on tests or assignments, ability to type rather than write) and high school graduation requirements (even for marginal students), we believe these estimates are of plausible magnitude. We consider the plausibility of these magnitudes in greater detail in Section 6.

5.3 Heterogeneous Impacts

We next explore whether there are differential impacts of the policy by race and family income. Ideally, we would first determine how the underlying severity of the conditions of marginal students compare across subgroups. On the one hand, if the underlying severity across subgroups were similar, we would be able to attribute differential responses to SE removal to differences in how subgroups respond to changes in SE access.[53] On the other hand, if the underlying severity across subgroups differed, then differential responses to SE removal could be driven by the conditions of marginal participants. Unfortunately, definitively establishing how the marginal SE student compares across student demographics is difficult with most available data (including our own).

[51] While we might typically expect to find that exiting SE is associated with increases in high school completion, the negative correlation can likely be explained by differences in high school graduation requirements. Despite the fact that students removed from SE programs are positively selected, it is more difficult to graduate outside of SE programs due to increased high school graduation requirements such as the high school exit exam.

[52] At the bottom of Table 4 we report the Kleibergen-Paap F-statistic to test whether our instrument is weak. The Kleibergen-Paap F-statistic of 17.02 is above typical critical values used to test for weak instruments.

[53] Even if marginal students across subgroups had similar underlying severity of conditions, differential responses to SE removal could emerge if more advantaged youth attended higher-resourced schools or had parents that were better able to offset the negative consequences of SE removal by paying for services outside of school.

Two recent papers that have been able to account for a large number of student characteristics, namely health endowments or early achievement measures, point to lower SE access for minority students (Elder et al., in press) with fewer differences in SE access by family income (Hibel, Farkas, & Morgan, 2010).[54] Despite having access to fewer covariates than these recent studies, we arrive at a similar conclusion based on models that predict 5th grade SE receipt based on demographics and 3rd grade achievement.[55] Although our predictive models only offer suggestive evidence of how the underlying conditions compare across subgroups, we find that at baseline minority students were likely to have more severe conditions than non-minority students, but there were fewer differences in disability severity across family income. Thus, the differences we document across race may partly reflect the fact that minority students were likely to have more severe conditions at baseline. However, the differences across family income are likely to reflect differences in how low-income students respond to reduced access to SE.

Our heterogeneity results suggest that low-income and minority students are significantly more likely to lose SE as a consequence of the policy. Panel A of Table 5 demonstrates that the likelihood of losing SE is driven by FRL students (Columns 1 vs. 2).[56] On average, students eligible for FRL are 5 percentage points more likely to lose SE after the policy, while non-FRL students are 3 percentage points more likely to lose SE. Moreover, this difference is statistically significant, with a p-value of 0.03. Similarly, minority students are more likely to lose SE than white students. On average, minority students are 5 percentage points more likely to lose SE, while white students are 3 percentage points more likely to lose SE. This difference, however, is not statistically significant. These results are consistent with less advantaged parents being less able to challenge SE removal decisions made by school personnel under pressure to reduce SE access (Koseki, 2017).

[54]Elder et al. (in press) link a rich set of health and economic endowments at birth to later SE participation. Hibel et al. (2010) utilize information on achievement prior to Kindergarten entry to predict SE participation.

[55]Specifically during the pre-policy period, we predict the likelihood of SE participation in 5th grade using 3rd grade characteristics. Before accounting for 3rd grade achievement, minority and FRL students are more likely to be enrolled in SE programs by 5th grade (Column 1, Appendix Table A.13). Once we condition on 3rd grade achievement, however, FRL status displays a relatively weak relationship with the likelihood of SE placement in 5th grade (i.e. only 0.5 percentage points less likely to be enrolled in SE), while being a minority student is a stronger predictor of not being enrolled in SE as of 5th grade (i.e. 4 percentage points less likely).

[56]This sample includes 5th grade SE cohorts between 1999-00 and 2004-05 in our high-impact sample.

We find that the reductions in educational attainment are also driven by low-income and minority students. In Table 5 we show DiD and IV estimates for high school completion in Panel B and for college enrollment in Panel C. IV estimates reveal that marginal FRL students are 49 percentage points less likely to graduate from high school and enroll in college if removed from SE. IV estimates reveal that marginal minority students are 56 percentage points less likely to graduate from high school and 66 percentage points less likely to enroll in college if removed from SE. In contrast, non-FRL and white students do not experience statistically significant declines in educational attainment. There is only one instance where we find an impact of the policy on longer-run outcomes for non-FRL students. DiD estimates reveal that non-FRL students are less likely to complete high school. However, this could be driven by higher income parents moving their children into private school or home schooling after 9th grade.

When interpreting these differences by race, it is important to highlight that districts were separately under pressure to limit SE enrollment for minority students if the rate of minority students in SE exceeded the rate of minority students in the district (referred to as "disproportionality") under the PBMAS. Districts facing both policy pressures would have more incentives to reduce SE enrollment among minority students, which could partly explain the larger impacts of SE removal among these groups.[57] In Ballis and Heath (2020), we show that limiting disproportionality has a separate effect on minority student outcomes compared to the effect of reducing overall access to SE programs. Interestingly, while reducing access to SE programs has a negative effect on later life outcomes, in Ballis and Heath (2020) we find that black students in districts with relatively higher rates of disproportionality experience small gains in long-run outcomes if removed from SE programs. We explore the mechanisms that drive these differences in Ballis and Heath (2020).

[57] Importantly, controlling for the additional pressure to reduce disproportionality of minority groups leaves our overall and minority group estimates unchanged. We present results for all SE students that include controls for disproportionality in Column 8 of Appendix Table A.2. Results for minority students are available upon request.

5.4 Robustness

The key identification assumption in our analysis is that districts under greater policy pressure to reduce SE enrollment had similar counterfactual trends relative to districts facing less pressure. We have presented evidence in support of this assumption in Section 4.1. However, an additional concern is that population differences by initial SE rates could have led to a later divergence in trends. For instance, the estimated effect of the policy could be driven by differential trends in outcomes due to population differences between more and less treated districts as highlighted in Section 3.3. Reassuringly, Appendix Table A.2 demonstrates that our results are robust to the inclusion of time trends interacted with the baseline fraction of Hispanic students, fraction of FRL students, and total cohort size, measured in the 2004-05 school year. This helps to rule out the possibility that differential trends driven by demographic differences are driving our results.

The other key assumption is that there were no contemporaneous shocks that differentially impacted districts in a way that correlates with district SE rates. To address this, we consider the other components of the PBMAS. While the SE enrollment target was the major component of the PBMAS, as previously noted, there were other targets put in place at the same time to reduce the services and accommodations being provided to SE students. In addition to evidence presented in Section 4.2 that the policy pressure to reduce SE enrollment did not led to changes in time spent in resource rooms or modified exam taking, we perform several checks to rule out the possibility that our estimates are confounded by other changes for SE students as a result of the PBMAS.

First, we re-estimate our results dropping districts under pressure to reduce the amount of time spent in separate classrooms or the number of SE students taking modified exams. Columns 2 and 3 of Appendix Table A.2 present these results, which are nearly identical to our main estimates. This suggests that the small number of districts facing these additional pressures are not driving our results. Second, we rule out the possibility that districts facing additional pressures under the PBMAS were on differential trends by including trends interacted with the 2005 rating in each area of the PBMAS monitoring. Columns 4-6 of Appendix Table A.2 demonstrate that our results are robust to the inclusion of such trends. Finally, we re-estimate all of our results on a subset of students who were receiving minimal accommodations at baseline (i.e. those taking unmodified exams and who spent minimal time in separate classrooms) in Appendix Table A.14. Focusing on this sample ensures we are estimating the effect of the policy on students who would have been exclusively affected by the policy pressure to reduce SE enrollment (these were the students who were already receiving the level of services and accommodations that were deemed compliant under the PBMAS). Taken together, these four checks provide compelling evidence that other aspects of the PBMAS SE monitoring are unlikely to be driving our results.

Next, we consider other education policies affecting Texas public school students during this period. To our knowledge, the only other policy change around this time that could have influenced long-run trajectories was the introduction of No Child Left Behind (NCLB) in 2003. Since many features of NCLB mirrored those of the existing accountability system that had been in place in Texas since 1993, we do not expect that NCLB played a large role in Texas. Nonetheless, it did introduce one important change, namely that SE subgroups were held accountable as a separate group under accountability. Prenovitz (2017) demonstrates that in North Carolina, NCLB's implementation led to incentives to alter the set of SE test-takers in order to improve the test performance of SE students. If low-performing SE students are losing SE in order to boost the SE subgroup's performance on standardized exams, we may be over-estimating the negative impact of SE removal for students on the margin. We present results that account for differences in pre-policy math test scores (measured in fourth grade) in Appendix Table A.15.[58] We find that the highest performing students were most likely to lose SE, ruling out this type of strategic placement.

We also consider the potential impact of the great recession, which officially occurred between December 2007 and June 2009. We believe it is unlikely that the great recession is influencing our results. The great recession was relatively mild in Texas. Between 2004 - 2006 and 2007 - 2009, the unemployment rate rose from 5.4 to 5.6 percent in Texas. In contrast, over the

[58] We augment Equation 1 by including a term that interacts 4th grade standardized math test scores with treatment and 4th grade standardized test scores.

same two time periods, the unemployment rate rose from 5.1 to 6.6 percent nationwide (Andrews, Li, & Lovenheim, 2014). Moreover, we do not have any reason to believe that the great recession differently affected students in high vs. low treated districts. As demonstrated in Appendix Table A.16, our results are robust to the fraction of a cohort that was FRL and the share of a district located in rural areas in 2004-05 interacted with time trends. This helps to rule out the possibility that educational attainment in districts that may have responded differently to the great recession, either due to their geography or underlying economic conditions, were on differential trends.[59]

As a final robustness check, we re-estimate all of our college enrollment estimates for the subgroup of students for whom we have National Student Clearinghouse (NSC) data. These additional data allow us to address whether the lack of out of state college enrollment is biasing our results. For 5th grade cohorts from 2000-01 through 2004-05, we are able to follow out of state college enrollment up to two years after expected high school graduation. Panel A of Appendix Table A.18 presents results were we omit out of state college enrollment, and Panel B of Appendix Table A.18 includes out of state college enrollment within two years of expected high school graduation. We find minimal differences across panels. In our fully specified model in Column 5, we find nearly identical effects regardless of whether out of state enrollments are included.[60]

5.5 Impacts on General Education Students

Finally, we examine the impacts of the policy on GE students. There are several reasons why GE students may have been affected by the large reductions in SE access within school districts. First, GE students could have been affected by peer-to-peer influences: policy induced SE removal could have led to more disruptive classroom behavior which could have negatively impacted GE peers. Second, any declines in learning or behavior among SE students may have changed the way teachers allocated resources within the classroom. For instance, if teachers tried to compensate for the loss in

[59]Moreover, we do not find a statistically significant impact of the policy on low-performing GE students on the reading exam and a small marginally significant impact on low-performing GE students on the math exam (see Appendix Table A.17). If the great recession was driving our results, we would expect it to have a similar impact on low-achieving GE students.

[60]Finally, we note that our results are not sensitive to the lag used to define high school completion or college enrollment.

services among special needs students, this could have taken their attention away from others in the classroom. Third, GE students may have benefited from the additional resources SE students bring to GE classrooms, such as classroom aides or co-teachers, and may have been harmed if they were removed. Finally, GE students may have been directly affected by the policy if they were also on the margin of SE participation. As previously noted, students with mild disabilities often transition in and out of SE programs. Therefore, it is possible that some GE students who would have been deemed eligible for services in later grades are now not as a result of the policy.

Table 6 shows the impact of the policy for all students enrolled in 5th grade together, and then separated by their SE participation status as of 5th grade.[61] Panel A demonstrates that across all samples there were declines in SE participation as of 9th grade. Although the declines in SE access were significantly smaller for GE students (0.7 percentage points or 25 percent) relative to SE students in the high-impact sample (4 percentage points or 14 percent), the fact that GE students experienced statistically significant declines in SE access during high school suggests that the impacts on GE students will in part be driven by the direct effects of this policy.[62] These results also demonstrate that the reductions in educational attainment documented in the combined sample of SE and GE students are largely driven by SE students, with smaller and less significant declines in educational attainment for GE students. GE students did not experience statistically significant declines in high school completion, although the point estimates are negative, suggesting a potential decline in high school completion. At the average district, GE students experienced a 1 percentage point (or 1 percent) decline in college enrollment, relative to a 2 percentage point (or 5 percent) decline among SE students in the high-impact sample.[63] To better understand which GE students are driving the declines in college enrollment, we estimate effects by baseline achievement (measured

[61] All specifications include 5th grade cohort and district fixed effects, as well as the full set of individual and cohort demographic controls detailed in Section 4. The combined SE and GE sample (Column 1) additionally controls for SE participation as of 5th grade and interacts all covariates with SE status as of 5th grade. For the GE sample (Column 2), we control for baseline performance on the math and reading standardized exams.

[62] Therefore, GE students should not be viewed as a purely untreated group, but rather a *less* treated group.

[63] Appendix Figure A.7 presents event-study estimates for the full population of students (i.e. SE and GE students combined). In support of our identification assumption, there were no pre-treatment trends in outcomes for 5th grade cohorts unexposed to the policy or with late exposure. For 5th grade cohorts exposed after 9th grade, the impacts of the policy are increasing with the number of years of exposure.

in 4th grade) in Appendix Table A.17. The negative impacts on college enrollment are largest for those who were initially in the middle of the reading (Panel A) and math (Panel B) achievement distribution.[64]

We view the declines in college enrollment among GE students as being driven by a combination of spillover and direct effects, rather than an internal-validity threat (i.e. unobserved and confounding factors unique to high SE rate districts that were occurring at the same time that the policy was introduced). If bias was driving our results, we would expect negative effects to occur for all GE students, regardless of their initial achievement level.[65] The pattern of heterogeneity that we document, with negative impacts being driven by low and middle achievers, suggests that a combination of direct and spillover effects is a more likely explanation for our findings. GE students in the lower half of the achievement distribution experienced the largest declines in SE access in later grades, and are also the group most likely to have been more negatively affected by changes in classroom resources driven by the SE removal of their peers.[66] Higher-achieving GE students may not have been harmed if they were in separate classrooms (e.g. honors classes) or if they had the skills and ability to compensate for changes in classroom resources on their own.

6 Discussion

We find that removal from SE programs for marginal students significantly reduces educational attainment. This suggests that the potential drawbacks of SE program participation (e.g. stigma effects or lowered expectations) are outweighed by the benefits. Ideally we would put our estimated effects in context by comparing them to other studies on SE effectiveness. Yet, as previously

[64] We also note that those at the bottom of the math achievement distribution were significantly less likely to graduate from high school.

[65] Additional evidence to support that our results are not driven by an internal-validity threat is that we do not estimate any negative impacts of the policy on SE students with physical or more cognitively severe disabilities, who did not lose access to SE services due to the SE enrollment target.

[66] SE resources could directly or indirectly benefit GE students. It is likely that additional resources in the GE classroom, such as teacher's aides, small group work, and extra time on tests improves SE students' focus and attention, thereby reducing behavioral challenges that could have negative spillover effects on GE students. However, these resources may also directly benefit GE students if for example, teacher's aides are able to help both GE and SE students in the classroom.

discussed, causal evidence on SE placement is sparse and primarily focuses on short-run outcomes. Instead, we can benchmark our results by comparing our estimates to the long-run impacts of other school-based programs. One caveat of this comparison, however, is that these other school-based programs affect all students, rather than only those with special needs. Other studies have found that reduced kindergarten classroom size increases college enrollment by 2.7 percent (Dynarski et al., 2013) and a 10 percent increase in school spending leads to 0.27 additional years of completed school (Jackson et al., 2015). We estimate that SE removal decreases college enrollment by 37 percentage points. While our effects are significantly larger, we are focused on a program that significantly alters a students' learning environment for a much longer time frame. Also, we are focused on students with disabilities, who are a particularly vulnerable group.

Despite our focus on school-aged youth, it is also relevant to compare our estimates to the long-run impacts of early childhood programs. Similar to SE programs, which target additional resources to students at risk of lower achievement, early education programs such as Head Start also target additional resources to vulnerable groups at younger ages. Head Start, an early childhood education program that provides additional services (i.e. educational, health, and nutrition related) to low-income children has been shown to have long-run positive impacts.[67] Garces et al. (2002) estimate that Head Start participation increases college enrollment by 9.2 percentage points. A rough back of the envelope calculation suggests returns to SE for marginal students that are nearly identical to the returns to early childhood programs such as Head Start.[68]

[67]Other preschool programs such as Abecedarian Project and Perry Preschool have also been shown to have long-run positive impacts. Campbell, Ramey, Pungello, Sparling, and Miller-Johnson (2002) estimate that the Abecedarian Project increased college enrollment by 22 percentage points. Schweinhart et al. (2005) estimate that Perry Preschool increased high school graduation rates by 50 percentage points for females, with no effect on males.

[68]The average additional yearly cost to educate an SE student is $7,016.66 in Texas ($12,573.37 for SE students vs. $4,292.71 GE students). The estimated increase in high school graduation over the four years after 5th grade is 52 percentage points, yielding a per-graduate cost to educate an additional marginal SE student of $52,955.94 (=(100/53) *(7016.66*4)). Using the social cost of a high school drop-out of $256,000 estimated by (Levin et al., 2007), this suggests a benefit cost ratio of 4.8(=(256,000/52,955.94)). Based on a similar calculation, Deming (2009), who identifies the long-run impacts of Head Start participation, estimates a benefit cost ratio of 4(=256,000/65,116).

6.1 The Role of the High School Exit Exam

Given the large estimated impact SE removal has on educational attainment, it is important to consider how much of this effect is driven by changes in graduation requirements (i.e. a "mechanical effect") versus changes in human capital accumulation driven by reduced services as a consequence of SE removal. As previously noted, SE students can be exempt from the high school exit exam, which is a high school graduation requirement for GE students. Thus, SE removal is associated with an increase in graduation requirements for students who were previously exempt from the exit exam.[69] This could be an important factor in explaining the reductions in high school completion.

Indeed, we find that the policy led to significant increases in exit-exam taking. DiD estimates are shown in Panel A (Column 1) of Table 7, where the outcome is an indicator for whether students ever took the math or reading exit exam. Among the full and high-impact samples, we find that the policy led to roughly a 3 percentage point increase in exit-exam taking. IV estimates presented in Appendix Table A.19 show that SE removal leads to very large increases in exit-exam taking. For students on the margin of SE placement decisions, SE removal increases the likelihood of exit exam test-taking by about 70 percentage points. This indicates that students on the margin of SE placement decisions typically do not take the exit exam. These increases in exit-exam taking, in particular for marginal SE students, suggest that being subject to the exit exam could explain some of the declines in educational attainment that we document.

However, we view it as unlikely that increases in exit-exam taking alone are driving our results. First, we do not find that the policy led to meaningful changes in the ability to pass the exit exam. We show DiD estimates in Panel A (Columns 2-3) of Table 7, where the outcomes are indicators for whether a student passed the math or reading exit exam.[70] We do not document significant changes in the likelihood of passing the math or reading exam. Moreover, pass rates

[69]Most SE students are exempt from taking the exit exam, as only 22 percent take it. However, even among SE students who take the high school exit exam, they may not be required to pass it in order to graduate.

[70]The sample for this outcome is restricted to those taking the exit exam, which is changing over time. Therefore, in Columns 4-5 of the same table, we show estimates for being unable to pass the exit exam, which is equal to 1 if a student took the exam and failed it, and 0 if the student did not take the exit exam or they took it and passed it.

on the exit exam are relatively high among SE students: 83 percent pass the reading portion and 53 percent pass the math portion. For students on the margin of SE placement decisions, our IV estimates in Appendix Table A.19 suggest that SE removal decreases the likelihood of passing the math exit exam by 33 percentage points, although this is not precisely estimated. Despite the lack of significance, we interpret this result as providing suggestive evidence against a purely mechanical. The magnitude of this coefficient is smaller than the magnitude of the declines in high school completion (a 52 percentage point decline), which suggests that declines in high school completion cannot fully be explained by exit exam failure.[71]

Next, we present evidence which suggests that the declines in high school completion are driven by students who were unlikely to be affected by policy-driven increases in exit-exam taking. Specifically, we use all of our covariates to predict the likelihood of being exempt from the exit exam for SE students during the pre-policy period. Appendix Table A.20 shows DiD estimates across quartiles of the predicted likelihood of taking the exit exam. From left to right, each column of the table spans SE students who were very likely to have been exempt from the exit exam in the absence of policy-driven SE removals, to those who would have very likely taken the exit exam regardless of the policy change. Panel A of Appendix Table A.20 documents significant increases in SE removal for all students, regardless of their predicted likelihood of taking the exit exam, of roughly 4 percentage points. However, the increases in exit exam test-taking are *exclusively* driven by students who were least likely to take the exit exam pre-policy (Panel B, Column 1), while the decreases in high school completion are *exclusively* driven by those who were most likely to take the exit exam pre-policy (Panel C, Column 4).[72]

Finally, we document policy-driven declines in high school enrollment *before* the first attempt at the exit exam for the sample of regular test-takers at 4th grade. Panel C of Table 7

[71] Additionally we note that for minority students, the impacts of college-going are larger than those on high school completion (Table 5). This further suggests that potentially most, but not all, of the effect is driven by mechanical changes in graduation requirements.

[72] We come to similar conclusions if we focus on regular test-takers at 4th grade. Table 7 show that regular test-takers do not experience differences in the likelihood of taking the exit exam (Panel A Column 1) or passing the math or reading exit exam (Panel A, Columns 2-3) as a result of the policy. However, relative to the overall sample, regular test-takers experienced the largest declines in high school completion (Column 4 of Panel B of Table 7).

presents DiD estimates of the policy on enrollment at each high school grade separately. While we do not find that the policy impacted enrollment before the exit exam was first administered for the full and high-impact samples, we find significant declines in 11th grade enrollment for the regular test-taker sample, right before students would have first attempted the exit exam. This provides evidence against a purely mechanical effect, at least for the regular test-taking sample.[73]

Taken together, we believe these analyses provide evidence that the effects we find are not solely driven by mechanical increases in exit-exam taking. However, we cannot rule out this mechanism entirely, as we do not ultimately observe the reason why a student decided to drop out of high school. Additionally, we cannot rule out the possibility that regular test-takers, for whom we find the strongest evidence against an increase in exit exam taking, did not experience increases in graduation requirements. As previously noted, if they had stayed in SE they still may have been able to graduate despite not passing the exit-exam. However, a rough back of the envelope calculation suggests that changes in graduation requirements for those who took and failed the exit exam cannot fully explain our findings. Among regular test-takers, the estimated decrease in high school graduation of 3.2 percentage points yields 407 fewer graduates in one cohort at the average district.[74] The estimated increase in SE removal of 5.7 percentage points yields 724 fewer students enrolled in SE in one cohort at the average district. In order for the mechanical effect to fully explain our results, at least 56 percent of these 724 students who lost SE due to the policy would have had to fail the exit exam. However, given that on average, 29 percent of regular test-takers who take the exit exam, fail it, we view it as unlikely that exit exam failure is explaining the full increases in high school graduation.

6.2 Other Mechanisms

We also explore other mechanisms to help shed light on whether the declines in educational attainment are driven by reductions in human capital accumulation as a result of losing access to SE

[73]In fact, the regular test-takers experienced the largest policy-driven increases in SE removal (Appendix Table A.14).

[74]This is computed for one fully exposed cohort. Specifically, (0.032*76238/6), where 76,238 is the sample size, which is divided by 6 since that is the number of cohorts we have.

services. We look at absences, discipline, grade repetition, and standardized test performance, again estimating effects separately for the full and high-impact samples. Panel C of Table 7 presents DiD estimates of the impact of the policy on these intermediate outcomes, measured during expected 9th grade. On the whole, we do not find significant declines in absences, discipline, grade repetition, or standardized test performance for any of the three samples.

One interpretation of these results is that the policy did not have any ill effects on human capital accumulation. However, we caution against this interpretation, given the limited outcomes we analyze, and our inability to track changes in achievement for SE students taking modified versions of exams.[75] Moreover, we cannot rule out the possibility that the policy affected the accumulation of non-cognitive skills such as motivation and interpersonal interactions, which could also be an important factor in driving the reductions in educational attainment that we document.

6.3 Mitigating Factors

We next explore school-based factors that could plausibly mediate the impacts of SE removal. For instance, better resourced or higher-performing districts may have been able to mitigate the negative impact of SE removal on student outcomes. First, we explore differences by district wealth. Columns 1 and 2 of Table 8 show that while students in high-wealth districts are more likely to experience SE removal (6 percentage points) relative to students in low-wealth districts (4 percentage points), they are less likely to experience negative long-run consequences associated with this SE removal.[76] Students attending low-wealth districts are 2 percentage points less likely to graduate and enroll in college, while those in high-wealth districts do not experience statistically significant changes in long-run outcomes. These differences suggest that either high-wealth districts are able to help struggling students through better resources in general (e.g. more qualified teachers,

[75] It is important to note that at the time the SE enrollment target was introduced, SE students' behavioral outcomes began to be monitored. As we have argued in Section 5.4, the pressure to reduce discipline among SE students was minimal. Nonetheless, there may have been incentives for districts to reduce discipline among SE students over this period.

[76] We classify high-wealth districts as the top 12% of districts in terms of tax base wealth per-pupil during 2004-05. These are the districts that had to re-distribute their local tax revenues to poorer districts in 2004-05 as part of school finance equalization policy (Cullen, 2003).

better facilities), or are better able to target additional resources to struggling students.

Another way districts may have been able to accommodate students losing SE was through locally funded 504 plans (Samuels, 2018). 504 plans are an alternative way students with disabilities are provided accommodations in school.[77] In fact, after the enrollment target was implemented many districts expanded access to 504 plans in Texas, despite little change in 504 plan enrollment nationally. Appendix Figure A.8 demonstrates that while the fraction of students enrolled in SE rapidly declined after the enrollment target was introduced (2005-2010), there was a corresponding increase in the fraction of students with a 504 plan (2005-2010). Ideally, we would estimate whether the students who transitioned from SE programs to 504 plans were differentially impacted by SE removal. However, we do not have access to student level 504 plan data. Instead, we address the potential role that access to 504 plans had on longer-run outcomes by testing whether there were differential impacts across districts that experienced large growth in 504 plan enrollment, after the SE enrollment target was introduced. In Columns 3 and 4 of Table 8 we find that while both types of districts reduced access to SE, the negative impacts on long-run educational attainment were more negative for students in districts that had lower growth in 504 plan enrollment.

Finally, we explore differences across various measures of average district performance to address whether being in a higher-performing district helps mitigate the negative long-run impact of SE removal. To explore this possibility, we compare across district-level average test scores (Table 8, Columns 5-6) and district value-added (Table 8, Columns 7-8). Regardless of what measure of district performance we rely on, we find that while both high and low performing districts reduced SE enrollment, the negative impacts of SE removal on educational outcomes were concentrated among students in lower-performing districts.

[77] While students with 504 plans receive all of their instruction in GE classrooms, they receive additional accommodations intended to make the GE curriculum more accessible. Typical accommodations include preferential seating, extra time on tests, daily check-ins with teachers, verbal testing, or modified assignments (KidsHealth, 2016).

7 Conclusion

In this paper, we present evidence on how access to SE programs affects long-run educational attainment. Specifically, we focus on how reduced access to SE programs during middle school and early high school ultimately affected high school completion and post-secondary enrollment decisions. Both of our identification strategies are based on the implementation of an SE enrollment target, which required school districts to have no more than 8.5 percent of their students enrolled in SE. This policy change led to an immediate drop in SE enrollment, which varied across districts depending on their initial SE rates.

We find that SE services prepare students with disabilities for long-run success. In the average school district (with initial SE enrollment of 13 percent), fully exposed 5th grade SE cohorts experienced a 3.7 percentage points increase in the likelihood of losing SE four years after 5th grade, a 2.0 percentage points decrease in the likelihood of high school completion, and a 1.2 percentage points decrease in the likelihood of college enrollment. These outcomes are strong predictors of adult success. The magnitude of the estimates is larger among less-advantaged youth and among those attending school in districts with lower wealth and lower average achievement. Our results are robust to a number of specification checks, including student attrition from the sample and differences in trends across the types of districts that would have been closer to or further from compliance with the 8.5 percent threshold prior to implementation.

Having demonstrated that the implementation of the SE enrollment target impacted the likelihood of SE participation, we employ an IV approach that allows us to identify how SE *removal* impacts long-run educational outcomes. We use policy exposure as an instrument for SE removal and find that SE removal decreases the likelihood a student completes high school by 51.9 percentage points and decreases the likelihood of college enrollment by 37.9 percentage points. Although very large, we view these results as plausible given the potential number of and intensity of supports. SE students receive a variety of services, even within the GE classroom, including (but not limited to) teachers aides or paraprofessionals, small group instruction, additional time on assignments and tests, and specific seating assignments. Again, we find that these results are driven by less-advantaged youth. Our results suggest there are large, meaningful, long-run returns to investing in SE services in the public K-12 school setting for students on the margin of placement, especially those from disadvantaged backgrounds.

We also find smaller and less precise declines in educational attainment among GE students. We use a similar DiD model to identify the causal impacts of reduced access to SE on GE students. While we do not estimate statistically significant impacts on high school completion for GE students, we do find significant declines in college enrollment. Our estimates imply that fully exposed GE students enrolled in the average district experienced a 1 percentage point (or 1.6 percent) decline in the likelihood of college enrollment. We view the declines in college enrollment among GE students as being driven by a combination of spillover and direct effects. SE services may indirectly benefit GE students if the resources provided to SE students within the GE classroom help to reduce behavioral challenges that disrupt both GE teachers' ability to teach and GE students' ability to focus. SE resources may directly benefit GE students if, for example, teachers aides in the classroom are also able to help GE students. In addition, GE students may directly benefit from expanded SE access, and be harmed if they are less likely to access SE services in later grades.

While this paper shows robust evidence on the impacts of SE placement on educational attainment decisions, the limited time after the policy does not yet allow us to fully follow students into the labor market. The large wage differential associated with one's decision to enroll in college suggests that reduced college enrollment is likely to have negative effects on later labor market outcomes, once these outcomes are able to fully realize. Understanding the longer-run labor market effects will be the focus of future research.

References

Aisch, G., Gebeloff, R., & Quealy, K. (2014, August 19). Where We Came From and Where We Went, State by State. *New York Times*.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (dsm-5®)*. American Psychiatric Pub.

Andrews, R., Li, J., & Lovenheim, M. (2014). Quantile Treatment Effects of College Quality on Earnings. *The Journal of Human Resources*, *51*(1), 200–238.

Ballis, B., & Heath, K. (2020). Direct and Spillover Effects of Limiting Minority Student Access to Special Education. *Unpublished manuscript*.

Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the abecedarian project. *Applied developmental science*, *6*(1), 42–57.

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.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from head start. *American Economic Journal: Applied Economics*, *1*(3), 111–34.

Dynarski, S., Hyman, J., & Schanzenbach, D. W. (2013). Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion. *Journal of Policy Analysis and Management*, *32*(4), 692–717.

Elder, T. E., Figlio, D. N., Imberman, S. A., & Persico, C. I. (in press). School segregation and racial gaps in special education identification. *Journal of Labor Economics*.

Garces, E., Thomas, D., & Currie, J. (2002). Longer-term effects of head start. *American Economic Review*, *92*(4), 999–1012.

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.

Hawkins, B. (2019). 250,000 Kids. $277 Million in Fines. Itâs Been 3 Years Since Feds Ordered a Special Ed Reboot in Texas â Why Are Students Still Being Denied? *https://www.the74million.org/article/250000-kids-277-million-in-fines-its-been-3-years-since-feds-ordered-a-special-ed-reboot-in-texas-why-are-students-still-being-denied/*. Retrieved from `https://www.the74million.org/article/250000-kids-277-million-in-fines-its-been-3-years-since-feds-ordered-a-special-ed-reboot-in-texas-why-are-students-still-being-denied/`

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

Hill, I. (2004). State responses to budget crises in 2004: Texas. *The Urban Institute*.

Jackson, C. K., Johnson, R. C., & Persico, C. (2015). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *The Quarterly Journal of Economics*, *131*(1), 157–218.

Kauffman, J. M., Hallahan, D. P., & Pullen, P. C. (2017). *Handbook of special education*. Routledge.

KidsHealth. (2016). 504 Education Plans. *https://kidshealth.org/en/parents/504-plans.html*.

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.

Levin, H. M., Belfield, C., Muennig, P., & Rouse, C. (2007). The public returns to public educational investments in african-american males. *Economics of Education Review*, *26*(6), 699–708.

National Center for Education Statistics. (2015). The Condition of Education At A Glance. *http://nces.ed.gov/programs/coe/indicator_cgg.asp*.

National Center for Education Statistics. (2019). Digest of Education Statistics. *https://nces.ed.gov/fastfacts/display.asp?id=59*.

Newman, L., Wagner, M., Knokey, A.-M., Marder, C., Nagle, K., Shaver, D., & Wei, X. (2011). The Post-High School Outcomes of Young Adults with Disabilities up to 8 Years after High School: A Report from the National Longitudinal Transition Study-2 (NLTS2). *National Center for Special Education Research*, *2*, 1–218.

Parrish, T., & Mahitivanichcha, K. (2005). The Implications of Fiscal Incentives on Identification Rates and Placement in Special Education: Formulas for Influencing Best Practice. *Journal of Education Finance*.

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

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

Samuels, A. (2018). Texas Has a Plan to Fix Special Education. Now what? *The Texas Tribune, https://www.texastribune.org/2018/05/03/special-education-texas-fix-disability-funding/*.

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C., & Nores, M. (2005). The high/scope perry preschool study through age 40. *Ypsilanti, MI: High Scope Educational Research Foundation*.

Setren, E. (in press). The Impact of Specialized Services vs. High Quality General Education on Special Education and English Language Learner Students. *Journal of Human Resources*.

Texas Education Agency. (2004). Performance Based Monitoring Analysis System 2004-2005 Manual.

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

Texas Education Agency. (2018). Secondary School Completion and Dropouts in Texas Public Schools 2016-17. Retrieved from `https://tea.texas.gov/acctres/dropcomp_index.html`

Texas House of Representatives. (2004). Interim Report to the 79th Texas Legislature, Committee on Public Education. *https://house.texas.gov/_media/pdf/committees/reports/78interim/public_education.pdf*.

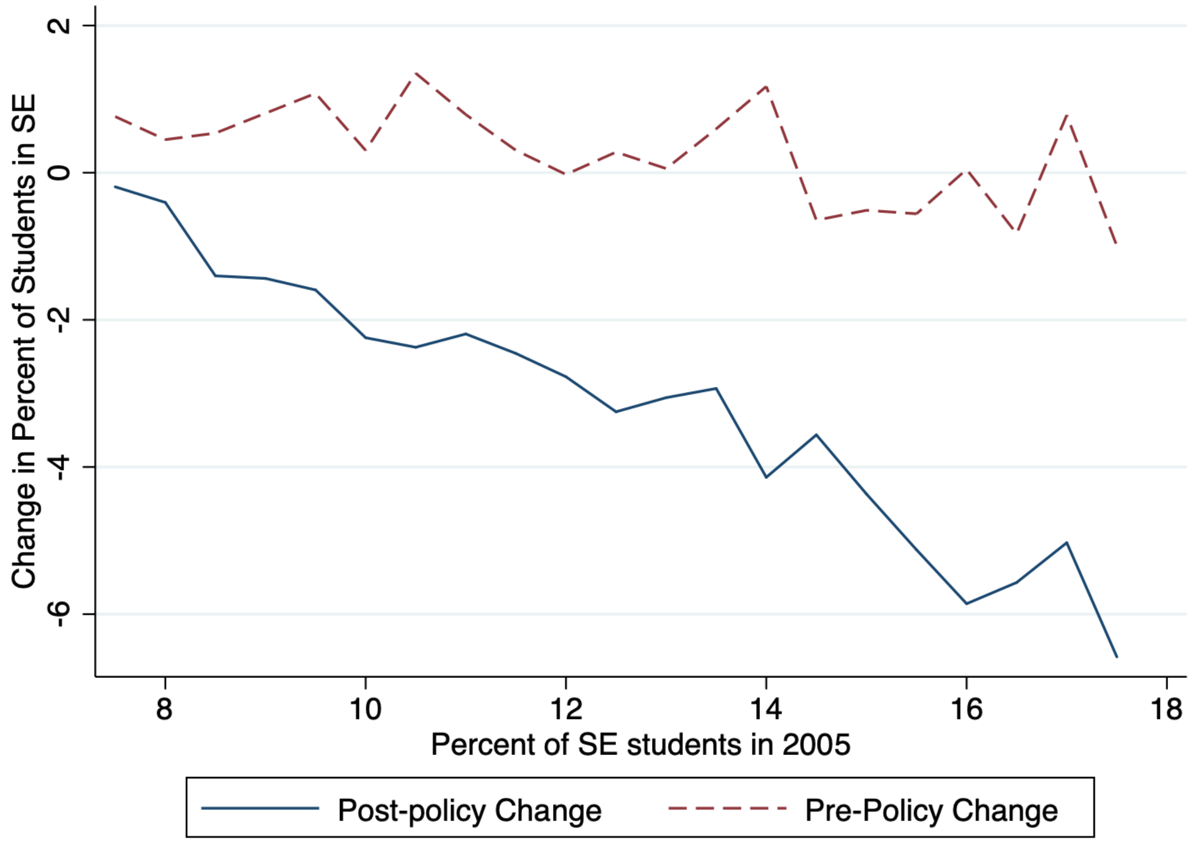
U.S. Department of Education. (2018). U.S. Department of Education Issues Findings in Texas Individuals with Disabilities Education Act Monitoring. *https://www.ed.gov/news/press-releases/us-department-education-issues-findings-texas-individuals-disabilities-education-act-monitoring*.

Webb, S. (2019). Denied Again Students Still Fighting for Special Education. *https://www.houstonchronicle.com/news/houston-texas/houston/article/Students-denied-special-education-failing-schools-14831755.php*.

White, S., Potter, L., You, H., Valencia, L., Jordan, J., & Pecotte, B. (2016). Introduction to Texas Domestic Migration. *Office of the State Demographer*.

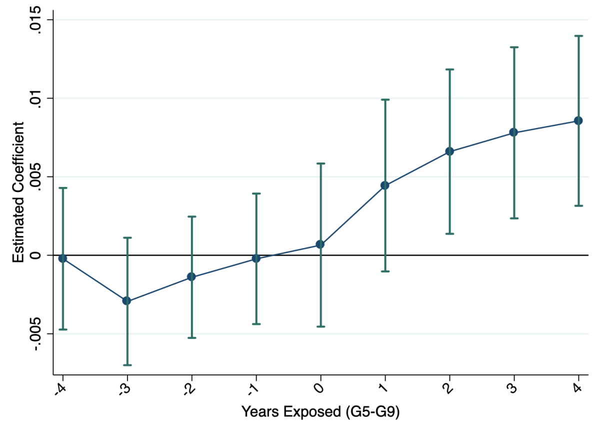
Tables and Figures

Figure 1: Change in District Level Special Education Enrollment During the Pre-Policy Period (2000-2005) and the Post-Policy Period (2005-2010)

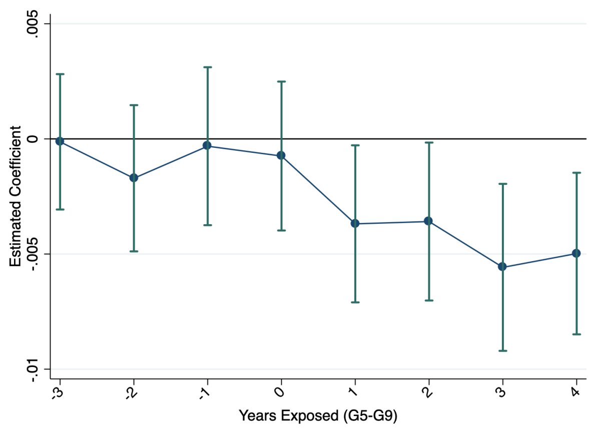


Note: This figure shows the change in Special Education enrollment between the pre-policy period (2000-2005) and the post-policy period (2005-2010) by Special Education enrollment in 2004-05. The district level changes are weighted by 2004-05 district enrollment.

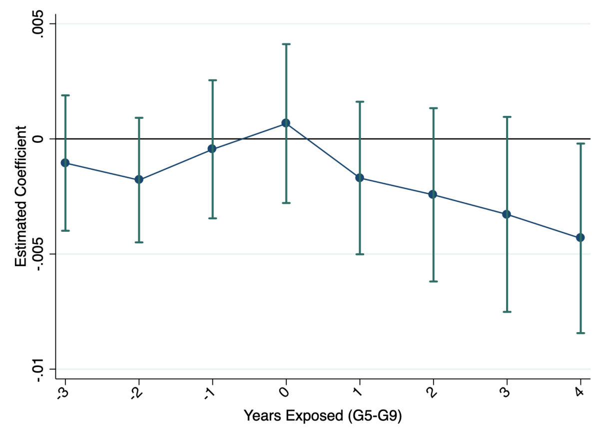
Figure 2: Event Study Estimates of the Impact of the Policy on Special Education Removal and Educational Attainment (High-Impact Sample)



(a) SE Removal



(b) High School Completion



(c) College Enrollment

Note: These figures plot coefficients and 95% confidence intervals from event-study regressions that estimate interactions between 5th grade cohort dummies and the 2004-05 district Special Education (SE) rate. The outcome in figure (a) is SE Removal by expected 9th grade. Figure (b) shows high school completion, and Figure (c) shows college enrollment, measured within four years of expected high school graduation. Event time is computed by subtracting 9 from the grade each 5th grade cohort was expected to be enrolled in during the first year of the policy (or the 2005-06 school year). The sample includes 5th grade cohorts enrolled in SE between 1995-96 to 2004-05 in our high-impact sample, which includes students with a malleable disability (i.e. learning disabilities, speech impairments, other health impairments, or emotional disturbance) who spent more than 50 percent of their instruction in general education classrooms at baseline (measured as of 5th grade). The 5th grade cohort from 1995-96 is omitted, so estimates are relative to that cohort. This regression includes controls for 5th grade cohort indicators, district fixed effects, gender, race, FRL status, ELL classification, gender-race interactions, baseline primary disability, an indicator for whether a student took the unmodified version of the exam, level of classroom inclusion (all measured at baseline in 5th grade). This regression also includes controls for district-level tax base wealth per-pupil and the percent of tax base wealth that is residential, as well as the percentage of students in a district and cohort belonging to each racial group, receiving FRL, classified as ELL, and who are male for the SE sample and the full sample of SE and general education students. Standard errors are clustered at the district level.

Table 1: Summary Statistics - 5th Grade Cohorts Between 2000-2005

	General Education	**Special Education**	**Special Education Removal by G9**	
			No	**Yes**
	(1)	(2)	(3)	(4)
Hispanic	0.430	0.412	0.426	0.374
Black	0.138	0.187	0.199	0.157
White	0.400	0.386	0.363	0.449
FRL	0.528	0.644	0.684	0.537
ELL	0.118	0.159	0.177	0.112
Male	0.486	0.655	0.665	0.629
Std Math Score (G4)	0.0625	-0.537	-0.884	-0.193
Std Reading Score (G4)	0.0492	-0.543	-0.959	-0.217
Taking Reg Test Math (G4)	0.847	0.353	0.243	0.644
Taking Reg Test Reading (G4)	0.847	0.285	0.173	0.582
Long-Run Outcomes				
High School Completion	0.793	0.718	0.724	0.701
Attend College	0.562	0.327	0.270	0.479
College Completion	0.197	0.063	0.034	0.139
Employed	0.701	0.637	0.612	0.706
Annual Earnings ($)	15,678	11,475	10,408	14,294
Disability Type				
Learning Disability	-	0.601	0.647	0.479
Speech Impairment	-	0.135	0.039	0.389
Other Health Impairment	-	0.105	0.119	0.0673
Emotional Disturbance	-	0.066	0.073	0.048
Intellectual Disability	-	0.047	0.064	0.003
Autism	-	0.018	0.024	0.002
Orthopedic Impairment	-	0.011	0.013	0.005
Auditory Impairment	-	0.011	0.013	0.005
Visual Impairment	-	0.005	0.006	0.001
Deafness and Blindness	-	0.0002	0.0002	0
Malleable		0.907	0.878	0.983
Less Malleable		0.093	0.122	0.0166
Classroom Setting				
Mainstream	-	0.236	0.126	0.526
Separate Classroom ($\leq$50%)	-	0.635	0.706	0.448
Separate Classroom (>50%)	-	0.128	0.167	0.026
Total Students	1,448,003	227,555	165,043	62,512

Note: This table presents summary statistics on outcomes and baseline demographics which are measured in 5th grade, except for measures of achievement which are measured in 4th grade. The sample includes 5th grade cohorts enrolled in Texas public schools between 1999-00 to 2004-05. The first column includes students enrolled in General Education as of 5th grade, the second column includes students enrolled in Special Education as of 5th grade, the third column includes students enrolled in Special Education as of 5th grade who were still enrolled in Special Education as of expected 9th grade, and the fourth column includes students enrolled in Special Education as of 5th grade who were no longer enrolled in Special Education as of expected 9th grade. FRL is students receiving free and reduced-price lunch, ELL is students in the English language-learner program. Malleable disabilities include learning disability, emotional disturbance, other health impairments, and speech impairments and Less Malleable is students with all other disability types. College enrollment is measured within four years of expected high school graduation. College Completion, Employment, and Earnings are measured six years after expected high school graduation. Earnings reported are not conditional on being employed. Those not employed are assigned 0 yearly earnings.

Table 2: The Impact of the Policy on Special Education Removal

	(1)	(2)	(3)	(4)	(5)
Panel A: Full Sample					
Treatment	0.0102	0.0096	0.0082	0.0083	0.0083
	(0.0021)	(0.0021)	(0.0021)	(0.0019)	(0.0019)
	[0.046]	[0.043]	[0.037]	[0.037]	[0.037]
Mean (Y)	0.275	0.275	0.275	0.275	0.275
N	227,555	227,555	227,555	227,555	227,555
Panel B: High-Impact Sample					
Treatment	0.0108	0.0100	0.0096	0.0097	0.0096
	(0.0030)	(0.0030)	(0.0023)	(0.0022)	(0.0022)
	[0.0484]	[0.0452]	[0.0433]	[0.0434]	[0.0431]
Mean (Y)	0.317	0.317	0.317	0.317	0.317
N	189,042	189,042	189,042	189,042	189,042
Controls					
Individual		X	X	X	X
Individual Disability			X	X	X
District-Cohort				X	X
District Finance					X

Note: This table shows difference-in-differences (DiD) estimates of the impact of the policy on Special Education (SE) removal at expected 9th grade. Within each panel, each column reports estimates of δ_1 from a separate regression of Equation 1. The outcome variable is an indicator for whether a student lost SE services the year they were expected to be enrolled in 9th grade. The sample includes 5th grade cohorts enrolled in SE between 1999-00 to 2004-05. Panel A includes estimates for the full sample. Panel B includes estimates for our high-impact sample, where “high-impact” is defined as students with a malleable disability (i.e. learning disabilities, speech impairments, other health impairments, or emotional disturbance) who received more than 50 percent of their instruction in GE classrooms at baseline (i.e 5th grade). All regressions control for 5th grade campus and cohort fixed effects. Individual controls include gender, race, FRL, ELL, gender-race interactions, and an indicator for whether a student took the unmodified version of the exam (all measured at baseline in 5th grade). Individual disability controls include baseline primary disability and level of classroom inclusion (measured at baseline in 5th grade). District demographic-cohort controls include the percent of students belonging to each racial group, receiving FRL, classified as ELL, and who are male for the SE sample and the sample of all students in SE and GE measured at baseline. District financial controls include tax base wealth per-pupil and the percent of tax base wealth that is residential. The effect for the fully exposed student at the average district is shown in brackets, and is defined as the coefficient multiplied by 4.5. Standard errors in parentheses are clustered at the district level.

Table 3: The Impact of the Policy on Educational Attainment

	(1)	(2)	(3)	(4)	(5)
High School Completion					
Panel A: Full Sample					
Treatment	-0.0041	-0.0046	-0.0043	-0.0044	-0.0043
	(0.0014)	(0.0014)	(0.0015)	(0.0015)	(0.0015)
	[-0.0184]	[-0.0205]	[-0.0192]	[-0.0198]	[-0.0195]
Mean (Y)	0.718	0.718	0.718	0.718	0.718
N	227,555	227,555	227,555	227,555	227,555
Panel B: High-Impact Sample					
Treatment	-0.0047	-0.0052	-0.0052	-0.0051	-0.0050
	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)
	[-0.0209]	[-0.0232]	[-0.0235]	[-0.0229]	[-0.0224]
Mean (Y)	0.710	0.710	0.710	0.710	0.710
N	189,042	189,042	189,042	189,042	189,042
College Enrollment					
Panel C: Full Sample					
Treatment	-0.0012	-0.0022	-0.0022	-0.0026	-0.0026
	(0.0013)	(0.0013)	(0.0014)	(0.0013)	(0.0013)
	[-0.0053]	[-0.0097]	[-0.0101]	[-0.0118]	[-0.0118]
Mean (Y)	0.327	0.327	0.327	0.327	0.327
N	227,555	227,555	227,555	227,555	227,555
Panel D: High-Impact Sample					
Treatment	-0.0028	-0.0038	-0.0034	-0.0037	-0.0036
	(0.0014)	(0.0014)	(0.0015)	(0.0015)	(0.0015)
	[-0.0126]	[-0.0170]	[-0.0153]	[-0.0164]	[-0.0163]
Mean (Y)	0.354	0.354	0.354	0.354	0.354
N	189,042	189,042	189,042	189,042	189,042
Controls					
Individual		X	X	X	X
Individual Disability			X	X	X
District-Cohort				X	X
District Finance					X

Note: This table shows DiD estimates of the impact of the policy on educational attainment decisions. Within each panel, each column reports estimates of δ_1 from a separate regression of Equation 1. The dependent variable is shown in bold. College enrollment is measured within four years of each student's expected high school graduation. All regressions control for 5th grade campus and cohort fixed effects. See Table 2 for more detail on the set of controls. The effect for the fully exposed student at the average district is shown in brackets, and is defined as the coefficient multiplied by 4.5. Standard errors in parentheses are clustered at the district level.

Table 4: OLS and IV Estimates of the Impact of Special Education Removal on Educational Attainment (High-Impact Sample)

	First Stage	Reduced Form	OLS	IV
	(1)	(2)	(3)	(4)
Dependent Variable:	*Special Education Removal*	*High School Completion*		
Treatment	0.0096	-0.0050		
	(0.0022)	(0.0016)		
	[0.0431]	[-0.0224]		
Mean (Y)	0.317	0.710		
Special Education Removal			-0.0803	-0.519
			(0.0053)	(0.183)
Dependent Variable:	*Special Education Removal*	*College Enrollment*		
Treatment	0.0096	-0.0036		
	(0.0022)	(0.0015)		
	[0.0431]	[-0.0163]		
Mean (Y)	0.317	0.354		
Special Education Removal			0.0714	-0.379
			(0.0036)	(0.180)
Kleibergen-Paap F-Statisitic	19.62			

Note: This table reports DiD estimates of the impact of the policy on Special Education (SE) removal (by expected 9th grade) in Column (1). Column (2) reports DiD estimates on high school completion and college enrollment. The sample includes 5th grade cohorts enrolled in SE between 1999-00 to 2004-05 in our high-impact sample (N=189,042), where “high-impact” is defined as students with a malleable disability (i.e. learning disabilities, speech impairments, other health impairments, or emotional disturbance) who received more than 50 percent of their instruction in General Education classrooms at baseline (i.e 5th grade). OLS and IV estimates of SE removal on educational attainment are reported in Columns 3 and 4. The dependent variable is shown in the panel headings. See Table 2 for the full list of controls used. The effect for the fully exposed student at the average district is shown in brackets, and is defined as the coefficient multiplied by 4.5. Standard errors in parentheses are clustered at the district level.

Table 5: Heterogeneity by Race and Free and Reduced-Price Lunch (FRL) Status (High-Impact Sample)

	FRL (1)	Non-FRL (2)	Minority (3)	White (4)
Panel A: Special Education Removal				
DiD Estimates				
Treatment	0.012	0.006	0.011	0.006
	(0.002)	(0.003)	(0.003)	(0.002)
P-value		0.034		0.083
Mean (Y)	0.265	0.409	0.281	0.367
Panel B: High School Completion				
DiD Estimates				
Treatment	-0.006	-0.005	-0.006	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)
P-value		0.778		0.105
IV Estimates				
SE Removal	-0.487	-0.828	-0.558	-0.32
	(0.181)	(0.510)	(0.231)	(0.303)
Mean (Y)	0.653	0.811	0.678	0.756
Panel C: College Enrollment				
DiD Estimates				
Treatment	-0.006	-0.000	-0.007	0.001
	(0.001)	(0.003)	(0.002)	(0.002)
P-value		0.046		0.000
IV Estimates				
SE Removal	-0.494	-0.078	-0.664	0.214
	(0.167)	(0.438)	(0.256)	(0.287)
Mean (Y)	0.268	0.506	0.313	0.41
N	120,565	68,429	112,462	73,959
KP F-Statistic	26.385	4.423	16.646	7.705

Note: Panel A reports DiD estimates of the impact of the policy on Special Education (SE) removal. Panel B and C report DiD and IV estimates on high school completion and college enrollment, respectively. See Table 2 for more detail on the full list of controls used. The sample includes 5th grade cohorts enrolled in SE between 1999-00 to 2004-05 in our high-impact sample (N=189,042), where "high-impact" is defined as students with a malleable disability (i.e. learning disabilities, speech impairments, other health impairments, or emotional disturbance) who received more than 50 percent of their instruction in General Education classrooms at baseline (i.e 5th grade). The first column includes student participating in the Free and Reduced-Price Lunch (FRL) Program at baseline, the second column includes students who did not participate in the FRL Program at baseline, the third column includes Hispanic and black students, and the fourth column includes white students. The p-value row presents the p-value associated with the test of equality across the two coefficients (either Columns 1 vs. 2 or Columns 3 vs. 4). Standard errors in parentheses are clustered at

Table 6: The Impact of the Policy on Outcomes by Special Education Status as of 5th Grade

	General and Special Education (1)	General Education (2)	Special Education (3)	High-Impact Special Education (4)
Panel A: Special Education Participation				
Treatment	-0.0029	-0.0015	-0.0083	-0.0096
	(0.0004)	(0.0003)	(0.0019)	(0.0022)
	[-0.0129]	[-0.0067]	[-0.0372]	[-0.0431]
Mean (Y)	0.122	0.0274	0.725	0.683
Panel B: High School Completion				
Treatment	-0.0027	-0.0009	-0.0043	-0.0050
	(0.0010)	(0.0009)	(0.0015)	(0.0016)
	[-0.0121]	[-0.0041]	[-0.0195]	[-0.0224]
Mean (Y)	0.783	0.793	0.718	0.710
Panel C: College Enrollment				
Treatment	-0.0020	-0.0020	-0.0026	-0.0036
	(0.0009)	(0.0010)	(0.0013)	(0.0015)
	[-0.0091]	[-0.0090]	[-0.0118]	[-0.0163]
Mean (Y)	0.530	0.562	0.327	0.354
N	1,675,558	1,217,393	227,555	189,042

Note: This table shows DiD estimates of the impact of the policy on Special Education (SE) removal (at expected 9th grade) and educational attainment for students based on whether they were classified as SE or General Education (GE) as of 5th grade. Within each panel, each column reports estimates of δ_1 from a separate regression of Equation 1. The dependent variable is shown in the panel headings. See Table 2 for more detail on full set of controls. The sample includes 5th grade cohorts enrolled in Texas public schools between 1999-00 to 2004-05. The first column includes all students enrolled in Texas public schools in 5th grade, the second column includes students enrolled in GE as of 5th grade, the third column includes students enrolled in SE as of 5th grade, and the fourth column includes students enrolled in our high-impact SE sample, which is defined as students with a malleable disability (i.e.learning disabilities, speech impairments, other health impairments, or emotional disturbance) who received more than 50 percent of their instruction in GE classrooms at baseline. The first column additionally controls for SE status as of 5th grade and interacts all covariates with SE status as of 5th grade. The second column additionally controls for baseline performance on the math and reading standardized exams. The effect for the fully exposed student at the average district is shown in brackets, and is defined as the coefficient multiplied by 4.5. Standard errors in parentheses are clustered at the district level.

Table 7: Mechanisms

Panel A: High School Exit Exams

	Took Exit Exams	Ever Pass		Unable Pass	
		Math	Reading	Math	Reading
	(1)	(2)	(3)	(4)	(5)
Full Sample					
Treatment	0.0061	-0.0036	0.00049	0.0042	0.0015
	(0.0019)	(0.0022)	(0.0016)	(0.0015)	(0.0008)
Mean (Y)	0.345	0.525	0.831	0.191	0.0678
N	227,555	91,456	91,456	227,555	227,555
High-Impact Sample					
Treatment	0.0068	-0.0027	0.0006	0.0047	0.0017
	(0.0021)	(0.0022)	(0.0016)	(0.0016)	(0.0009)
Mean (Y)	0.389	0.528	0.837	0.212	0.0734
N	189,042	85,110	85,110	189,042	189,042
Reg Test-Taker Sample					
Treatment	0.0026	-0.0042	-0.0004	0.0038	0.0010
	(0.0027)	(0.0028)	(0.0015)	(0.0022)	(0.0010)
Mean (Y)	0.658	0.602	0.920	0.278	0.0559
N	54,337	37,919	37,919	54,337	54,337

Panel B: Enrollment and Attainment

	Enrolled G10	Enrolled G11	Enrolled G12	High School Completion	College Enrollment
	(1)	(2)	(3)	(4)	(5)
Full Sample					
Treatment	-0.0003	-0.0011	-0.0017	-0.0043	-0.0026
	(0.0005)	(0.0008)	(0.0011)	(0.0015)	(0.0013)
Mean (Y)	0.944	0.877	0.776	0.718	0.327
N	227,555	227,555	227,555	227,555	227,555
High-Impact Sample					
Treatment	-0.00002	-0.0009	-0.0016	-0.0050	-0.0036
	(0.0005)	(0.0009)	(0.0013)	(0.0016)	(0.0015)
Mean (Y)	0.944	0.874	0.769	0.710	0.354
N	189,042	189,042	189,042	189,042	189,042
Reg Test-Taker Sample					
Treatment	-0.0007	-0.0039	-0.0036	-0.0083	-0.0046
	(0.0009)	(0.0014)	(0.0018)	(0.0022)	(0.0028)
Mean (Y)	0.957	0.907	0.817	0.762	0.518
N	54,337	54,337	54,337	54,337	54,337

Panel C: Intermediate Outcomes (During Expected 9th Grade)

	Share Absent	Repeated Grade	Ever Disciplined	Std Exam Score	
				Math	Reading
	(1)	(2)	(3)	(4)	(5)
Full Sample					
Treatment	-0.00004	0.0003	-0.0017	0.0075	0.0159
	(0.0003)	(0.0005)	(0.0016)	(0.0046)	(0.0053)
Mean (Y)	0.0693	0.0262	0.405	-0.454	-0.554
N	227,363	227,555	227,555	100,339	103,713
High-Impact Sample					
Treatment	-0.0001	-0.0002	-0.0013	0.0061	0.0147
	(0.0003)	(0.0004)	(0.0016)	(0.0045)	(0.0055)
Mean (Y)	0.0685	0.0197	0.419	-0.448	-0.538
N	188,952	189,042	189,042	93,978	96,693
Reg Test-Taker Sample					
Treatment	0.0003	0.0007	-0.0013	0.0052	0.0096
	(0.0004)	(0.0006)	(0.0025)	(0.0051)	(0.0052)
Mean (Y)	0.0538	0.0156	0.329	-0.106	-0.166
N	54,322	54,337	54,337	44,454	45,362

Note: This table shows DiD estimates of the impact of the policy on high school exit-exam taking and passing (Panel A), high school enrollment and educational attainment (Panel B), and intermediate outcomes (Panel C). Within each panel, each column reports estimates of δ_1 from a separate regression of Equation 1. Within each panel, we report estimates for the full sample, high-impact sample, and regular test-taker sample. The regular test-taker sample includes students who took unmodified standardized exams in 4th grade. The outcomes in Panel A are measured during the year a student was expected to be enrolled in 11th grade. The outcomes in Panel C are measured in expected 9th grade. In Panel A columns (2) and (3), Ever Pass is defined as an indicator for whether students passed the math or reading exit exam, conditional on taking it. In columns (4) and (5), Unable to Pass is defined as an indicator for not passing the exit exam, and assigns a value of 1 to students who fail the exam and a value of 0 to students who pass or do not take the exam. In Panel C column (1), Share Absent is the fraction of days a student was absent. Repeated Grade (column 2) is an indicator for repeating 9th grade. Ever Disciplined (column 3) is an indicator for ever being suspended or expelled. See Table 2 for more detail on the sample and full set of controls. Standard errors in parentheses are clustered at the district level.

Table 8: Heterogeneity by School-Based Factors (High-Impact Sample)

	District Wealth		504 Plan Growth		Average Test Scores		District Value Added	
	High	Low	High	Low	High	Low	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Special Education Removal								
Treatment	0.0123	0.0096	0.0164	0.0080	0.0063	0.0125	0.0103	0.0093
	(0.0074)	(0.0023)	(0.0047)	(0.0025)	(0.0027)	(0.0027)	(0.0028)	(0.0033)
	[0.0554]	[0.0433]	[0.0736]	[0.0360]	[0.0284]	[0.0561]	[0.0462]	[0.0417]
Mean (Y)	0.364	0.315	0.284	0.325	0.359	0.293	0.332	0.308
Panel B: Took HS Exit Exams								
Treatment	0.0041	0.0069	0.0149	0.0051	0.0102	0.0042	0.0102	0.0054
	(0.0065)	(0.0022)	(0.0032)	(0.0026)	(0.0024)	(0.0029)	(0.0029)	(0.0028)
	[0.0184]	[0.0309]	[0.0671]	[0.0227]	[0.0460]	[0.0188]	[0.0459]	[0.0241]
Mean (Y)	0.454	0.386	0.363	0.395	0.458	0.348	0.422	0.368
Panel C: High School Completion								
Treatment	-0.0087	-0.0049	-0.00073	-0.0057	-0.0021	-0.0067	-0.0032	-0.0056
	(0.0065)	(0.0017)	(0.0030)	(0.0019)	(0.0021)	(0.0023)	(0.0024)	(0.0021)
	[-0.0392]	[-0.0221]	[-0.00327]	[-0.0257]	[-0.00950]	[-0.0303]	[-0.0142]	[-0.0251]
Mean (Y)	0.799	0.706	0.712	0.710	0.759	0.682	0.746	0.688
Panel D: College Enrollment								
Treatment	-0.0061	-0.0037	-0.0028	-0.0039	0.00002	-0.0066	-0.0020	-0.0049
	(0.0044)	(0.0015)	(0.0026)	(0.0018)	(0.0021)	(0.0019)	(0.0018)	(0.0021)
	[-0.0273]	[-0.0168]	[-0.0124]	[-0.0176]	[0.0001]	[-0.0296]	[-0.0091]	[-0.0218]
Mean (Y)	0.445	0.350	0.329	0.360	0.416	0.317	0.390	0.331
N	8,579	180,463	37,561	151,481	69,748	119,294	73,598	115,444

Note: This table shows DiD estimates of the impact of the policy on Special Education (SE) removal and educational attainment for different types of districts. The sample includes 5th grade cohorts enrolled in SE between 1999-00 to 2004-05 in our high-impact sample, where "high-impact" is defined as students with a malleable disability (i.e. learning disabilities, speech impairments, other health impairments, or emotional disturbance) who received more than 50 percent of their instruction in GE classrooms at baseline (i.e 5th grade). High wealth districts made up the top 12% of school districts in terms of tax-base wealth. High 504 Plan growth districts grew 504 plan participation by 2 percent between 2005 and 2010 and make up the top 25% of school districts in terms of 504 plan participation growth. Value-added is measured by regressing average standardized test scores on lagged test scores, indicators for a student's race, gender, SE status, ELL status, and FRL status. We split districts according to the median of average test scores and value-added, where those above the median are labelled "High" and those below are labelled "Low". Within each panel, each column reports estimates of δ_1 from a separate regression of Equation 1. The dependent variable is shown in the panel headings. See Table 2 for more detail on the full set of controls. The effect for the fully exposed student at the average district are shown in brackets, and is defined as the coefficient multiplied by 4.5. Standard errors in parentheses are clustered at the district level.