EPIC
Education Policy Innovation Collaborative

Following the Letter of the Law: 2020-21 Retention Outcomes Under Michigan's Read by Grade Three Law

March 2023

Education Policy Innovation Collaborative

COLLEGE OF EDUCATION | MICHIGAN STATE UNIVERSITY

236 ERICKSON HALL, 620 FARM LANE, EAST LANSING, MI48824 | www.EPICedpolicy.org

DISCLAIMER

The Education Policy Innovation Collaborative (EPIC) at Michigan State University is an independent, non-partisan research center that operates as the strategic research partner to the Michigan Department of Education (MDE) and the Center for Educational Performance and Information (CEPI). EPIC conducts original research using a variety of methods that include advanced statistical modeling, representative surveys, interviews, and case study approaches. This research used data structured and maintained by the MERI-Michigan Education Data Center (MEDC). MEDC data are modified for analysis purposes using rules governed by MEDC and are not identical to those data collected and maintained by the MDE and/or CEPI. Results, information, and opinions solely represent the author(s) and are not endorsed by, nor reflect the views or positions of, grantors, MDE and CEPI, or any employee thereof.

ACKNOWLEDGEMENT

The authors are grateful to the many staff at the Michigan Department of Education and Michigan's Center for Educational Performance Information for providing the data necessary to conduct this research, in particular Trina Anderson, Rod Bernosky, Vic Bugni, Dr. Sue Carnell, Dr. Delsa Chapman, Kellie Flaminio, Tom Howell, Mike McGroarty, Shelly Proebstle, and Dr. Michael Rice. We also thank Dr. Tara Kilbride for her excellent research support and Drs. Brian Jacob and Christina Weiland for their helpful feedback.

MARCH 2023

Following the Letter of the Law: 2020-21 Retention Outcomes Under Michigan's Read by Grade Three Law

AUTHORS

John Westall, *EPIC, Michigan State University*

Andrew Utter, *EPIC, Michigan State University*

Katharine O. Strunk, *EPIC Faculty Director, Professor of Education Policy, MSU*

ABSTRACT

Early literacy skills are critical to the educational outcomes of young students. Accordingly, 19 states have early literacy policies that require grade retention for underperforming readers at the end of third grade. However, there is mixed evidence about retention's effectiveness and concerns that retention may disproportionately impact traditionally disadvantaged student groups. Using regressions and a regression discontinuity design, we examine retention outcomes under Michigan's early literacy law, the Read by Grade Three Law. We find that Black and economically disadvantaged students are more frequently eligible for retention and retained than their peers. While controlling for students' test performance, particularly their math scores, eliminates this disparity for Black students, it persists for economically disadvantaged students. We show that differences in average math performance, exemption characteristics, district characteristics, and eligibility-induced student mobility across districts do not explain the disparities in the implementation of retention by economic disadvantage status.

INTRODUCTION

The ability to read and write is a critical building block in the education of young students. Students who can read at or above grade level by the end of third-grade graduate high school and attend college at higher rates than students who cannot (Hernandez, 2011; Lesnick et al., 2010). However, proficiency in these areas has stagnated or even declined in many states across the country. For instance, in 2019, the average reading score on the National Assessment of Educational Progress declined in 17 states, and overall scores nationwide have plateaued over the last decade (Green & Goldstein, 2019). In response, over the last 20 years, most states have passed laws to improve literacy instruction and outcomes in students' early elementary years (Cummings et al., 2021).

These early literacy laws provide both support and accountability for students, schools, and districts working to improve literacy outcomes in kindergarten through third grade. At least 33 states provide intervention resources during the summer or before and after school, 32 states notify parents of students identified with a "reading deficiency," and 22 states require retention for third-grade students who do not perform at a specific level on the states' end-of-year summative English Language Arts (ELA) achievement tests (ExcelinEd, 2021). These 22 states often provide extra resources to retained students to enable them to read on grade level by the time they move on to fourth grade.

Michigan passed its early literacy law, the Read by Grade Three Law, in 2016. Like policies in many other states, the Read by Grade Three Law (RBG3) provides intensive support for early elementary teachers through literacy coaching and literacy-focused professional development. It requires that students with a "reading deficiency," defined under the Law as "scoring below grade level or being determined to be at risk of reading failure based on a screening assessment, diagnostic assessment, standardized summative assessment, or progress monitoring," (Act No. 306, 2016) be identified as early as possible, receive an individualized reading improvement plan, and receive additional one-on-one and small-group instruction, among other interventions. In addition, the Law requires districts to retain third-grade students if they do not meet a specified cut-point on the state's end-of-year standardized achievement test, the Michigan Student Test of Educational Progress (M-STEP). However, students who score at or below the cut-point but fall into one of six categories may qualify for "good cause exemptions," allowing them to be promoted to fourth grade despite their M-STEP score with additional intensive assistance. In addition, the Law requires that districts promote students who score just above the ELA cut score but provide them with additional intensive assistance and support through assignments to highly-effective teachers, as determined by the state's rating system, given that these students are still substantially below grade level.

While educators, advocates, and researchers have generally lauded the support and assistance aspects of these early literacy laws, the retention elements are highly controversial. There is mixed evidence on the effectiveness of grade retention for improved short- and longer-term outcomes (Greene & Winters, 2004, 2006, 2007; Hwang & Koedel, 2022; Jacob, 2005; Jacob & Lefgren, 2004; Lorence, 2014; Roderick et al., 2002; Roderick & Nagaoka, Jenny, 2005; Schwerdt et al., 2017; Weiss et al., 2018; Winters & Greene, 2012) and on the potential impacts on student self-efficacy, sense of belonging, and future criminal activity (Crego et al., 2009; Eren et al., 2021; Holmes & Matthews, 1983; Hong & Yu, 2007; Jimerson, 2001; Wu et al., 2010). Moreover, there are concerns that retention will disproportionately impact traditionally disadvantaged student groups, resulting in inequitable harm (Greene & Winters, 2009; Licalsi et al., 2019).

In this paper, we examine retention outcomes under RBG3 in the 2020-21 school year, the first year the retention element of the law was implemented. We define retention outcomes as whether a student scores low enough on the third-grade ELA M-STEP to qualify for retention and whether the student is eventually retained. We focus on the potential disparate implementation of retention for traditionally disadvantaged students, including students of color, low-income students, students with disabilities, English learners, and students in specific kinds of districts (e.g., by urbanicity, size, prior achievement). In so doing, we ask two research questions: 1) *Is retention implemented differentially under the Read by Grade Three Law by student and district characteristics?*; and 2) *Are differences in retention outcomes explained by exemption eligibility?* In addition, we examine the roles of math performance, good cause exemption qualification, district characteristics, and student mobility in explaining disparities in retention outcomes.

To answer these questions, we analyze administrative data from the Michigan Department of Education (MDE), including student and district characteristics and indicators of retention outcomes and the good cause exemptions awarded by districts. We use regression analysis and exploit the sharp test-score cut-off for retention eligibility under the Read by Grade Three Law in a regression discontinuity (RD) research design.

Our paper provides several substantive findings. First, we show descriptive evidence of disparate retention outcomes; Black and economically disadvantaged students are more frequently eligible for retention based on their third-grade ELA M-STEP scores, and districts retain these same groups at higher rates than their peers. Conversely, while male students are more frequently eligible for retention, districts retain eligible female students at higher rates. However, when we account for district characteristics and student test scores, especially math scores, female and Black students are not retained at higher rates than their similarly-performing peers. In contrast, economically disadvantaged students are still retained at higher rates.

We use a regression discontinuity design to further investigate this discrepancy, finding evidence of differential implementation of the retention policy by economic status. While retention-eligible economically disadvantaged students are *more likely* to be retained than their peers just above the test-score cut-off, retention-eligible non-economically disadvantaged students are *no more likely* to be retained than their peers with higher scores. We provide evidence that differences in other student characteristics correlated with economic disadvantage status, including race, math M-STEP performance, qualifications for retention exemptions, and district characteristics, are not driving this disparity. Furthermore, we find no evidence that differential student mobility across school districts by economic disadvantage status eligibility explains these differentials. Altogether, our results support the hypothesis of Licalsi et al. (2019) that students from more advantaged backgrounds can avoid retention through better advocacy by their families.

Our paper proceeds with a description of the Read by Grade Three Law in Section 2. We then place this study in the context of prior research in Section 3. Next, we describe our data and methods in Section 4. Then, we review our results in Section 5. Finally, we discuss our results and their policy implications in Section 6.

BACKGROUND

The Michigan Read by Grade Three Law

The Michigan House of Representatives passed HB 4822, known as the Read by Grade Three Law, in 2016 to improve upon decades of low statewide performance on the NAEP literacy exams. The Law was modeled, in part, after Florida's "Just Read, Florida!" initiative, which advocates credit as helping the state to improve student literacy and increase its national NAEP ranking from 31st to 8th over a decade (Cummings et al., 2021). Like Florida's and other states' early literacy laws, Michigan's Read by Grade Three Law includes requirements for additional early literacy coaches, literacy instructional time, and early intervention strategies (Cummings et al., 2021). Also similar to 22 other state's early literacy laws, the Read by Grade Three Law stipulates that beginning with the 2019-2020 academic year (later postponed until the 2020-21 school year given the COVID-19 pandemic and associated pause in end-of-year standardized testing), all third graders scoring at least one grade level behind expectations on the ELA portion of the M-STEP (Michigan's statewide standardized assessment) would be retained in third grade unless exempted by one of six specific good cause exemptions.

Third-grade students are subject to different outcomes under the Read by Grade Three Law based on their ELA M-STEP scores. Students who score a 1272 or above are promoted to fourth grade. If they score between 1253 and 1271 they are promoted to fourth grade, but with recommended extra reading support. Last, if students score 1252 or below they are eligible for retention in third grade. Districts can provide a good cause exemption to students who score below the cut-off, promoting them to the fourth grade with additional literacy support. Students who do not take the third-grade ELA M-STEP cannot be retained based on their test scores and are therefore not subject to the Read by Grade Three Law retention policy.

Districts have substantial autonomy in retaining or granting good cause exemptions to retention-eligible students. Good cause exemptions (GCEs) enable districts to promote retention-eligible students to the fourth grade if they fall within one or more of the following categories: 1) English learners with fewer than three years of English language instruction; 2) students with disabilities (i.e., those with an Individualized Education Program [IEP] or Section 504 Plan); 3) students who were previously retained and received intensive reading interventions for two or more years; 4) students who have been enrolled in their current district for less than two years and were not provided with an appropriate individual reading improvement plan (IRIP); 5) students who demonstrated proficiency in other subject areas and/or through an alternative assessment or portfolio of work; and 6) students whose parents or guardians requested an exemption, provided that their superintendent agrees that retention is not in the best interest of the student (referred to as a "Parent Request" GCE in the rest of this paper).

Of all the elements of the Read by Grade Three Law, the retention component has been the most controversial and received the most pushback from state legislators, school administrators, and the public (Cummings et al., 2021). The state legislature introduced a bill to remove the retention component of the Law while maintaining the early literacy interventions, although it did not pass (Michigan State Senate, 2019). Some district superintendents declared that they would use good cause exemptions for all eligible students to prevent the retention of any third graders (French & Kalakailo, 2021). Anticipating the potential for inequitable retention outcomes, Michigan governor Gretchen Whitmer worked with non-governmental organizations to educate parents, especially in low-income areas, on their rights to request good cause exemptions (French, 2020).

A Brief Review of the Literature on Student Retention

This paper contributes to the literature related to retention policies in K-12 settings and, in particular, to the implementation of retention policies. To establish the importance of retention for students and communities, we first summarize extant studies related to grade retention's academic and social effects and the effects of the threat of retention on school communities. We then review literature that examines the previous implementation of retention exemptions and the potential for bias in retention policies.

There is mixed evidence to support retention as a mechanism for improving learning. Indeed, a recent meta-analysis of research on the effectiveness of retention found that retention has an average zero effect, but found substantial heterogeneity across contexts, outcomes, and research design (Goos et al., 2021). Numerous studies have examined the effects of test-score-based retention policies by using clearly defined retention-eligibility criteria to generate credible control groups (e.g., Greene & Winters, 2006, 2007, 2009; Jacob & Lefgren, 2004; Mariano & Martorell, 2013; Roderick et al., 2005; Schwerdt et al., 2017; Weiss et al., 2018; Winters & Greene, 2012). These studies include retention policies in various grades, as prior research suggests that the overall differences between early-grade and late-grade retention are minimal (Anderson et al., 2002; Diris, 2017; Silberglitt et al., 2006).

In Chicago, New York, and Florida, researchers found positive short-term effects of third-grade retention on student reading achievement, including modest gains on standardized tests in the years immediately following retention (Greene & Winters, 2004; Jacob & Lefgren, 2004; Mariano & Martorell, 2013; Schwerdt et al., 2017). In addition, retaining low-performing students in the third grade led to improved math achievement beyond third grade (Figlio & Özek, 2020; Greene & Winters, 2004; Hwang & Koedel, 2022). A recent paper found that retaining third grade students in Indiana led to immediate improvements in ELA achievement, with effects persisting into middle school (Hwang & Koedel, 2022). Other studies have found longer-lasting impacts; students retained by their districts at younger ages based on their standardized test scores had higher high school GPAs and took fewer remedial courses (Schwerdt et al., 2017; Winters & Greene, 2012).

In contrast, Weiss et al., 2018 found no evidence of gains in reading achievement for retained third-grade students in North Carolina. Moreover, elementary and middle school retention policies adversely impacted longer-term outcomes by reducing retained students' high school credit hours, decreasing state high school exit exam participation, and increasing high school dropout rates (Eren et al., 2017; Mariano et al., 2018).

Related literature also provides mixed evidence regarding retention's effects on retained students' social-emotional development. Meta-analyses found that retained students demonstrate lower levels of social-emotional development relative to their promoted peers (Holmes & Matthews, 1983; Jimerson, 2001). Other research shows that retention policies can lead to the increased likelihood of disciplinary incidents and suspensions for retained students and connects retention and increased crime later in life (Eren et al., 2021; Ozek, 2015)

However, context may matter a great deal. In New York City, retention in grades 3 through 8 did not significantly affect school suspensions (Mariano et al., 2018). Louisiana's eighth-grade retention policy decreased the probability of a student being convicted of a juvenile offense (Eren et al., 2017). Moreover, first and third-grade retention positively affected students' perceived sense of belonging within a school (Crego et al., 2009; Hong & Yu, 2007; Wu et al., 2010).

Of course, retention policies may indirectly impact adult practices and student outcomes, even for students who are not retained. The mere threat of retention may change teachers' and administrators' practices and student effort or outcomes. For instance, teachers in Florida, Arizona, and New York City changed their instruction practice due to the threat of retention, focusing more on teaching test-taking strategies (Crego et al., 2009; Perrault & Winters, 2020). However, the effect of the threat of retention on student outcomes is not well studied. Existing evidence suggests a slightly positive impact on student achievement. For example, Roderick et al. (2002) found that a policy in Chicago that would retain students at various points in elementary and middle school based on their performance on standardized tests improved reading and math scores for all sixth and eighth-grade students, on average, and not just for those who were retained. However, the same policy did not significantly impact reading achievement outcomes for third-grade students (Roderick et al., 2002). Related work suggests that the threat of retention did not affect student achievement in non-tested subjects or subjects not used in retention decisions (Jacob, 2005). This finding suggests that the retention-induced student achievement growth demonstrated in earlier studies may not translate into improvements in other academic outcomes.

In addition to examining the benefits and costs of retention to students, a related literature estimates that the cost of retention to taxpayers could be as high as $42,000 per student (Babcock & Bedard, 2011; Eide & Goldhaber, 2005). However, recent research suggests that these estimates overstate the cost of retention, finding that third-grade retention under Florida's early literacy policy was 45% less costly than previously thought because these previous estimates failed to account for the time value of money and that retention leads to less than a full year of additional instruction for retained students (Winters, 2022).

Other research has investigated potential disproportionality in test-based retention policies across multiple states. For instance, before policymakers implemented retention in Georgia, Livingston and Livingston (2002) predicted that because Black and poor students were more likely to fail the state's standardized test, they would also be more likely to be retained. In Texas, Valencia and Villareal (2005) found that Latino/a students were more likely to be retained based on their test scores.

As in Michigan, retention laws in other states allow for good cause exemptions to grade retention. Such exemptions are common. For instance, in 2002-03, only 54% of third-grade students in Florida who scored below the threshold for promotion were retained, 25% were promoted due to an exemption, and the rest were promoted for no specific reason (Greene & Winters, 2009). However, there are differences in the types of students who receive GCEs; in Florida, Black and Latino/a students were less likely to receive exemptions even when controlling for academic achievement (Greene & Winters, 2009). Also, in Florida, Licalsi et al. (2019) found that retention-eligible students with more educated mothers were more likely to receive GCEs and be promoted.

Together, these mixed findings contribute to the controversy surrounding grade retention, especially given the extent to which these policies disproportionately impact the most vulnerable students. While the evidence clearly demonstrates that students of color and low-income students are most often retained, the effects of such retention have varied between contexts. This paper contributes to the literature on retention by examining the first year of retention implementation in Michigan to understand better the impact of such policies on historically marginalized communities. In particular, our study builds on prior work examining retention requirements (see e.g., Greene & Winters, 2009; Licalsi et al., 2019; Livingston and Livingston, 2002; Valencia and Villareal, 2005) but is most similar to Licalsi et al. (2019), who use an RD design to examine the differential implementation of retention in Florida by socioeconomic status. They found that students with less educated mothers were more likely to be retained under Florida's law and that much of the disparity is driven by the fact that students with more educated mothers are more likely to receive subjective retention exemptions. They hypothesize that more advantaged households are better able to advocate for exemptions.

Our study is distinguished from and contributes to the extant body of work in several ways. First, the context of our study is substantively different. Retention under Michigan's RBG3 Law started nearly two decades after Florida began retaining third-grade students, and contexts and laws have changed substantially in that time. Additionally, in Michigan, overall retention rates before the Read by Grade Three Law's retention component's implementation were roughly 0.7%, relative to just over 3% in Florida (Licalsi et al., 2019). This difference in the prevalence of retention likely reflects differences in beliefs about retention's effectiveness, meaning educators in Michigan might be more reluctant to retain even the lowest-performing students.

Differences in retention-eligibility criteria further evidence this reluctance to retain. Whereas roughly 20% of Florida's third-grade students were retention-eligible each year studied (Licalsi et al., 2019), just 4.8% of tested Michigan third-grade students were retention eligible in 2020-21. This implies that retention-eligible students in Michigan are substantially lower performing than those in Florida, and because RD designs identify local average treatment effects, the findings in Licalsi et al. (2019) may not be valid at lower quantiles of the achievement distribution. There are also differences in the retention exemptions provided under the Michigan Law. Notably, parents in Michigan can request an exemption solely because they believe it is in the best interest of their child, and can work with local administrators to ensure their child is promoted (Act No. 306, 2016). Florida's policy does not provide an exemption of this type. The "parent request" exemption in Michigan offers another avenue for differential implementation by socioeconomic status and other student characteristics.

Second, our study examines the implementation of a retention requirement during the COVID-19 pandemic. The pandemic has had particularly detrimental impacts on historically disadvantaged students (Hatch & Harbatkin, 2021). Examining the implementation of a universal retention policy during the pandemic and comparing to previous studies in Florida (Greene & Winters, 2009; Licalsi et al., 2019), Georgia (Livingston and Livingston, 2002), and Texas (Valencia and Villareal, 2005) will help us understand if the pandemic has exacerbated inequities in retention outcomes.

Finally, we examine the role of student mobility in the differential implementation of retention. We examine whether differential changes in the probability of district switching by economic disadvantage explain some inequities in retention outcomes. While economically disadvantaged students tend to have higher mobility rates, wealthier students will potentially have more resources to switch districts to avoid grade retention in response to retention eligibility, as retention decisions are made at the district rather than the state level under the Read by Grade Three Law.

DATA

We use administrative education data from the Michigan Department of Education and the Center for Educational Performance Information regarding Michigan third-grade students in 2020-21. These data contain math and ELA M-STEP performance and participation, Read by Grade Three retention eligibility status, special education and English learner program participation, and other demographic characteristics. There were 102,138 third-grade students in Michigan in the 2020-21 school year who we examined and followed into the fall semester of 2021-22 to determine their final grade retention status. We exclude from our analyses six students who are missing demographic information and 2,079 students who leave Michigan public schools before the fall semester of 2021-22 and are missing grade information because we cannot determine their final retention status.[1] We retain 100,053 Michigan third-grade students in our analysis sample.

Our primary outcomes of interest relate to the final grade retention status of Michigan's 2020-21 third-grade students. First, we define $RBG3Retention_i$ to indicate retention attributable to the Read by Grade Three Law. We say a student is retained because of the Read by Grade Three Law ($RBG3Retention_i = 1$) if they were retention eligible, did not receive a GCE, and were enrolled in third grade in the Fall of 2021-22. Next, we define grade retention, $Retention_i$, more generally such that $Retention_i$ equals one if a student was again enrolled in third grade in the Fall of 2021-22 for any reason.

Table 1, Column (1), shows summary statistics of the third-grade analytic sample. Roughly half of the sample is female, and a little over half are economically disadvantaged.[2] Just over 60% are white, while a little under 20% are Black. About 10% of the total sample are designated English learners or previously retained in grade. Importantly because of good cause exemptions waiving required retention for certain groups, just over 15% are students with disabilities, and just under 8% have been enrolled in their current district for less than two years. In Michigan, about 12% of the third-grade sample attends a charter school, also known as Public School Academies

(PSAs). A majority of students attend school in a suburb or town or in a district that performed in one of the top two quartiles of the ELA M-STEP in 2019.

Michigan implemented RBG3's retention component amidst the COVID-19 pandemic, and a smaller proportion of third-grade students took the M-STEP than would be the case in a typical year. Only 72% of our analysis sample of Michigan's third-grade students, 72,495 students, took the ELA M-STEP exam in 2020-21, relative to nearly 96% in the 2018-19 school year (the last time Michigan administered the M-STEP, before the COVID-19 pandemic). Table 1, Column (2), shows summary statistics for Michigan third-grade students who participated in the 2020-21 ELA M-STEP. The statistics imply significant heterogeneity in participation rates across student subgroups. For example, Black and Latino/a students are under-represented among tested students and, thus, were less likely to take the exam than their White peers.

An important implication of these low participation rates arises from the fact that only tested students can be retained under RBG3. Since traditionally disadvantaged groups took the ELA M-STEP at lower levels than their peers, these students are at lower risk of grade retention under the Law, all else equal. Moreover, test participation is likely positively correlated with potential test performance, meaning higher performing students were more likely to take the ELA M-STEP. Altogether, this suggests that estimates of disparate outcomes or discrimination against historically disadvantaged student groups in retention decisions for this cohort of third-grade students represent lower bounds of the disparate outcomes occurring for these groups in a typical year.

Our analyses focus on retention-eligible students and students near the retention-eligibility threshold. Third-grade students who took the ELA M-STEP could be retention-eligible based on their score (1252 or below, which equates to a z-score of -1.59 or below). Roughly 4.8% of tested third-grade students were RBG3 retention-eligible (3,440 students). Table 1, Column (3), presents summary statistics for retention-eligible third-grade students. These statistics suggest substantial disparities in retention-eligibility between student subgroups. Historically marginalized groups, including Black, Latino/a, and economically disadvantaged students, and students with characteristics aligned with good cause exemptions (students with disabilities, English learners, previously retained students, and students who had been enrolled less than two years at their current district) are over-represented relative to the tested sample in Column (2), meaning they are retention-eligible at higher rates.

Table 1, Columns (4) and (5) compare the characteristics of retention-eligible students by their final retention status. Column (4) shows retention-eligible students promoted through GCEs, while Column (5) shows retained retention-eligible students. Male, White, and non-economically disadvantaged students are overrepresented in Column (4) relative to Column (5), suggesting they are promoted at higher rates than their female, non-white, and economically disadvantaged peers. English learners, previously retained students, students with disabilities, and students enrolled in their district for fewer than two years are also less likely to be retained than their peers without these characteristics, reflecting qualifications for exemptions. Students in charter schools, districts with the lowest levels of prior ELA performance, and rural districts are also more likely to be retained than their peers.

Table 1, Column (6), examines the characteristics of students who were retained for reasons other than the Read by Grade Three Law. We focus on these students because they likely differ from those retained under RBG3. Students retained for other reasons were less likely to participate in the ELA M-STEP (Westall et al., 2021a), and test participation is required for retention-eligibility under the Read by Grade Three Law. Indeed, Column (6) shows that these students are more likely to be Black, enrolled in a PSA, attend a low-performing ELA district, and attend school in a city than the retention-eligible population of third-grade students.

In the next section, we describe the analyses we apply to these data to understand the implementation of retention under RBG3 and any disparate retention outcomes across student subgroups.

METHODS

Regression Models

We begin our analyses by examining correlations between the characteristics of retention-eligible students and their retention outcomes. We estimate univariate regressions on the sample of retention-eligible third-grade students:

$$P(RBG3\ Retention_i) = \beta_0 + \beta_1 W_i + u_i \qquad (1)$$

Where the outcome, $P(RBG3\ Retention_i)$, represents the probability of retention because of RBG3 for student i and u_i are idiosyncratic errors. The variable W_i represents a single student or district characteristic. Thus, β_1 measures how much the probability of retention changes given a one-unit increase in W_i. We estimate separate regressions for each characteristic listed in Table 1 and cluster standard errors at the district level because retention decisions are localized by district.

This univariate regression analysis allows us to examine unconditional relationships between student characteristics and retention outcomes, but these characteristics are correlated with one another and with retention outcomes. In our subsequent analysis, we examine the relationship between student and district characteristics and retention outcomes, holding other factors constant. Through this analysis, we intend to measure whether GCE qualifications, M-STEP performance, and district characteristics explain the relationship between students' demographics or economic disadvantage status and retention outcomes. To do this, we estimate:

$$P(RBG3Retention_i) = \beta_0 + X_i\beta_1 + GCE_i\beta_2 + D_d\beta_3 + \epsilon_{id} \qquad (2)$$

Where the outcome, $P(RBG3\ Retention_i)$, again represents the probability of retention because of RBG3 for student *i* and u_i are idiosyncratic errors. We first control for X_i, a vector of individual characteristics including race, gender, and economic disadvantage. Then, we control for GCE_i, a vector of characteristics associated with the various GCEs, to examine whether eligibility for GCEs can explain differences in retention rates by student characteristics. Next, we include controls for D_d, a vector of district characteristics including M-STEP participation rates, student body demographics, district locales, previous M-STEP performance quartiles, and district size, to account for relationships between district and student characteristics that may explain disparities in retention across student groups.

We estimate both sets of models using a linear probability model. While linear probability models can estimate predicted probabilities of greater than one or less than zero, the parameter estimates are unbiased. However, results using logistic regression models return substantially the same results (available from the authors upon request).

Regression Discontinuity Design

In our previous analyses, we examine whether disparities in retention outcomes exist for retention-eligible students. However, differences in the implementation of retention under RBG3 might still exist. Next, we examine heterogeneity in the causal effect of RBG3-retention-eligibility on grade retention by student and district characteristics to understand whether retention is implemented differentially across student groups near the retention-eligibility cut-off.

We identify the causal effect of retention-eligibility on retention by leveraging the retention-eligibility cut-off in a regression discontinuity design (RDD). Our identification strategy compares changes in the probability of retention across the retention-eligibility threshold for students with test scores close to the cut-off. We estimate the following local linear regression model:

$$P(Retention_i) = \delta_0 + \delta_1 \mathbf{1}[Score_i \leq 1252] + \delta_2 f(Score_i) + e_i \qquad (3)$$

As before, $P(Retention_i)$ is the probability of retention for student *i*. The indicator, $\mathbf{1}[Score_i \leq 1252]$, equals one if student *i* had an ELA M-STEP scale score less than or equal to the 1252 retention-eligibility threshold. We control for $f(Score_i)$, a flexible function of the ELA M-STEP score. The coefficient of interest is δ_1, which represents the local average treatment effect (LATE) of being just below the retention-eligibility cut-off on the probability of retention. We then examine heterogeneity in the local average treatment effect by student characteristics. We estimate the RD model separately on different subgroups of students, then compare their LATEs.

In our preferred specification, $f(Score_i)$ is linear, implying a local linear regression model. We select the bandwidth around the test score cut-off using the mean squared error-optimal bandwidth selection procedure in Calonico et al. (2020). Our preferred method allows the bandwidth to differ above and below the cut-off. The optimal bandwidth varies by subgroup. We ensure changes in the bandwidth across subgroups do not drive our findings by selecting a single bandwidth for all our analyses. We first compute the MSE-optimal bandwidth overall and for each subgroup. We then use the average of these bandwidths rounded down to the nearest scale score. This procedure gives us a bandwidth of 12 points below and 30 points above the cut-off. Students within this bandwidth are very low performing relative to the average student. In particular, 12 scale score points below the cut-off captures 80% of all retention-eligible students.

To understand whether our estimates are sensitive to different bandwidths, we test the robustness of our estimates to other bandwidth selection methods. First, we use +/- 12 scale score points, representing the average bandwidth from the mean squared error-optimal bandwidth selection procedure (Calonico et al., 2020) across subgroups with equal bandwidths above and below the threshold. Second, we use +/-7 scale score points from the cut-off, selected because this is the minimum bandwidth from the mean squared error-optimal bandwidth selection procedure (Calonico et al., 2020) across subgroups with equal bandwidths above and below the threshold.

After selecting the estimation bandwidth, we use triangular weights to limit our estimation sample to individuals within the bandwidth around the cut-off. We perform inference using robust nonparametric confidence intervals following Calonico et al. (2014). To generate these confidence intervals, we select a bandwidth following the same procedure outlined above, which produces a bandwidth of 18 scale score points below the cut-off and 47 points above. We implement estimation using software developed by Calonico et al. (2017).

Validity of the RD Design

The data must satisfy two key assumptions for the RD estimates to have a causal interpretation. First, there must be continuity of expected potential outcomes around the cut-off. In other words, in the absence of the RBG3-retention requirement, the probability of retention would have changed smoothly across the test-score threshold. We test the continuity assumption by examining how average individual and district characteristics change across the cut-off. Moreover, we perform placebo analyses around other test-score cut-offs, including a 1272 scale score ELA M-STEP, representing an achievement level at which the state recommends promotion without additional literacy support, and a 1252 scale score Math M-STEP, which has no policy relevance to retention decisions under RBG3. We find no evidence of other discontinuities and describe the results of these robustness tests in Section 6.

A second assumption is that there is no manipulation of ELA M-STEP scores around the cut-off. Direct manipulation of M-STEP scores is unlikely because students and teachers do not have a priori knowledge of what a student's test score will be, even given the number of correct or incorrect item responses.[3] If manipulation were occurring, one would expect discontinuities in the distribution of ELA M-STEP scores around the cut-off, and there could be a bunching of test scores just above the threshold relative to just below. We examine the smoothness of the distribution of ELA M-STEP scores near the cut-off, shown in Figure 1, and find no evidence of bunching just above or below the cut-off.

A final caveat relates to the discrete nature of ELA M-STEP scores. Continuity-based RD designs assume that the forcing variable is continuous. M-STEP scores are discrete by design, with one scale score increments. Thus, our running variable contains mass points, or values shared by many observations. When mass points exist, local polynomial methods behave as if there are as many observations as mass points – in our case, scale score points – not the total number of observations, and estimation is based on extrapolation from the points just above and just below the cut-off (Lee & Card, 2008). However, in the case of M-STEP scale scores, where scores are implicitly measured with error, this extrapolation around the cut-off is natural and makes intuitive sense. In our main analysis, we follow Lee & Card (2008) and cluster the standard errors at the scale-score level.

RESULTS

Student, School, and District Characteristics and Retention Outcomes

We begin by examining differences in the probability of retention for retention-eligible students by estimating Equation (1) with the various student and district characteristics described in Section 3. The sample consists of all retention-eligible students, and the outcome of interest is grade retention related to RBG3. As in other states with similar retention policies and as suggested in Table 1, we find significant disparities in the probability of retention across student subgroups. Each row in Column (1) of Table 2 shows coefficients from univariate regressions such that each row is a coefficient from a separate univariate regression. We find that historically marginalized groups are more likely to be retained than other groups of retention-eligible students. Relative to students of other races, retention-eligible Black students were 2.2 percentage points more likely to be retained. Economically disadvantaged students were 3.3 percentage points more likely than their wealthier peers to be retained, and female students were 2.0 percentage points more likely to be retained than their male peers.

Students with characteristics aligned with good cause exemptions (students with disabilities, English learners, and previously retained students) were generally less likely to be retained -- with the notable exception of students who had been enrolled less than two years in their current districts. These students were potentially more likely to be retained because we imperfectly measure eligibility for the mobility GCE. The GCE requires that students have been enrolled for less than two years at their current district *and* that the student's current district has not provided them with an IRIP. We can only measure the length of attendance in a student's district.

Disparities in retention also exist across districts. Students in districts that performed worse on the ELA M-STEP in 2019 were more likely to be retained. Students in charter schools were also more likely to be retained than students in traditional public schools. Additionally, retention-eligible students in districts with higher test participation rates and lower enrollments were more likely to be retained.

This analysis provides suggestive evidence of disparities in the retention outcomes of Michigan's third-grade students. While these analyses paint a clear picture of student retention outcomes across the state, they do not provide information about whether these results are in some way different for students with similar backgrounds and achievement profiles. For instance, if more Black and low-income students are retained, and fewer are given GCEs, even holding constant their performance on state standardized tests, this might point to concerning biases ingrained in the implementation of the Read by Grade Three Law. We, therefore, turn to multivariate regression analysis to understand whether districts are more likely to retain students in historically marginalized groups once we account for student characteristics associated with GCEs, student performance on the state summative standardized achievement tests, and district characteristics.

In Table 2, Column (2), we provide results from multivariate regressions that control for student-level covariates, including gender, race, and economic disadvantage status. We find that, when controlling only for students' race and ethnicity and economically disadvantaged status, economically disadvantaged students are 2.7 percentage points more likely, and female students are 1.8 percentage points more likely than their wealthier and male retention-eligible peers to be retained. Black students were also 1.9 percentage points more likely to be retained than their White peers, but this is only marginally significant at $p<.10$. Thus, when controlling for economic disadvantage, it does not appear that districts are significantly more likely to retain Black or Latino/a students than White or Asian students.

In Column (3), we add the set of student characteristics explicitly associated with GCEs. Because students with these characteristics should automatically qualify for a retention exemption, adding these controls to the model should explain student retention outcomes and potentially further mitigate the relationships between race, income, and eventual retention. As expected, we see that the covariates related to GCEs significantly negatively correlate with retention. Students with disabilities are 3.0 percentage points less likely, English learners are 2.3 percentage points less likely, and previously retained students are 3.3 percentage points less likely to be retained than their other retention-eligible peers. However, economically disadvantaged students are still significantly more likely to be retained than their wealthier peers, and the inclusion of the GCE covariates does not change the magnitude of this relationship. In other words, even controlling for students' statuses as special education students or English learners, as well as their previous mobility and retention patterns, retention-eligible lower-income students are still 2.8 percentage points more likely to be retained in the third grade. However, the significance and magnitude of gender's impact decrease and race is no longer a significant predictor of retention, even at the 0.10 level. Thus, these results suggest that the correlations between economic disadvantage and GCE-related characteristics do not explain the economic disadvantage status gap in retention outcomes, but they contribute to the explanation of the relationship between gender and race in retention outcomes.

Columns (4) and (5) add measures of student M-STEP scores in ELA (Column 4) and ELA and math (Column 5). ELA scores are never a significant predictor of retention in our models. This is likely because there is little variation in the ELA scores of retention-eligible students – all by definition score below 1252 scale score points – so they have little predictive power. However, math scores significantly predict retention outcomes in all models; a one standard deviation increase in math scores is associated with a 2.5 to 3.1 percentage point decrease in the probability of retention.

Moreover, once we account for both ELA and math M-STEP scores, female students are no longer significantly more likely to be retained than their male peers at any conventional significance level. One potential explanation for the relationship between math scores and retention could be that districts and parents use students' math proficiency to indicate that students do not need to be retained. However, controlling for ELA and math achievement does not fully explain the discrepancies in retention outcomes for economically disadvantaged students; we still find that economically disadvantaged students are 2.4 percentage points more likely to be retained than their wealthier peers. Together, these results suggest that test scores, particularly math scores, explain much – although not all – of the apparent disparities in retention outcomes across student groups.

As discussed above, individual districts have substantial discretion over retention decisions. 76.8% of Michigan school districts chose to promote all of their retention-eligible students, and another 19.1% and 4.2% chose to promote only some or all of their retention-eligible third-graders at the end of the 2020-21 school year (Westall et al., 2021). Thus, it is important to account for district characteristics to determine if individual districts drive potential inequitable retention outcomes.

In Column (6), we add controls for district characteristics. We find charter schools were significantly more likely to retain eligible students, as were schools in rural districts. Given that charter schools are home to a substantial proportion of Michigan's Black student population (52% of third-grade charter school students in Michigan are Black compared to 19% of Michigan's total third-grade student body) and that both rural and charter schools serve high proportions of economically disadvantaged students (80% of charter school and 54% of rural third-graders are considered economically disadvantaged), some of the disparities in retention rates for Black and economically disadvantaged students may be attributable to these differences in retention outcomes across districts. Indeed, once we control for district characteristics, the difference in retention probability between Black and White students becomes negative, although not statistically significant. However, economically disadvantaged students are still significantly more likely to be retained than wealthier students.

Differential Implementation of Retention

We have thus far demonstrated disparities in retention outcomes for retention-eligible students, particularly economically disadvantaged students, who are retained at higher rates than their wealthier peers with the same test performance, demographics, and observable GCE qualifications. In this section, we look for further evidence that retention under RBG3 was implemented differently for different student groups. We do this by examining heterogeneity in the effect of retention eligibility on retention across student subgroups in an RD design.

We begin by estimating the overall impact of retention eligibility on the probability of retention for students just above and below the 1252 scale score cut-off on the ELA M-STEP. Figure 2 shows a graphical representation of the RD. The x-axis is the ELA M-STEP scale score recentered such that a 1253 scale score equals zero, and the y-axis is the probability of retention. There is one dot per scale score point, representing the average retention probability for students with a given ELA M-STEP scale score. The figure also has best-fit lines from a local linear regression using a triangular kernel. We see that, as expected, retention-eligible students scoring just below the cut-off are more likely to be retained than students scoring just above the cut-off. Table 3 provides more detailed estimates. Column (1) contains the point estimate of the impact of retention eligibility on the probability of retention, Column (2) is the robust p-value, Column (3) shows the robust 95% confidence interval, and Columns (4) and (5) are the number of observations above and below the cut-off, respectively. The first row of Table 3 indicates that scoring just below the RBG3-retention cut-off causes a 3.8 percentage point increase in the probability of third-grade retention, significant at less than the 0.1% level. Relative to the 0.8% probability of retention just above the cut-off, this represents a substantial 450% increase in the probability of retention relative to the average just above the cut-off.

Next, we turn to heterogeneity in the effect of retention eligibility on retention. We focus on heterogeneity by three student characteristics for which we have established substantial disparities in retention outcomes in our descriptive analyses above: gender, race (specifically Black relative to White students), and economic disadvantage status.[4] We present these findings in Figure 3 and in more detail in Table 3. We find no substantial differences in the impact of retention eligibility across gender (Panels E and F) or race (Panels C and D), consistent with the findings from our regressions once we had controlled for M-STEP scores, GCE qualification, and district characteristics. Specifically, Black students are 4.3 percentage points more likely to be retained if they are retention eligible, which is close in magnitude to the 3.4 percentage point increase in White retention-eligible students' propensity to be retained. The confidence intervals of the estimates for Black and White students overlap, suggesting that they are not significantly different. Similarly, the increases in the likelihood that retention-eligible females and males are retained are statistically indistinguishable. However, retention eligibility has a substantially larger impact on the probability of retention for economically disadvantaged students than for their wealthier peers. Panels A and B show that Retention-eligible economically disadvantaged students are 4.5 percentage points more likely to be retained if just below the cut-off. In contrast, non-economically disadvantaged students are no more likely to be retained if they are just retention eligible than if they were just above the cut-off.

Our findings are broadly consistent with those of the regression analysis above and raise the question, why does retention-eligibility impact economically disadvantaged students differently than their peers with otherwise similar characteristics? In the following sections, we examine the roles of math performance, good cause exemption qualification, district characteristics, and student mobility in explaining this disparity.

The Role of Math Performance

We have shown that math M-STEP scores significantly predict retention for retention-eligible students, potentially because districts and parents gauge the necessity of retention based partly on math proficiency. However, controlling for math M-STEP performance did not fully explain the disparities in retention rates for economically disadvantaged retention-eligible students relative to their non-economically disadvantaged peers. In this section, we examine the role of differences in math performance in explaining the differential implementation of retention-eligibility by economic disadvantage status.

We begin by estimating differences in the effect of retention eligibility on retention across the math M-STEP performance distribution by grouping students into three policy-driven performance groups: (1) scale score <=1252 (retention eligible); (2) scale score of 1253 to 1272 (promote but with additional interventions); and (3) scale score >=1273 (promote without additional interventions). Roughly a third of tested third-grade students fall into each math performance category. Figure 4 shows graphical representations of the regression discontinuity estimates across math performance groups, and Table 4 shows point estimates and estimation details. We find little evidence of differences in the effect of retention-eligibility on retention by math M-STEP performance; even though only the lowest-performing students on the Math M-STEP are significantly more likely to be retained, the point estimates are quite similar across performance groups. The confidence intervals for these estimates overlap entirely, indicating that the estimates are statistically indistinguishable.

To determine whether systematic differences in math performance explain disparities in the effects of retention eligibility by economic disadvantage status, we examine whether there is math treatment effect heterogeneity by economic disadvantage status. In Table 5, Columns (1) through (4), we restrict the sample to only include non-economically disadvantaged students and estimate treatment effect heterogeneity across other characteristics. We perform a parallel analysis for economically disadvantaged students in Table 5, Columns (5) through (8). Column (1) provides no evidence that retention eligibility increases the probability of retention for non-economically disadvantaged students across the math performance distribution. In fact, students with near median levels of performance, scoring 1253 to 1272, are *less* likely to be retained if they score just below the retention-eligibility cut-off than above. Column (5) finds similar effect sizes across the math score distribution, consistent with our findings overall. Altogether, these results suggest that differences in math performance cannot explain disparities in the effects of retention eligibility by economic disadvantage.

The Role of Good Cause Exemption

Our results presented in Section 5.1 show that many of the ostensible discrepancies in retention outcomes across student groups are explained by differences in observable qualifications for GCEs. However, the discrepancy in retention rates between economically disadvantaged students and their wealthier peers persists even when we include those covariates in our model in Table 2, Column (3). In this section, we further explore the potential for GCE qualification to explain disparities in retention outcomes by economic status.

We first examine differential GCE qualification rates by student subgroups in Table 4. If non-economically disadvantaged students are more likely to qualify for GCEs, this could explain why we see no change in the probability of retention across the retention-eligibility threshold for non-economically disadvantaged students. The fifth and sixth rows of Table 4 show that among all third-grade students, economically disadvantaged students are more likely to qualify for a GCE. However, when we examine only retention-eligible students, non-economically disadvantaged students are substantially more likely to have a 504 Plan or IEP than economically disadvantaged students. For other observable GCE qualifications, retention-eligible economically disadvantaged students continue to qualify at slightly higher rates than their wealthier peers.

If we found a null effect of retention-eligibility on retention probabilities for non-economically disadvantaged students near the cut-off because those students qualify for a “students with disabilities GCE” at higher rates, we would expect to see an increase in the probability of retention across the eligibility threshold for non-economically disadvantaged students who are not students with disabilities. Table 5 provides no evidence that retention-eligible non-economically disadvantaged students who are *not* students with disabilities are more likely to be retained.[5] In fact, we find no significant impact of retention eligibility for non-economically disadvantaged students across all observable GCE qualification groups. On the other hand, economically disadvantaged students *with* a disability are more likely to be retained if they are just below the retention-eligibility threshold than just above, even though they qualify for a GCE. Thus, our findings suggest that differential qualifications for GCEs cannot explain differences in the implementation of retention under RBG3 by economic disadvantage status.

The Role of District Characteristics

Differences in retention decisions across districts could explain the apparent disparities in retention outcomes. For instance, charter schools and rural districts, which serve higher proportions of economically disadvantaged students, were also more likely to retain students under RBG3. Yet, even when accounting for these differences in district characteristics, economically disadvantaged students are still significantly more likely to be retained. In this section, we further examine the role of district characteristics in explaining disparities in retention outcomes.

If certain districts are more lenient with respect to their retention decisions, and those districts tend to serve wealthier students, this relationship could explain the apparent disparity in retention implementation by economic status. We examine whether this is the case by comparing the effect of retention eligibility on retention across district characteristics for non-economically disadvantaged students. The bottom rows of Table 5 present our results.[6] Column (1) provides no evidence that non-economically disadvantaged students in any of these observable district types were significantly more likely to be retained if they scored just below the eligibility threshold. In contrast, Column (5) shows that retention eligibility increased the probability of retention for economically disadvantaged students in all district types, except those in districts with the highest levels of prior ELA performance. Together, we find no compelling evidence that district characteristics explain the disparity in retention implementation by economic status.

The Role of Student Mobility

Another potential driver of disparities by economic disadvantage status is student mobility. Families might change school districts in response to retention eligibility for various reasons. Families might move in response to inferred poor school quality or switch to a district that will promote their retention-eligible child to fourth grade. While economically disadvantaged students tend to be more mobile on average than their wealthier peers, potentially due to differences in housing, family, and financial instability (Hanushek et al., 2004; Xu et al., 2009), wealthier families may be more responsive to the threat of retention because changing districts is costly. If district changes are associated with grade promotion for wealthier families, then this could explain the null effect of retention-eligibility on retention for non-economically disadvantaged students.

Table 6 and Figure 7 provide results from the RD analysis of the impact of retention-eligibility on the probability of switching districts between spring 2021 and fall 2021. The first row of Table 6 provides point estimates and estimation details. We find that retention-eligible students just below the cut-off are 3.3 percentage points more likely to switch districts than students just above the cut-off, implying a 41% increase relative to the 8% district -switching rate just above the cut-off. This is, on its own, an important finding, suggesting that families may be choosing to move districts in order to avoid being retained.

The next rows of Table 6 show how the retention-eligibility-induced mobility decisions of students differ by student subgroups.[7] We focus on the second and third rows, examining differences by economic status. Economically disadvantaged students and their wealthier peers are both 3.2 percentage points more likely to switch districts. While these point estimates are nearly identical, they represent substantially different effect magnitudes relative the baseline district switching rates of students by economic disadvantage status. Just over 9% of economically disadvantaged students just above the cut-off switched districts, implying that retention eligibility increased the probability of switching districts by 34%. On the other hand, 5% of non-economically disadvantaged students just above the cut-off switched districts, suggesting a 67% increase in the probability of district switching resulting from retention eligibility.

These results mean that differences in mobility could explain the disparities in the implementation of retention-eligibility by economic disadvantage. However, if this were the case, we would expect that retention eligibility would increase the probability of retention for non-economically disadvantaged students who do not switch districts. When we examine this in the bottom two rows of Table 5, we see this is not the case. Column (1) shows that retention-eligible, non-economically disadvantaged students who do not switch districts are no more likely to be retained than similar students just above the eligibility threshold. In contrast, retention eligible, economically disadvantaged students are significantly more likely to be retained than their peers above the cut-off regardless of whether they switch districts or not. These findings suggest that differences in mobility do not fully explain disparities in the implementation of retention by economic disadvantage status.

ADDITIONAL ROBUSTNESS CHECKS

Smoothness of Other Characteristics, Placebo Cut-Offs, and Quadratic Functional Form

The RD approach assumes that potential outcomes are continuous around the cut-off. While it is impossible to observe potential outcomes, we can infer information about the plausibility of this assumption by examining the smoothness of student and district characteristics, theoretically unaffected by the retention-eligibility cut-off, around the threshold. Appendix Figure A3 shows RD estimates using student characteristics, math M-STEP scores, and charter school status as outcomes. We see no compelling evidence of discontinuities in these outcomes across the retention-eligibility threshold, supporting the continuity assumption.

We further test the validity of our RD design by examining the robustness of our results to placebo test-score cut-offs that have no policy relevance to retention decisions under RBG3 or otherwise. A significant change in retention probability across the placebo cut-off could indicate the probability of retention is changing for reasons other than RBG3. In Appendix Figure A4, Panel A, we run our preferred model specification using ELA M-STEP scale scores as the running variable but with a placebo cut-off of 1272. In Appendix Figure A4, Panel B, we use math M-STEP scores as the running variable with a placebo cut-off of 1252. We find no evidence of a statistically significant change in the probability retention across either of these placebo cut-offs, providing further support for the validity of our design.

Finally, we test whether the choice of a linear function form for the running variable is driving our estimates by re-estimating the models with a quadratic functional form. We re-estimate Table 3, where we examine treatment effect heterogeneity by student characteristics with a quadratic functional form in Appendix Table A1. Appendix Table A2 replicates Table 5, examining subgroup heterogeneity by economic disadvantage status. Using a quadratic functional form increases the disparity in retention eligibility's effect on the probability of retention by economic disadvantage status. Appendix Table A1 shows that non-economically disadvantaged students are significantly *less* likely to be retained if they score just below the threshold. Moreover, Appendix Table A2 shows that more subgroups of non-economically disadvantaged students are *less* likely to be retained if they are retention eligible. Overall, our findings are robust to this alternative functional form.

Alternative Bandwidths

Our preferred bandwidth selection method draws from Calonico et al. (2020) and selects the average MSE-optimal bandwidth across all subgroups we analyze. We also allow the bandwidth to differ above and below the cut-off. Our preferred bandwidth is 12 scale score points below and 30 scale score points above the cut-off. In this section, we test the robustness of our estimates to other cut-offs, including +/- 12 and +/- 7 scale score points from the cut-off. We replicate our main findings in Table 3 and Table 5 using both alternative cut-offs. Appendix Tables 3 and 4 use a bandwidth of +/- 12 scale score points from the cut-off. The results are qualitatively similar to our preferred bandwidth. Appendix Tables 5 and 6 use +/- 7 scale score points from the cut-off. The only substantive difference here is that several subgroups of non-economically disadvantaged students are significantly less likely to be retained if they score just below the eligibility threshold, which does not change our conclusions. Overall, our results are robust to alternative cut-offs.

Differential Attrition

Our analyses exclude students who are missing information regarding their Fall 2022 grade level. This sample restriction generates attrition that could be a concern for our identification strategy if attrition is correlated to both retention eligibility and grade retention. In Appendix Table A7, we examine whether retention eligibility affects attrition rates overall and by sub-groups. We find that students just below the retention-eligibility cut-off are 0.2 percentage points less likely to be missing Fall 2022 grade information than their peers just above the cut-off (p-value of 0.07), implying that retention eligibility *reduces* the probability of sample attrition. Importantly, this differential attrition appears concentrated among non-economically disadvantaged students, for which retention eligibility reduces the probability of attrition by 0.8 percentage points (p-value of 0.01).

This type of differential attrition could bias our estimates toward zero if students who leave the data Fall 2022 and are not retention-eligible are more likely to be promoted than their peers who remain in the data. A substantial bias of this type could lead us to falsely conclude that retention eligibility has a null impact on retention for non-economically disadvantaged students. This is particularly salient because differential attrition in our setting appears to disproportionately impact this group. We test whether a bias of this type could possibly occur in our setting by bounding our estimates to account for attrition using a method similar to Lee (2009). In this bounding exercise, we estimate a worst-case scenario upper bound by assuming that all students missing Fall 2022 grade information *above* the retention-eligibility cut-off were promoted and all students missing data *below* the cut-off were retained. We estimate the lower bound by assuming the opposite. We show the upper and lower bound estimates of our primary RDD specification in Appendix Table A8 and our non-economically disadvantaged subgroup analysis in Appendix Table A9. These bounds show that our findings and conclusions would not change even in the worst-case scenario where the bias from differential attrition is as large as possible.

DISCUSSION AND CONCLUSION

This paper explores disparities in the retention outcomes of Michigan third-grade students under Michigan's Read by Grade Three Law, which mandates retention for students who score below an end-of-year assessment threshold unless they receive an exemption. We provide evidence that Black and economically disadvantaged students are retention eligible and retained at higher rates than their peers and that male students are more likely to be retention eligible but less likely to be retained than retention-eligible female students. However, these racial and gender disparities in retention outcomes disappear once we account for differences in test scores –

particularly math scores – as well as in exemption qualifications and district characteristics. But disparities persist by economic disadvantage status.

We further examine disparities in the implementation of the retention policy under RBG3 by economic disadvantage status. Although economically disadvantaged students are more likely to be retained if they score just below the eligibility threshold, non-economically disadvantaged students are no more likely to be retained if they are retention eligible. Differences in GCE eligibility, math performance, district characteristics, and student mobility do not explain any of this disparity.

Our findings support the hypothesis of Licalsi et al. (2019) that wealthier families are better able to advocate for exemptions than their economically disadvantaged peers. However, our findings differ from this earlier work in important ways. First, while more advantaged retention-eligible Florida students were more likely to receive a subjective exemption, this is not the case in Michigan. We find the opposite: economically disadvantaged students were more likely to receive a parent request or portfolio-based exemption than their wealthier peers. Second, Licalsi et al. (2019) find that, while the probability of retention was less affected by retention eligibility for more advantaged students, more advantaged students were still more likely to be retained if they scored just below the eligibility threshold. Our findings are starker. We find that non-economically disadvantaged students are no more likely to be retained if they score just below the eligibility threshold than just above.

The actual impacts of elevated retention rates for some students are not certain. As mentioned earlier, the research on retention outcomes produces conflicting inferences, both in terms of academic and social skills. Students required to repeat third grade under the Law might experience improved math and reading outcomes in the years to come (Greene & Winters, 2004; Hwang & Koedel, 2022) and could feel a higher degree of belonging within their schools (Crego et al., 2009). However, the effects could also be negative and consequential; they might drop out of high school at higher rates than their peers (Eren et al., 2017; Mariano et al., 2018), and they may be suspended at higher rates in the future (Ozek, 2015). It will be critical to continue to follow the outcomes of students who were eligible for retention but promoted through GCEs and who were eventually retained under the Read by Grade Three Law.

Our findings inform policymakers about the importance of equity considerations when implementing ostensibly universal education policies. While Michigan's retention mandate is neutral to race and economic disadvantage status, underlying inequalities in access to effective advocacy can lead to substantial disparities in outcomes. Even though these concerns were known to policymakers and administrators in Michigan, with outreach targeted to historically underserved families (French, 2020), our findings suggest efforts to inform families about their options under the Law were insufficient to eliminate disparities in the implementation of retention.

These results suggest that future research is needed to better understand the exemption distribution process. Our data only provide insights into which exemptions students received. However, students are often eligible for multiple exemptions. Further research into how schools determined which exemption students received would provide further insights into any potential bias in districts' retention decision processes. Additionally, examining exemption requests denied by superintendents could give more insight into potential biases that our present data and methods cannot detect. These avenues for future research will provide deeper insights into how retention requirements that are ostensibly neutral to socioeconomic factors can disproportionately retain historically marginalized student groups and whether this leads to disparate impacts on future outcomes.

REFERENCES

Anderson, G. E., Whipple, A. D., & Jimerson, S. R. (2002). *Achievement and Mental Health Outcomes*. 4.

Babcock, P., & Bedard, K. (2011). The Wages of Failure: New Evidence on School Retention and Long-Run Outcomes. *Education Finance and Policy*, *6*(3), 293–322. https://doi.org/10.1162/EDFP_a_00037

Calonico, S., Cattaneo, M. D., & Farrell, M. H. (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal*, *23*(2), 192–210. https://doi.org/10.1093/ectj/utz022

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2017). Rdrobust: Software for Regression-discontinuity Designs. *The Stata Journal: Promoting Communications on Statistics and Stata*, *17*(2), 372–404. https://doi.org/10.1177/1536867X1701700208

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, *82*(6), 2295–2326. https://doi.org/10.3982/ECTA11757

Crego, A., Gershwin, D., Ikemoto, G. S., Le, V.-N., Marsh, J. A., Naftel, S., Setodji, C. M., Xia, N., McCombs, J. S., Kirby, S. N., & Mariano, L. T. (2009). *Ending Social Promotion Without Leaving Children Behind: The Case of New York City*. RAND Corporation. https://www.rand.org/pubs/monographs/MG894.html

Cummings, A., Strunk, K. O., & De Voto, C. (2021). “A lot of states were doing it”: The development of Michigan’s Read by Grade Three Law. *Journal of Educational Change*. https://doi.org/10.1007/s10833-021-09438-y

Diris, R. (2017). Don’t Hold Back? The Effect of Grade Retention on Student Achievement. *Education Finance and Policy*, *12*(3), 312–341. https://doi.org/10.1162/EDFP_a_00203

Eide, E. R., & Goldhaber, D. D. (2005). Grade retention: What are the costs and benefits? *Journal of Education Finance*, *31*(2), 195–214.

Eren, O., Depew, B., & Barnes, S. (2017). Test-based promotion policies, dropping out, and juvenile crime. *Journal of Public Economics*, *153*(C), 9–31.

Eren, O., Lovenheim, M. F., & Mocan, H. N. (2021). The Effect of Grade Retention on Adult Crime: Evidence from a Test-Based Promotion Policy. *Journal of Labor Economics*, 715836. https://doi.org/10.1086/715836

ExcelinEd. (2021). *Comprehensive early literacy policy: State-by-state analysis of fundamental principles*. ExcelinEd. https://excelined.org/wp-content/uploads/2021/10/ExcelinEd_PolicyToolkit_EarlyLiteracy_StatebyStateAnalysis_2021.pdf

Figlio, D., & Özek, U. (2020). An extra year to learn English? Early grade retention and the human capital development of English learners. *Journal of Public Economics*, *186*, 104184. https://doi.org/10.1016/j.jpubeco.2020.104184

French, R. (2020, January 29). *Gov. Whitmer launching effort to undercut Michigan's third-grade reading law | Bridge Michigan*. https://www.bridgemi.com/talent-education/gov-whitmer-launching-effort-undercut-michigans-third-grade-reading-law

French, R., & Kalakailo, S. (2021, June 15). Michigan schools revolt, won't flunk struggling third-grade readers. *Bridge Michigan*. https://www.bridgemi.com/talent-education/michigan-schools-revolt-wont-flunk-struggling-third-grade-readers

Goos, M., Pipa, J., & Peixoto, F. (2021). Effectiveness of grade retention: A systematic review and meta-analysis. *Educational Research Review*, *34*, 100401. https://doi.org/10.1016/j.edurev.2021.100401

Green, E. L., & Goldstein, D. (2019, October 30). Reading Scores on National Exam Decline in Half the States. *The New York Times*. https://www.nytimes.com/2019/10/30/us/reading-scores-national-exam.html

Greene, J. P., & Winters, M. A. (2004). *An evaluation of Florida's program to end social promotion* (p. 24).

Greene, J. P., & Winters, M. A. (2006). *Getting farther ahead by staying behind: A second-year evaluation of Florida's policy to end social promotion* (No. 49; Civic Report). Center for Civic Innovation at the Manhattan Institute.

Greene, J. P., & Winters, M. A. (2007). Revisiting grade retention: An evaluation of Florida's test-based promotion policy. *Education Finance and Policy*, *2*(4), 319–342.

Greene, J. P., & Winters, M. A. (2009). The effects of exemptions to Florida's test-based promotion policy: Who is retained? *Economics of Education Review*, *28*(1), 135–142. https://doi.org/10.1016/j.econedurev.2008.02.002

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Disruption versus Tiebout improvement: The costs and benefits of switching schools. *Journal of Public Economics*, *88*(9), 1721–1746. https://doi.org/10.1016/S0047-2727(03)00063-X

Hatch, E., & Harbatkin, E. (2021). *COVID-19 and Michigan's Lowest Performing Schools*. EPIC. https://epicedpolicy.org/wp-content/uploads/2021/10/Partnership_COVID_Brief_Oct2021.pdf

Hernandez, D. J. (2011). Double Jeopardy: How Third-Grade Reading Skills and Poverty Influence High School Graduation. In *Annie E. Casey Foundation*. Annie E. https://eric.ed.gov/?id=ED518818

Holmes, C. T., & Matthews, K. M. (1983). *The Effects of Nonpromotion on Elementary and Junior High*. 24.

Hong, G., & Yu, B. (2007). Early-Grade Retention and Children's Reading and Math Learning in Elementary Years. *Educational Evaluation and Policy Analysis*, *29*(4), 239–261. https://doi.org/10.3102/0162373707309073

Hwang, N., & Koedel, C. (2022). Holding Back to Move Forward: The Effects of Retention in the Third Grade on Student Outcomes. In *EdWorkingPapers.com*. Annenberg Institute at Brown University. https://www.edworkingpapers.com/ai22-688

Jacob, B. A. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in Chicago Public Schools. *Journal of Public Economics*, *89*, 761–796.

Jacob, B. A., & Lefgren, L. (2004). Remedial Education and Student Achievement: A Regression-Discontinuity Analysis. *The Review of Economics and Statistics*, *86*(1), 226–244.

Jimerson, S. R. (2001). Meta-analysis of Grade Retention Research: Implications for Practice in the 21st Century. *School Psychology Review*, *30*(3), 420–437. https://doi.org/10.1080/02796015.2001.12086124

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *The Review of Economic Studies*, *76*(3), 1071–1102. https://doi.org/10.1111/j.1467-937X.2009.00536.x

Lee, D. S., & Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, *142*(2), 655–674. https://doi.org/10.1016/j.jeconom.2007.05.003

Lesnick, J., Goerge, R. M., Smithgall, C., & Gwynne, J. (2010). *Reading on Grade Level in Third Grade: How is it related to High School Performance and College Enrollment?* Chapin Hall at the University of Chicago. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.715.5162&rep=rep1&type=pdf

Licalsi, C., Ozek, U., & Figlio, D. (2019). The uneven implementation of universal school policies: Maternal education and Florida's mandatory grade retention policy. *Education Finance and Policy*, *14*(3), 383–413. https://doi.org/10.1162/edfp_a_00252

Livingston, D., & Livingston, S. (2002). Failing Georgia—The Case Against the Ban on Social Promotion. *Education Policy Analysis Archives*, *10*. https://doi.org/10.14507/epaa.v10n49.2002

Lorence, J. (2014). Third-grade retention and reading achievement in Texas: A nine year panel study. *Social Science Research*, *48*, 1–19.

Mariano, L. T., & Martorell, P. (2013, March). *The Academic Effects of Summer Instruction and Retention in New York City*. https://doi.org/10.3102/0162373712454327

Mariano, L. T., Martorell, P., & Berglund, T. (2018). *The Effects of Grade Retention on High School Outcomes: Evidence from New York City Schools*. RAND Corporation. https://www.rand.org/pubs/working_papers/WR1259.html

Michigan State Senate. (2019). *Senate Bill 0633*. http://www.legislature.mi.gov/(S(blwfs4ptdngtotdtrjd3hbv0))/mileg.aspx?page=GetObject&objectname=2019-SB-0633

Ozek, U. (2015). Hold Back To Move Forward? Early Grade Retention And Student Misbehavior. *Education Finance and Policy*, *10*(3), 350–377.

Perrault, P., & Winters, M. (2020). Test-Based Promotion and Student Performance in Florida and Arizona. Issue Brief. In *Manhattan Institute for Policy Research*. Manhattan Institute for Policy Research. https://eric.ed.gov/?id=ED608422

Act No. 306, 4822, State of Michigan 98th Legislature, Regular (2016).

Roderick, M., Jacob, B. A., & Bryk, A. S. (2002). The impact of high-stakes testing in Chicago on student achievement in promotional gate grades. *Educational Evaluation and Policy Analysis*, *24*(4), 333–357. https://doi.org/10.3102/01623737024004333

Roderick, M., Nagaoka, J., & Allensworth, E. (2005). Is the glass half full or mostly empty? Ending social promotion in Chicago. *Yearbook of the National Society for the Study of Education*, *104*(2), 223–259.

Roderick, M., & Nagaoka, Jenny. (2005). Retention under Chicago's high-stakes testing program: Helpful, harmful, or harmless? *Educational Evaluation and Policy Analysis*, *27*(4), 309–340.

Schwerdt, G., West, M. R., & Winters, M. A. (2017). The effects of test-based retention on student outcomes over time: Regression discontinuity evidence from Florida. *Journal of Public Economics*, *152*(C), 154–169.

Silberglitt, B., Jimerson, S., Burns, M., & Appleton, J. (2006). Does the Timing of Grade Retention Make a Difference? Examining the Effects of Early Versus Later Retention.

School Psychology Review, *35*, 134–141. https://doi.org/10.1080/02796015.2006.12088007

Valencia, R. R., & Villareal, B. J. (2005). Texas' Second Wave of High Stakes Testing: Anti-Social Promotion Legislation, Grade Retention, and Advserse Impact on Minorities. In *Leaving Children Behind: How "Texas-style" Accountability Fails Latino Youth*. SUNY Press.

Weiss, S., Stallings, D. T., & Porter, S. (2018). Is Read to Achieve Making the Grade? *Making the Grade*, 40.

Westall, J., Kilbride, T., & Strunk, K. O. (2021). *Read by Grade Three Law Initial Retention Decisions*. EPIC. https://epicedpolicy.org/read-by-grade-three-law-initial-retention-decisions/

Winters, M. A. (2022). The Cost of Retention Under a Test-Based Promotion Policy for Taxpayers and Students. *Educational Evaluation and Policy Analysis*, 01623737221138041. https://doi.org/10.3102/01623737221138041

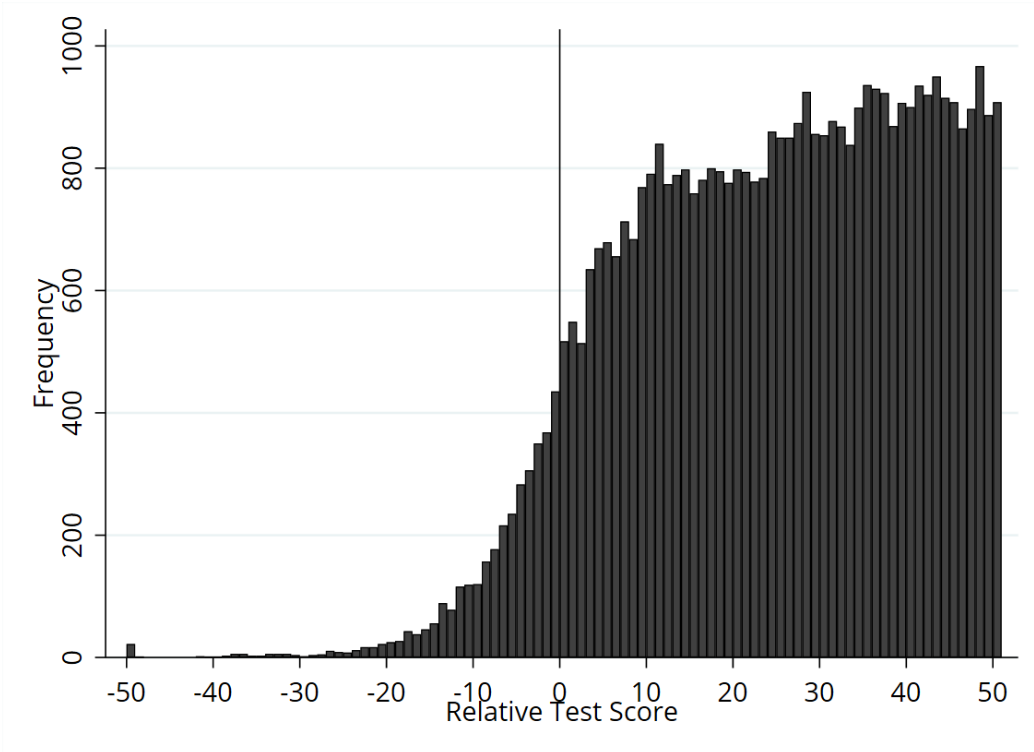
Winters, M. A., & Greene, J. P. (2012). The Medium-Run Effects of Florida's Test-Based Promotion Policy. *Education Finance and Policy*, *7*(3), 305–330. https://doi.org/10.1162/EDFP_a_00069

Wu, W., West, S. G., & Hughes, J. N. (2010). Effect of grade retention in first grade on psychosocial outcomes. *Journal of Educational Psychology*, *102*(1), 135–152. https://doi.org/10.1037/a0016664

Xu, Z., Hannaway, J., & D'Souza, S. (2009). Student Transience in North Carolina: The Effect of School Mobility on Student Outcomes Using Longitudinal Data. Working Paper 22. In *National Center for Analysis of Longitudinal Data in Education Research*. National Center for Analysis of Longitudinal Data in Education Research. https://eric.ed.gov/?id=ED509673

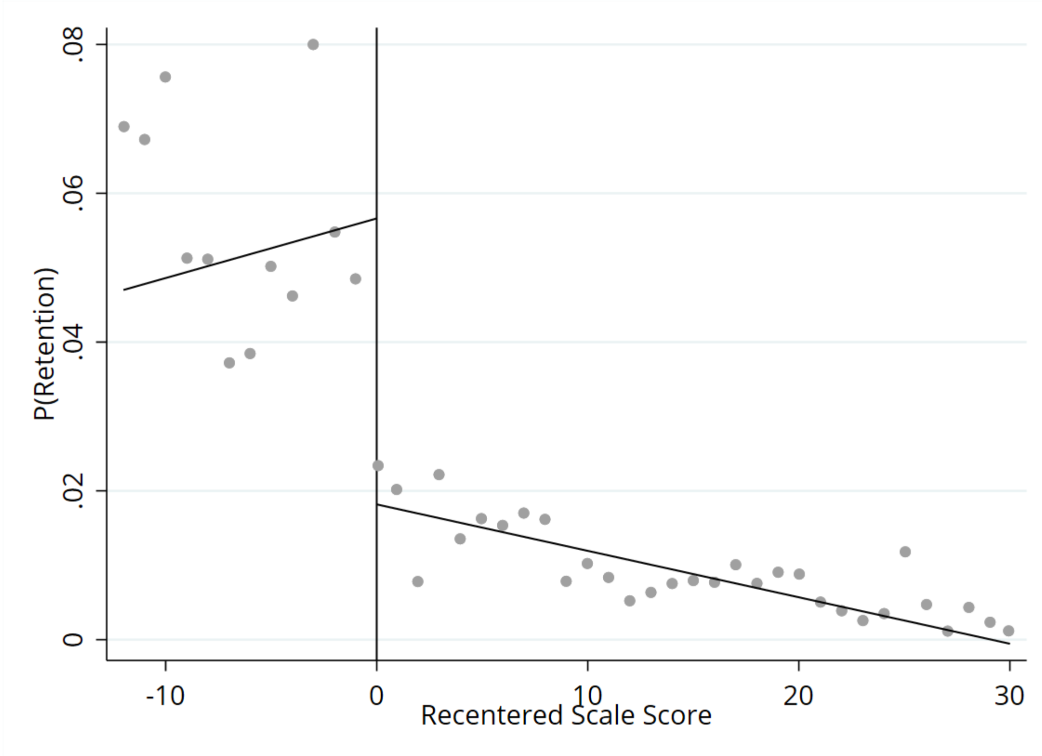
FIGURES AND TABLES

Figure 1: Distribution of ELA M-STEP Scores Near the 1252 Scale Score Cut-off



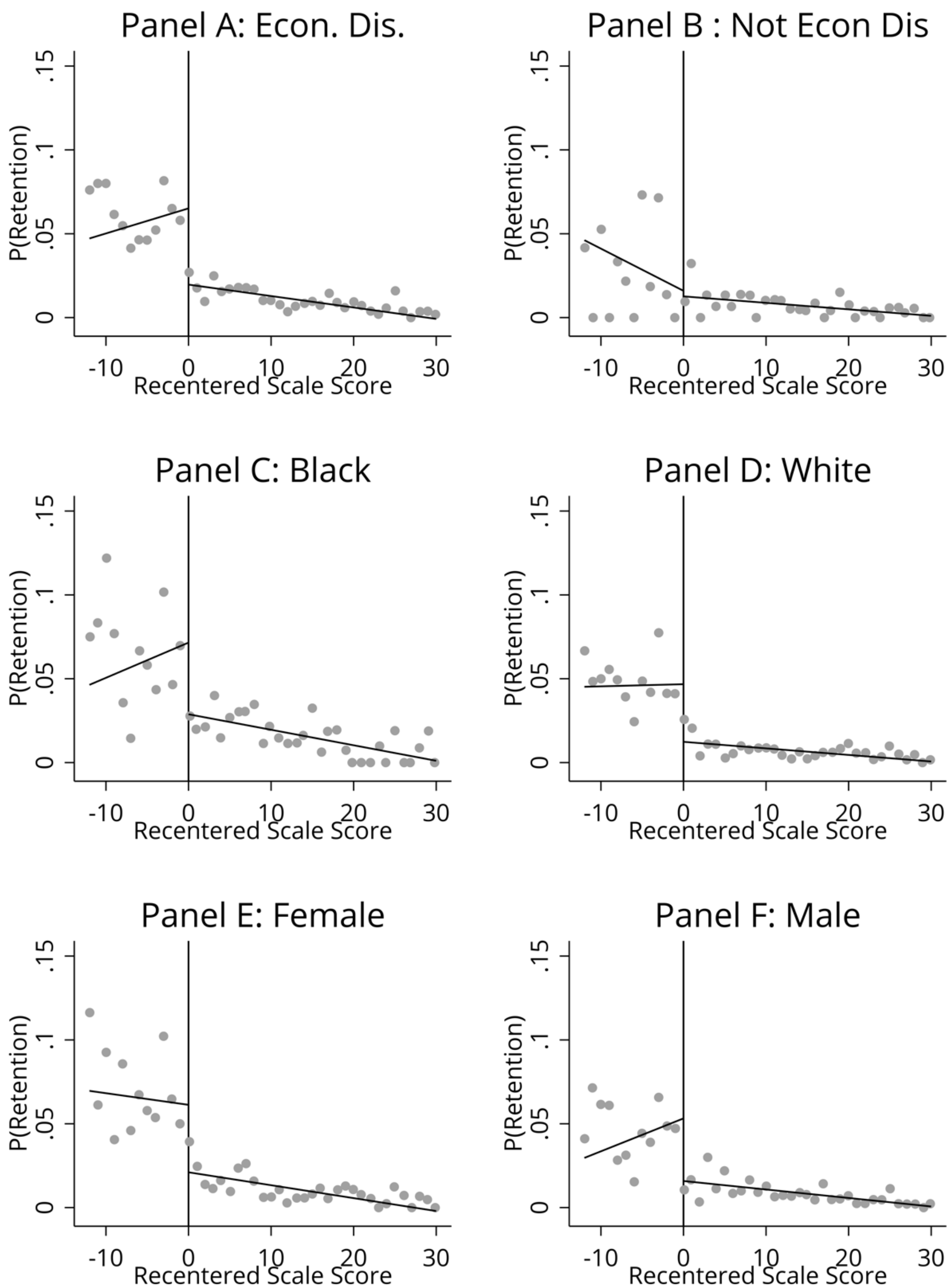
Note: The x-axis contains ELA M-STEP scale scores adjusted such that a 1253 scale score equals 0. The y-axis represents the frequency each scale score occurs in the data.

Figure 2: Retention Eligibility and the Probability of Retention



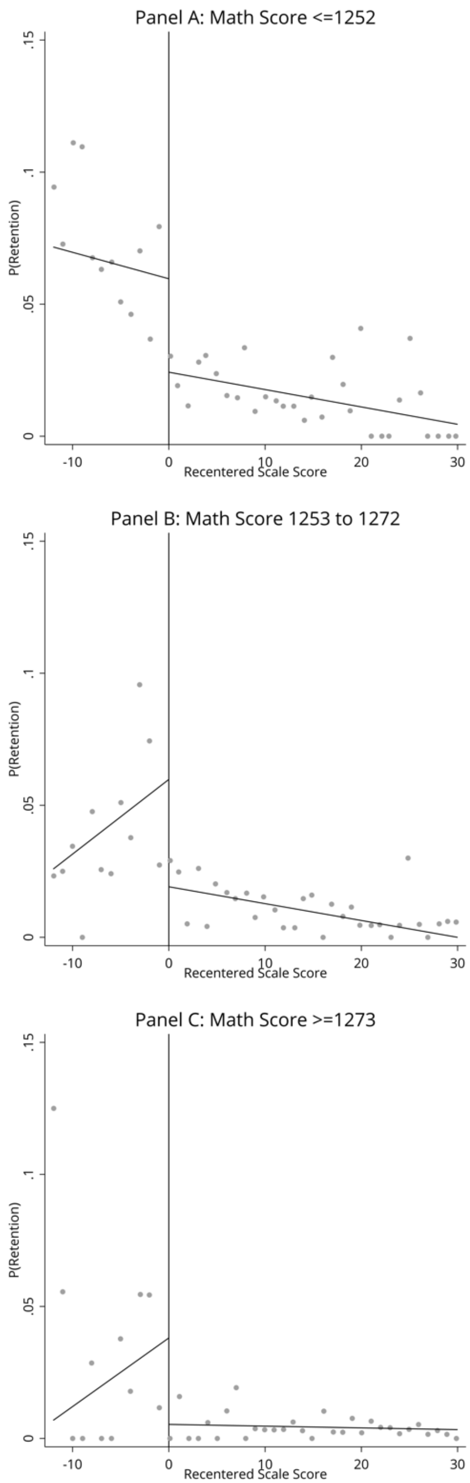
Note: The vertical line indicates the retention-eligibility threshold of 1252 scale score on the third-grade 2020-21 ELA M-STEP. There is one dot for each scale score. The fit lines are from local linear regressions with triangular weights.

Figure 3: Retention Eligibility and the Probability of Retention by Student Characteristics



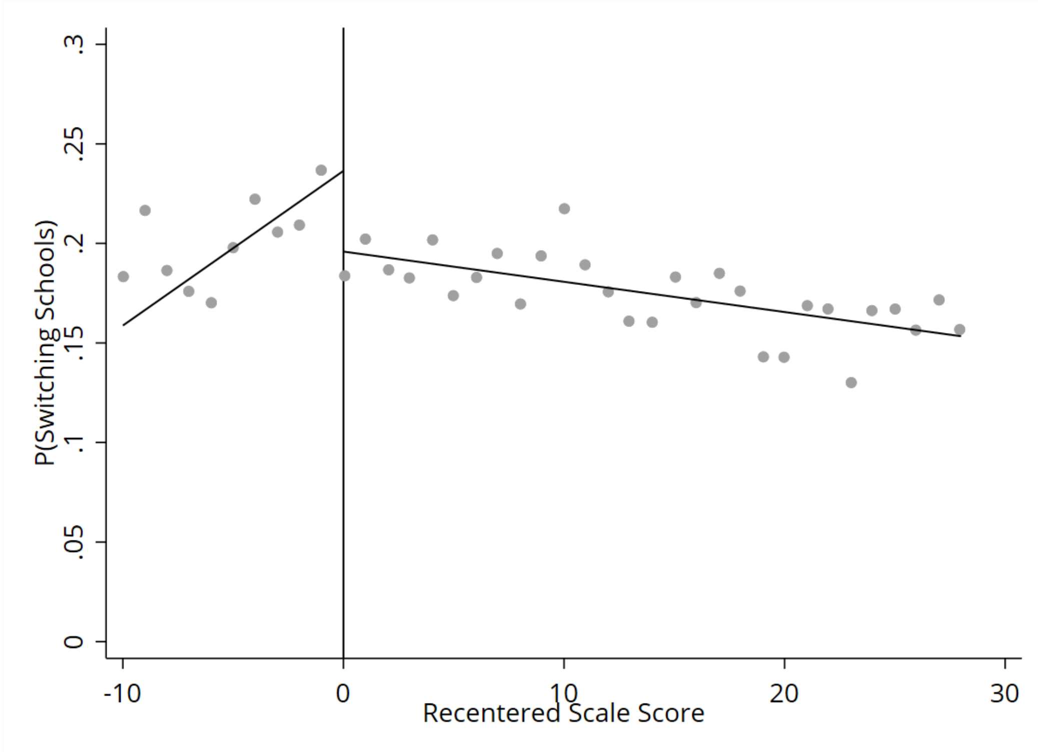
Note: The vertical line indicates the retention-eligibility threshold of 1252 scale score on the third-grade 2020-21 ELA M-STEP. There is one dot for each scale score. The fit lines are from local linear regressions with triangular weights. Each panel is estimated on the given subgroup.

Figure 4: Retention Eligibility and the Probability of Retention by 2020-21 M-STEP Math Performance



Note: The vertical line indicates the retention-eligibility threshold of 1252 scale score on the third-grade 2020-21 ELA M-STEP. There is one dot for each scale score. The fit lines are from local linear regressions with triangular weights. Each panel is estimated on the given subgroup.

Figure 5. Retention Eligibility and Student Mobility



Note: The vertical line indicates the retention-eligibility threshold of 1252 scale score on the third-grade 2020-21 ELA M-STEP. There is one dot for each scale score. The fit lines are from local linear regressions with triangular weights. Each panel is estimated on the given subgroup.

Table 1. Summary Statistics by Retention Outcome

	(1)	(2)	(3)	(4)	(5)	(6)
	All Third Graders	Took ELA M-STEP	Retention Eligible	Promoted w/ GCE	Retained: RBG3	Retained: Other
Number	100,053	72,495	3,440	3,240	175	824
Student Characteristics						
Male	51.5%	51.0%	57.4%	58.0%	48.0%	55.6%
Female	48.5%	49.0%	42.6%	42.0%	52.0%	44.4%
White	63.3%	71.8%	49.8%	50.2%	41.1%	38.7%
Asian	3.7%	3.4%	1.9%	2.0%	1.1%	1.8%
Black	18.6%	11.5%	31.1%	30.6%	40.6%	45.4%
Latino/a	8.5%	7.5%	10.4%	10.5%	9.7%	8.4%
Other	5.9%	5.7%	6.7%	6.6%	7.4%	5.7%
Not Econ Dis.	43.4%	48.4%	17.1%	17.7%	8.0%	17.2%
Econ Dis.	56.6%	51.6%	82.9%	82.3%	92.0%	82.8%
Not English Learner	90.8%	91.6%	87.5%	87.2%	91.4%	93.1%
English Learner	9.2%	8.4%	12.5%	12.8%	8.6%	6.9%
Not Previously Retained	89.2%	89.7%	82.9%	82.4%	92.6%	89.2%
Previously Retained	10.8%	10.3%	17.1%	17.6%	7.4%	10.8%
Not Student w/ Disabilities	84.5%	85.9%	70.4%	69.6%	85.7%	80.7%
Student with Disabilities	15.5%	14.1%	29.6%	30.4%	14.3%	19.3%
Enrolled > 2 Yrs	92.2%	92.4%	90.4%	90.6%	85.1%	81.6%
Enrolled < 2 Yrs	7.8%	7.6%	9.6%	9.4%	14.9%	18.4%
District Characteristics						
Traditional Public School	87.6%	89.1%	78.4%	79.4%	59.4%	64.3%
PSA (Charter)	12.4%	10.9%	21.6%	20.6%	40.6%	35.7%
District ELA Quartile (2019)						
Lowest	17.1%	10.2%	31.2%	30.8%	38.5%	50.3%
Med-Low	18.6%	17.5%	25.5%	25.1%	33.3%	19.1%
Med-High	25.6%	29.1%	23.7%	23.9%	17.2%	15.0%
Highest	38.7%	43.2%	19.6%	20.1%	10.9%	15.6%
City	25.5%	17.0%	31.6%	31.4%	37.4%	41.1%
Suburb/Town	56.9%	61.8%	50.3%	51.0%	37.9%	41.7%
Rural	17.6%	21.2%	18.1%	17.6%	24.7%	17.2%

Note: Each percentage indicates the proportional breakdown of each category by subgroup. For example, 51.5% of the overall third grade population is male and 48.5% is female. The sum of each box equals 100%, and if it does not, it is due to rounding. The last column includes 25 students who were retention-eligible, who received a GCE at some point in time, but were eventually retained in third grade. Due to the reception of a GCE while also still appearing in third grade, we do not consider these students as retained under RBG3 or as promoted with a GCE.

Table 2. Regression Estimates Predicting Retention Amongst Retention-Eligible Students

	Univariate (1)	Student (2)	GCE (3)	ELA Z-Score (4)	Math Z-Score (5)	Dis. Cov. (6)
Student Characteristics						
Female	0.020*	0.018*	0.014+	0.014+	0.005	0.005
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Asian	-0.021	-0.012	0.001	0.002	0.005	0.011
	(0.021)	(0.022)	(0.020)	(0.020)	(0.021)	(0.020)
Black	0.022*	0.019+	0.012	0.012	0.001	-0.005
	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Latino/a	-0.004	0.002	0.006	0.006	-0.000	-0.001
	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)
Other race(s)	0.006	0.010	0.008	0.008	-0.002	-0.003
	(0.016)	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)
Econ. Disad.	0.033***	0.027**	0.028***	0.028***	0.023**	0.018*
	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Eng. Learner	-0.018+		-0.021*	-0.021*	-0.019+	-0.019
	(0.010)		(0.010)	(0.011)	(0.011)	(0.012)
Prev. Retained	-0.035***		-0.032***	-0.032***	-0.033***	-0.036***
	(0.008)		(0.008)	(0.008)	(0.008)	(0.008)
Stud. w/ Disab.	-0.037***		-0.031***	-0.031***	-0.041***	-0.038***
	(0.007)		(0.007)	(0.007)	(0.008)	(0.008)
Enroll < 2 Yrs	0.031+		0.026	0.026	0.026	0.023
	(0.016)		(0.016)	(0.016)	(0.016)	(0.016)
ELA Z-Score	-0.007			-0.004	-0.006	-0.005
	(0.013)			(0.013)	(0.014)	(0.014)
Math Z-Score	-0.023***				-0.025***	-0.026***
	(0.005)				(0.006)	(0.006)
District Characteristics						
PSA (Charter)	0.057***					0.047**
	(0.014)					(0.016)
M-STEP Partic	0.001***					0.001**
	(0.000)					(0.000)
Urban	0.014					0.005
	(0.012)					(0.013)
Rural	0.023+					0.031**
	(0.012)					(0.012)
Enroll. (100s)	-0.002*					-0.000
	(0.001)					(0.001)
2019 ELA Perf.	-0.036**					-0.021
	(0.013)					(0.013)
Observations		3440	3440	3440	3290	3270
R-squared		0.006	0.016	0.016	0.022	0.040

*Note: The sample includes only retention-eligible 2020-21 third-grade students. The dependent variable is an indicator for being retained by your district under RBG3. In Column (1), each cell represents a separate univariate regression of Equation (1). No overall observation count is given because it varies between 3,440 and 3,270, depending in the regression. In Columns (2) through (6), each column is a single regression of Equation 2. Heteroskedasticity-robust standard errors clustered at the district-level are in parentheses. + p<0.1 * p<0.05 ** p<0.01 *** p<0.001*

Table 3. Regression Discontinuity Estimates of the Impact of Retention Eligibility on the Probability of Retention

	(1)	(2)	(3)	(4)
	RD Estimate - Impact of Retention Eligibility	Confidence Interval	N - Below Cut-off	N - Above Cut-off
Overall	0.038***	[0.022,0.060]	2,749	22,527
STUDENT CHARACTERISTICS				
Econ. Dis.	0.045***	[0.033,0.069]	2,274	15,894
Not Econ. Dis.	0.003	[-0.038,0.022]	475	6,632
White	0.034***	[0.019,0.054]	1,403	13,914
Black	0.043***	[0.021,0.078]	828	4,474
Female	0.040**	[0.013,0.069]	1,184	10,394
Male	0.037***	[0.025,0.058]	1,565	12,132
MATH PERFORMANCE				
Math Score <=1252	0.035*	[0.005,0.076]	1,195	4,269
Math Score 1253 to 1272	0.041	[-0.016,0.098]	930	6,934
Math Score >1272	0.032	[-0.012,0.085]	529	10,875
GCE QUALIFICATION				
Student with Disabilities	0.021*	[0.003,0.048]	805	5,328
Not Student with Disabilities	0.045**	[0.014,0.078]	1,944	17,198
English Learner	0.017	[-0.039,0.059]	355	2,671
Not English Learner	0.042***	[0.030,0.062]	2,394	19,855
Enrolled in District < 2 Years	0.006	[-0.055,0.068]	263	1,986
Not Enrolled in District < 2 Years	0.042***	[0.028,0.061]	2,486	20,541
Previously Retained	-0.001	[-0.037,0.013]	480	3,265
Not Previously Retained	0.048***	[0.035,0.071]	2,269	19,262
DISTRICT CHARACTERISTICS				
Charter School	0.031*	[0.010,0.075]	590	3,516
Traditional Public School	0.040***	[0.021,0.060]	2,159	19,011
District's Quartile for ELA in 2019				
Lowest	0.032+	[-0.003,0.076]	832	3,986
Mid-Low	0.060***	[0.039,0.078]	684	5,233
Mid-High	0.040***	[0.025,0.072]	652	6,444
High	0.014	[-0.032,0.051]	539	6,657
Urban	0.035***	[0.014,0.052]	834	4,780
Suburban and Town	0.039***	[0.018,0.057]	1,361	9,671
Rural	0.036***	[0.018,0.059]	2,259	17,796

*Note: Regression discontinuity estimates of the impact of retention-eligibility on the probability of retention. Estimates are from local linear models with triangular kernels. The running variable is students' ELA M-STEP scale score. Students are retention eligible if they score 1252 or below. Each row represents a separate model estimate on a given subpopulation. In each model the bandwidth is 12 scale score points below and 30 scale score points above the cut-off. Column (1) shows the LATE estimate of the impact of being just below the retention-eligibility threshold on the probability of retention. Column (2) shows robust confidence intervals. Columns (3) and (4) show the number of observations within the bandwidth below and above the cut-off. +p<0.1 *p<0.05 **p<0.01 ***p<0.001*

Table 4. Percentages of Student Subgroups Eligible for GCEs and Considered Economically Disadvantaged

	(1) SPED	(2) EL	(3) Enrolled < 2 Years	(4) Previously Retained
All Third Grade Students	15,548	9,230	7,848	10,772
White	15.7%	5.3%	7.2%	10.3%
Asian	6.8%	53.8%	9.7%	4.6%
Black	15.9%	2.1%	8.7%	12.6%
Latino/a	16.2%	39.9%	8.6%	12.9%
Other	17.5%	2.0%	9.6%	10.5%
Not Econ. Dis.	11.9%	5.3%	6.5%	6.9%
Econ. Dis.	18.3%	12.2%	8.8%	13.8%
Retention Eligible Students	1,018	431	331	587
White	34.6%	11.0%	8.9%	17.8%
Asian	20.9%	91.0%	9.0%	19.4%
Black	20.5%	2.1%	10.4%	16.3%
Latino/a	29.2%	42.9%	9.7%	16.2%
Other	37.7%	1.7%	11.7%	16.5%
Not Econ. Dis.	38.9%	7.3%	8.2%	13.1%
Econ. Dis.	27.7%	13.6%	9.9%	17.9%

Note: Each percentage indicates the percentage of each student subgroup who is a student with disabilities, an English Learner, a student enrolled in their district for less than two years, a student that was previously retained, or a student that is economically disadvantaged. The first four categories align with GCEs. The intersectionality between race and economic status and the differences between all third grade students and retention-eligible students is shown in the last column.

Table 5. Regression Discontinuity Estimates by Economic Status

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Non-Economically Disadvantaged				Economically Disadvantaged			
	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF
Overall	0.003	[-0.038,0.022]	475	6,632	0.045***	[0.033,0.069]	2,274	15,894
Student Characteristics								
White	0.000	[-0.036,0.013]	342	5,204	0.045***	[0.034,0.070]	1,061	8,710
Black	0.003	[-0.056,0.057]	61	524	0.046***	[0.026,0.081]	767	3,950
Female	-0.008+	[-0.079,0.001]	181	2,887	0.048**	[0.018,0.087]	1,003	7,507
Male	0.013	[-0.029,0.053]	294	3,745	0.044***	[0.036,0.064]	1,271	8,387
Math Performance								
Math Score <=1252	0.006	[-0.057,0.037]	136	657	0.039*	[0.009,0.087]	1,059	3,611
Math Score 1253 to 1272	-0.021**	[-0.047,-0.009]	163	1,556	0.053	[-0.013,0.121]	767	5,378
Math Score >1272	0.013	[-0.032,0.052]	167	4,351	0.040+	[-0.009,0.104]	362	6,524
GCE Qualification								
Student with Disabilities	0.013	[-0.037,0.047]	178	1,506	0.024+	[-0.005,0.072]	627	3,822
Not Student with Disabilities	-0.002	[-0.044,0.013]	297	5,126	0.052**	[0.019,0.092]	1,647	12,072
English Learner	0.015	[-0.098,0.101]	39	450	0.018	[-0.034,0.057]	316	2,221
Not English Learner	0.002	[-0.035,0.016]	436	6,182	0.050***	[0.042,0.074]	1,958	13,673
Enrolled in District < 2 Yrs	0.023	[-0.131,0.166]	33	458	0.003	[-0.050,0.066]	230	1,527
Not Enrolled in District < 2 Yrs	0.003	[-0.034,0.016]	442	6,174	0.050***	[0.039,0.072]	2,044	14,367
Previously Retained	-0.002	[-0.007,0.002]	64	680	-0.002	[-0.043,0.014]	416	2,585
Not Previously Retained	0.004	[-0.043,0.025]	411	5,952	0.057***	[0.048,0.084]	1,858	13,309

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Non-Economically Disadvantaged				Economically Disadvantaged			
	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF
District Characteristics								
Charter School	-0.081	[-0.193,0.027]	41	469	0.037**	[0.013,0.082]	549	3,047
Traditional Public School	0.010	[-0.036,0.030]	434	6,163	0.047***	[0.033,0.070]	1,725	12,847
District ELA Quartile 2018-19								
Lowest	-0.014	[-0.049,0.018]	56	351	0.035+	[-0.002,0.083]	776	3,635
Mid-Low	0.022	[-0.073,0.085]	80	1,034	0.065***	[0.044,0.089]	604	4,198
Mid-High	-0.022**	[-0.073,-0.011]	146	2,001	0.059***	[0.046,0.101]	506	4,443
High	0.022	[-0.031,0.060]	189	3,205	0.008	[-0.049,0.062]	350	3,452
Urban	0.007	[-0.033,0.039]	85	827	0.039***	[0.018,0.057]	749	3,953
Suburban and Town	-0.016	[-0.070,-0.001]	194	2,269	0.048***	[0.024,0.076]	1,167	7,401
Rural	0.003	[-0.038,0.022]	475	6,632	0.040***	[0.026,0.063]	1,887	12,569
Student Mobility								
Stayed in Same District	0.001	[-0.027,0.012]	438	6300	0.045***	[0.037,0.066]	1,989	14,327
Switched Districts	0.018	[-0.221,0.142]	37	320	0.044	[-0.013,0.095]	280	1,513

*Note: See Table 3 note for detail. Column (1) – (4) show models estimated for only non-economically disadvantaged students. Columns (5) – (8) are estimated with only economically disadvantaged students. +p<0.1 *p<0.05 **p<0.01 ***p<0.001*

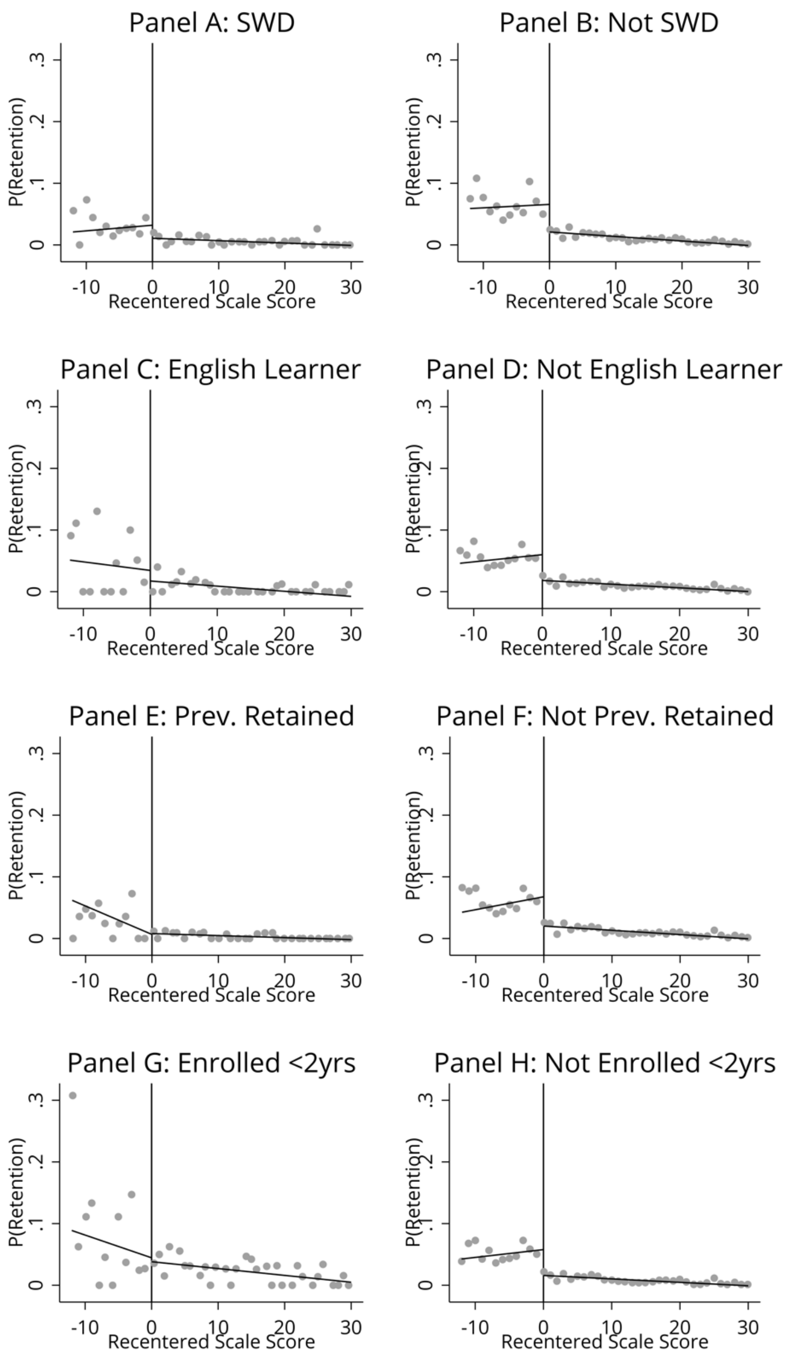
Table 6: Regression Discontinuity Estimates of the Impact of Retention Eligibility on Student Mobility

	(1)	(2)	(3)	(4)
	RD Estimate - Impact of Retention Eligibility	Confidence Interval	N - Below Cut-off	N - Above Cut-off
Overall	0.033***	[0.024,0.060]	2,511	22,527
STUDENT CHARACTERISTICS				
Econ. Dis.	0.032***	[0.026,0.058]	2,074	15,894
Not Econ. Dis.	0.032+	[-0.003,0.081]	437	6,632
White	0.015*	[0.003,0.038]	1,281	13,914
Black	0.039*	[0.005,0.103]	763	4,474
Female	0.034**	[0.013,0.054]	1,081	10,394
Male	0.033***	[0.027,0.072]	1,430	12,132
MATH PERFORMANCE				
Math Score <=1252	0.043*	[0.011,0.078]	1,077	4,269
Math Score 1253 to 1272	0.001	[-0.034,0.047]	861	6,934
Math Score >1272	0.045***	[0.035,0.116]	488	10,875
GCE QUALIFICATION				
Student with Disabilities	0.022*	[0.005,0.056]	719	5,328
Not Student with Disabilities	0.037***	[0.025,0.069]	1,792	17,198
English Learner	0.013	[-0.061,0.082]	328	2,671
Not English Learner	0.037***	[0.030,0.066]	2,183	19,855
Enrolled in District < 2 Years	0.035	[-0.080,0.205]	239	1,990
Not Enrolled in District < 2 Years	0.034***	[0.026,0.056]	2,272	20,537
Previously Retained	0.000	[-0.045,0.064]	431	3,265
Not Previously Retained	0.040***	[0.032,0.066]	2,080	19,262
DISTRICT CHARACTERISTICS				
Charter School	0.109*	[0.032,0.254]	539	3,516
Traditional Public School	0.013+	[-0.002,0.035]	1,972	19,011
DISTRICT'S QUARTILE FOR ELA IN 2019				
Lowest	0.032	[-0.024,0.107]	757	3,986
Mid-Low	0.012	[-0.042,0.055]	638	5,233
Mid-High	0.049+	[-0.001,0.124]	596	6,444
High	0.056***	[0.043,0.122]	484	6,657
Urban	0.073***	[0.062,0.125]	765	4,780
Suburban and Town	0.065***	[0.066,0.101]	1,244	9,671
Rural	0.031**	[0.015,0.063]	2,064	17,796

*Note: Regression discontinuity estimates of the impact of retention-eligibility on the probability of switching Districts. Estimates are from local linear models with triangular kernels. The running variable is students' ELA M-STEP scale score. Students are retention eligible if they score 1252 or below. Each row represents a separate model estimate on a given subpopulation. In each model the bandwidth is 12 scale score points below and 30 scale score points above the cut-off. Column (1) shows the LATE estimate of the impact of being just below the retention-eligibility threshold on the probability of switching schools. Column (2) shows robust confidence intervals. Columns (3) and (4) show the number of observations within the bandwidth below and above the cut-off. +p<0.1 *p<0.05 **p<0.01 ***p<0.001*

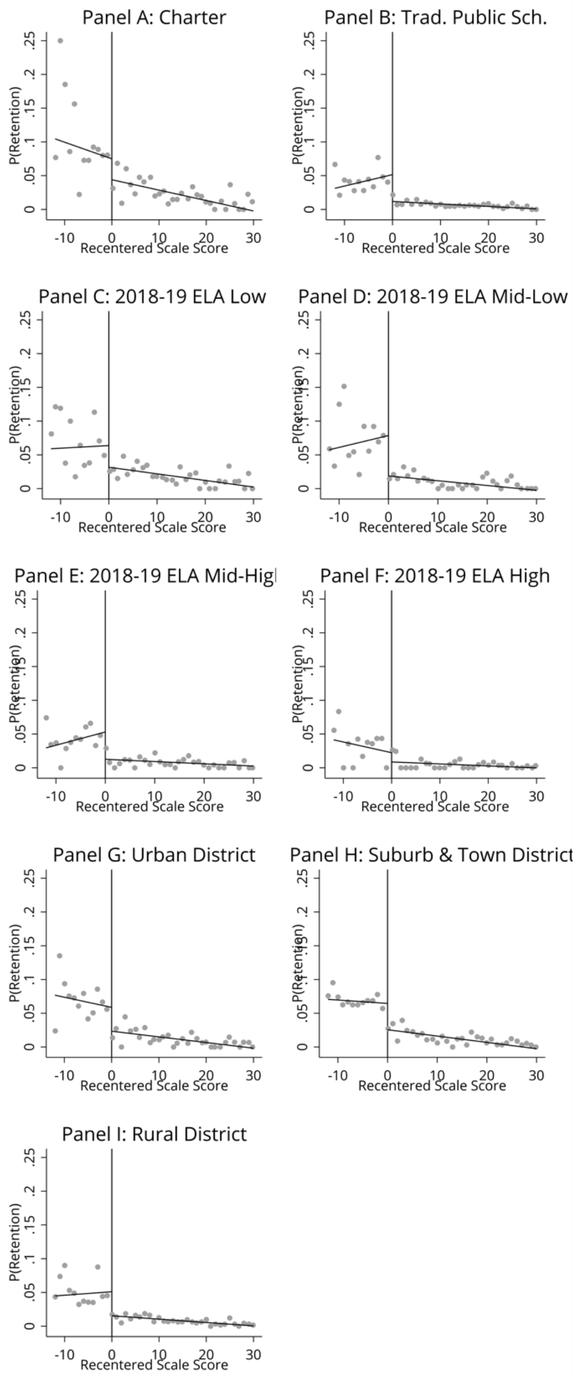
ONLINE APPENDIX

Appendix Figure A1. Retention Eligibility and the Probability of Retention by GCE Qualification



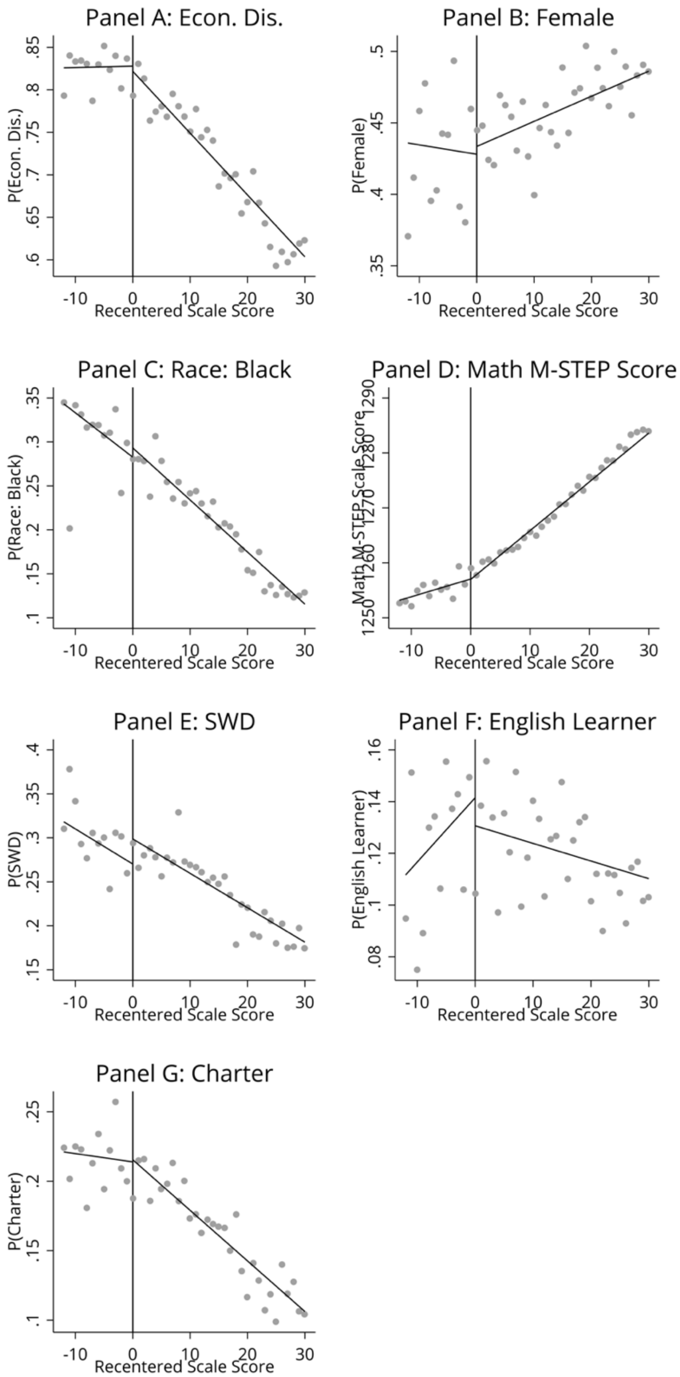
Note: The vertical line indicates the retention-eligibility threshold of 1252 scale score on the third-grade 2020-21 ELA M-STEP. There is one dot for each scale score. The fit lines are from local linear regressions with triangular weights. Each panel is estimated on the given subgroup.

Appendix Figure A2: Retention Eligibility and the Probability of Retention by District Characteristics



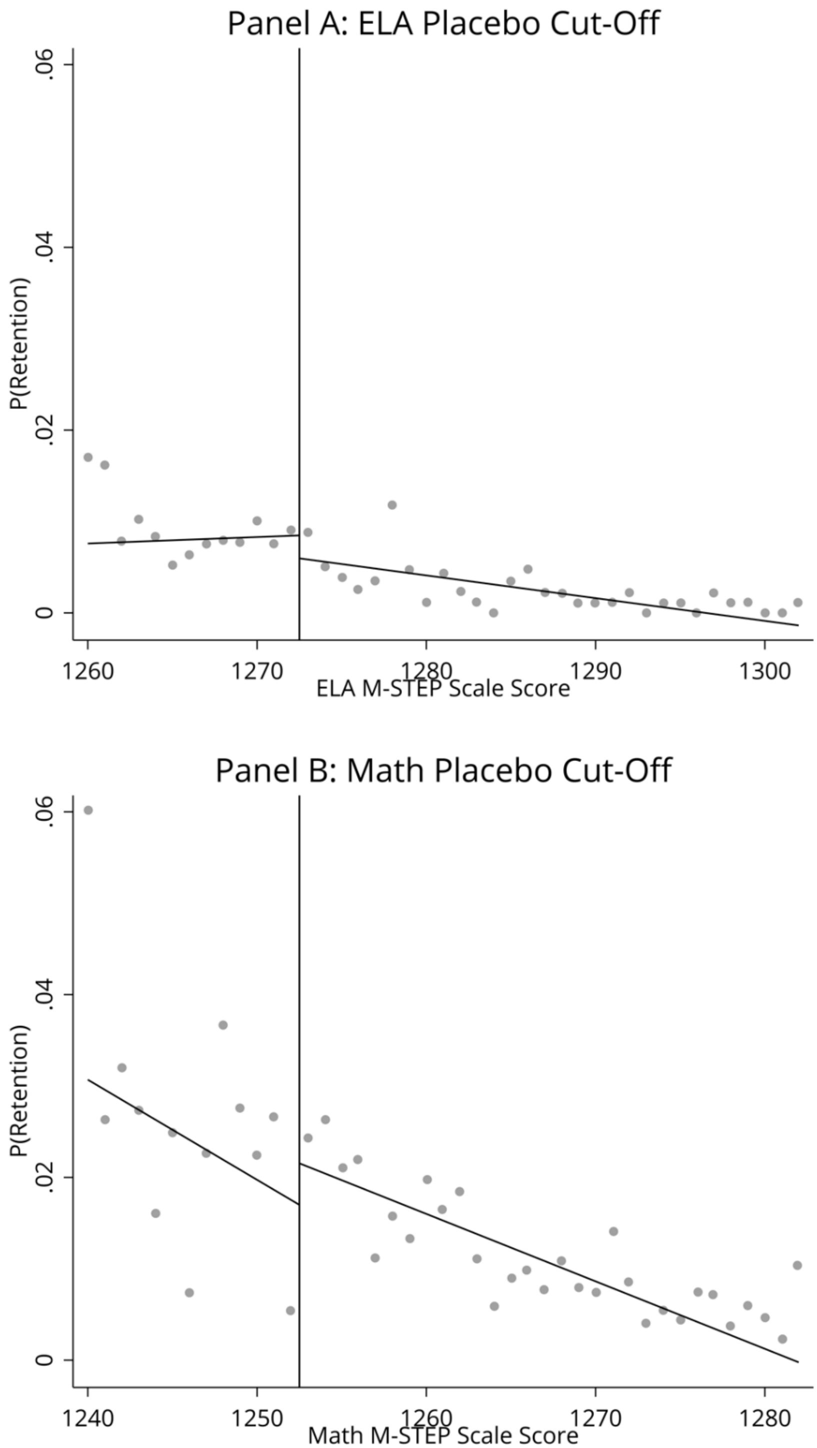
Note: The vertical line indicates the retention-eligibility threshold of 1252 scale score on the third-grade 2020-21 ELA M-STEP. There is one dot for each scale score. The fit lines are from local linear regressions with triangular weights. Each panel is estimated on the given subgroup.

Appendix Figure A3. Falsification Tests



Note: The vertical line indicates the retention-eligibility threshold of 1252 scale score on the third-grade 2020-21 ELA M-STEP. There is one dot for each scale score. The fit lines are from local linear regressions with triangular weights. Each panel is estimated on a different outcome given by the panel title.

Appendix Figure A4. Placebo Cut-Offs



Note: The vertical line indicates the placebo thresholds. In Panel A, the threshold is a 1272 scale score on the third-grade ELA M-STEP. In Panel B, the threshold is a 1252 scale score on the third-grade Math M-STEP. There is one dot for each scale score. The fit lines are from local linear regressions with triangular weights.

Appendix Table A1. Regression Discontinuity Estimates of the Impact of Retention Eligibility on the Probability of Retention (Quadratic Functional Form)				
	(1)	(2)	(3)	(4)
	RD Estimate - Impact of Retention Eligibility	Confidence Interval	N - Below Cut-off	N - Above Cut-off
Overall	0.035**	[0.011,0.052]	2749	22527
STUDENT CHARACTERISTICS				
Econ. Dis.	0.047***	[0.022,0.067]	2274	15894
Not Econ. Dis.	-0.028*	[-0.074,-0.004]	475	6632
White	0.026*	[0.004,0.038]	1403	13914
Black	0.052***	[0.024,0.079]	828	4474
Female	0.025	[-0.010,0.044]	1184	10394
Male	0.043***	[0.020,0.065]	1565	12132
MATH PERFORMANCE				
Math Score <=1252	0.058***	[0.042,0.095]	1195	4269
Math Score 1253 to 1272	0.013	[-0.055,0.041]	930	6934
Math Score >1272	0.018	[-0.045,0.061]	529	10875
GCE QUALIFICATION				
Student with Disabilities	0.033***	[0.016,0.062]	805	5328
Not Student with Disabilities	0.035	[-0.007,0.061]	1944	17198
English Learner	0.005	[-0.062,0.070]	355	2671
Not English Learner	0.039***	[0.020,0.051]	2394	19855
Enrolled in District < 2 Years	-0.027	[-0.100,0.031]	263	1986
Not Enrolled in District < 2 Years	0.041***	[0.019,0.057]	2486	20541
Previously Retained	-0.020*	[-0.051,-0.005]	480	3265
Not Previously Retained	0.048***	[0.024,0.068]	2269	19262
DISTRICT CHARACTERISTICS				
Charter School	0.055**	[0.014,0.099]	590	3516
Traditional Public School	0.030*	[0.003,0.047]	2159	19011
District's Quartile for ELA in 2019				
Lowest	0.038	[-0.015,0.090]	832	3986
Mid-Low	0.064***	[0.034,0.097]	684	5233
Mid-High	0.026	[-0.007,0.038]	652	6444
High	-0.011	[-0.064,0.020]	539	6657
Urban	0.042**	[0.011,0.073]	834	4780
Suburban and Town	0.031*	[0.006,0.050]	1361	9671
Rural	0.039**	[0.013,0.062]	2259	17796

*Note: Regression discontinuity estimates of the impact of retention-eligibility on the probability of retention. Estimates are from local linear models with triangular kernels. The running variable is students' ELA M-STEP scale score. Students are retention eligible if they score 1252 or below. Each row represents a separate model estimate on a given subpopulation. In each model the bandwidth is 12 scale score points below and 30 scale score points above the cut-off. Column (1) shows the LATE estimate of the impact of being just below the retention-eligibility threshold on the probability of retention. Column (2) shows robust confidence intervals. Columns (3) and (4) show the number of observations within the bandwidth below and above the cut-off. +p<0.1 *p<0.05 **p<0.01 ***p<0.001*

Appendix Table A2. Regression Discontinuity Estimates by Economic Status (Quadratic Functional Form)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Non-Economically Disadvantaged				Economically Disadvantaged			
	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF
Overall	-0.028*	[-0.074,-0.004]	475	6632	0.047***	[0.022,0.067]	2274	15894
Student Characteristics								
White	-0.031^{+}	[-0.088,0.007]	342	5204	0.044***	[0.018,0.060]	1061	8710
Black	0.024	[-0.074,0.133]	61	524	0.054***	[0.029,0.077]	767	3950
Female	-0.066^{+}	[-0.178,0.011]	181	2887	0.038	[-0.009,0.070]	1003	7507
Male	-0.008	[-0.064,0.031]	294	3745	0.055***	[0.037,0.077]	1271	8387
Math Performance								
Math Score <=1252	-0.033	[-0.107,0.016]	136	657	0.070***	[0.056,0.111]	1059	3611
Math Score 1253 to 1272	-0.036^{+}	[-0.081,-0.001]	163	1556	0.022	[-0.058,0.059]	767	5378
Math Score >1272	-0.006	[-0.065,0.044]	167	4351	0.027	[-0.042,0.071]	362	6524
GCE Qualification								
Student with Disabilities	-0.021^{+}	[-0.078,0.001]	178	1506	0.051***	[0.030,0.099]	627	3822
Not Student with Disabilities	-0.031	[-0.099,0.024]	297	5126	0.044^{+}	[-0.003,0.076]	1647	12072
English Learner	-0.041	[-0.177,0.031]	39	450	0.012	[-0.050,0.078]	316	2221
Not English Learner	-0.027*	[-0.071,-0.002]	436	6182	0.053***	[0.032,0.069]	1958	13673
Enrolled in District < 2 Yrs	-0.042	[-0.238,0.078]	33	458	-0.023	[-0.094,0.036]	230	1527
Not Enrolled in District < 2 Yrs	-0.026	[-0.068,-0.002]	442	6174	0.055***	[0.031,0.076]	2044	14367
Previously Retained	0.001	[-0.002,0.007]	64	680	-0.023*	[-0.058,-0.005]	416	2585

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Non-Economically Disadvantaged				Economically Disadvantaged			
	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF
Not Previously Retained	-0.031*	[-0.084,-0.002]	411	5952	0.064***	[0.039,0.089]	1858	13309
District Characteristics								
Charter School	0.006	[-0.184,0.293]	41	469	0.053*	[0.012,0.092]	549	3047
Traditional Public School	-0.031**	[-0.081,-0.012]	434	6163	0.045***	[0.019,0.067]	1725	12847
District ELA Quartile 2018-19								
Lowest	-0.025	[-0.075,0.021]	56	351	0.044	[-0.013,0.100]	776	3635
Mid-Low	-0.044	[-0.164,0.025]	80	1034	0.078***	[0.052,0.114]	604	4198
Mid-High	-0.042+	[-0.098,0.008]	146	2001	0.045**	[0.009,0.059]	506	4443
High	-0.022	[-0.071,-0.005]	189	3205	-0.005	[-0.076,0.048]	350	3452
Urban	-0.014	[-0.073,0.021]	85	827	0.050**	[0.019,0.085]	749	3953
Suburban and Town	-0.052+	[-0.120,0.003]	194	2269	0.045**	[0.011,0.072]	1167	7401
Rural	-0.010	[-0.053,0.012]	372	5227	0.049***	[0.025,0.073]	1887	12569

*Note: Note: Regression discontinuity estimates of the impact of retention-eligibility on the probability of retention. Estimates are from local linear models with triangular kernels. The running variable is students' ELA M-STEP scale score. Students are retention eligible if they score 1252 or below. Each row represents a separate model estimate on a given subpopulation. In each model the bandwidth is 12 scale score points below and 30 scale score points above the cut-off. Column (1) – (4) show models estimated for only non-economically disadvantaged students. Columns (5) – (8) are estimated with only economically disadvantaged students. +p<0.1 *p<0.05 **p<0.01 ***p<0.001*

Appendix Table A3. Regression Discontinuity Estimates of the Impact of Retention Eligibility on the Probability of Retention (+/- 12 Scale Score Point Bandwidth)

	(1)	(2)	(3)	(4)
	RD Estimate - Impact of Retention Eligibility	Confidence Interval	N - Below Cut-off	N - Above Cut-off
Overall	0.037***	[0.021,0.060]	2749	7959
STUDENT CHARACTERISTICS				
Econ. Dis.	0.044***	[0.031,0.069]	2274	6213
Not Econ. Dis.	0.003	[-0.039,0.026]	475	1746
White	0.029***	[0.013,0.049]	1403	4466
Black	0.046***	[0.026,0.084]	828	2041
Female	0.035*	[0.007,0.065]	1184	3507
Male	0.039***	[0.028,0.062]	1565	4452
MATH PERFORMANCE				
Math Score <=1252	0.034*	[0.005,0.075]	1195	2483
Math Score 1253 to 1272	0.038	[-0.018,0.098]	930	2853
Math Score >1272	0.034	[-0.010,0.087]	529	2422
GCE QUALIFICATION				
Student with Disabilities	0.019+	[-0.001,0.046]	805	2220
Not Student with Disabilities	0.043**	[0.014,0.080]	1944	5739
English Learner	0.018	[-0.038,0.066]	355	1009
Not English Learner	0.040*	[0.028,0.062]	2394	6950
Enrolled in District < 2 Years	0.001	[-0.061,0.063]	263	765
Not Enrolled in District < 2 Years	0.041***	[0.028,0.062]	2486	7194
Previously Retained	-0.002	[-0.037,0.016]	480	1293
Not Previously Retained	0.046***	[0.034,0.071]	2269	6666
DISTRICT CHARACTERISTICS				
Charter School	0.031*	[0.008,0.086]	590	1531
Traditional Public School	0.038***	[0.018,0.060]	2159	6428
DISTRICT'S QUARTILE FOR ELA IN 2019				
Lowest	0.036*	[0.004,0.084]	832	1900
Mid-Low	0.058***	[0.041,0.079]	684	1978
Mid-High	0.039*	[0.021,0.072]	652	2067
High	0.006	[-0.040,0.043]	539	1922
Urban	0.037***	[0.016,0.058]	834	1966
Suburban and Town	0.034**	[0.013,0.056]	1361	3625
Rural	0.037***	[0.020,0.063]	2259	6369

*Note: Regression discontinuity estimates of the impact of retention-eligibility on the probability of retention. Estimates are from local linear models with triangular kernels. The running variable is students' ELA M-STEP scale score. Students are retention eligible if they score 1252 or below. Each row represents a separate model estimate on a given subpopulation. In each model the bandwidth is 12 scale score points below and 12 scale score points above the cut-off. Column (1) shows the LATE estimate of the impact of being just below the retention-eligibility threshold on the probability of retention. Column (2) shows robust confidence intervals. Columns (3) and (4) show the number of observations within the bandwidth below and above the cut-off. +p<0.1 *p<0.05 **p<0.01 ***p<0.001*

Appendix Table A4. Regression Discontinuity Estimates by Economic Status (+/- 12 Scale Score Point Bandwidth)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Non-Economically Disadvantaged				Economically Disadvantaged			
	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF
Overall	0.002	[-0.034,0.017]	475	1746	0.044^{***}	[0.023,0.077]	2274	6213
Student Characteristics								
White	-0.002	[-0.044,0.019]	342	1305	0.039^{*}	[0.007,0.081]	1061	3161
Black	0.016	[-0.042,0.069]	61	188	0.049^{*}	[0.008,0.110]	767	1853
Female	-0.011+	[-0.080,0.007]	181	676	0.042^{*}	[0.005,0.089]	1003	2831
Male	0.011	[-0.017,0.041]	294	1070	0.045^{**}	[0.018,0.088]	1271	3382
Math Performance								
Math Score <=1252	0.012	[-0.043,0.037]	136	327	0.038^{*}	[0.001,0.092]	1059	2156
Math Score 1253 to 1272	-0.027^{+}	[-0.072,0.002]	163	560	0.051^{*}	[0.013,0.097]	767	2293
Math Score >1272	0.011	[-0.044,0.059]	167	839	0.042^{*}	[0.008,0.091]	362	1583
GCE Qualification								
Student with Disabilities	0.018	[-0.007,0.034]	178	526	0.020	[-0.021,0.074]	627	1694
Not Student with Disabilities	-0.007	[-0.058,0.019]	297	1220	0.051^{***}	[0.025,0.089]	1647	4519
English Learner	0.002	[-0.108,0.094]	39	130	0.020	[-0.033,0.067]	316	879
Not English Learner	0.001	[-0.035,0.018]	436	1616	0.048^{***}	[0.026,0.086]	1958	5334
Enrolled in District < 2 Yrs	0.016	[-0.121,0.146]	33	129	-0.001	[-0.071,0.077]	230	636
Not Enrolled in District < 2 Yrs	0.002	[-0.034,0.017]	442	1617	0.049^{***}	[0.027,0.084]	2044	5577
Previously Retained	0.002	[-0.003,0.009]	64	189	-0.003	[-0.038,0.012]	416	1104
Not Previously Retained	0.002	[-0.040,0.020]	411	1557	0.055^{***}	[0.033,0.097]	1858	5109
District Characteristics								
Charter School	-0.092^{+}	[-0.221,0.018]	41	147	0.039	[-0.011,0.121]	549	1384
Traditional Public School	0.009	[-0.027,0.023]	434	1599	0.045^{***}	[0.020,0.077]	1725	4829

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Non-Economically Disadvantaged				Economically Disadvantaged			
	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF
District ELA Quartile 2018-19								
Lowest	-0.031	[-0.108,0.033]	56	129	0.040^{+}	[0.000,0.098]	776	1771
Mid-Low	0.020	[-0.074,0.096]	80	305	0.062^{*}	[0.010,0.121]	604	1673
Mid-High	-0.017	[-0.083,0.018]	146	522	0.055^{*}	[0.014,0.121]	506	1545
High	0.018	[-0.019,0.039]	189	777	0.000	[-0.049,0.043]	350	1145
Urban	0.014	[-0.011,0.038]	85	245	0.040	[-0.008,0.090]	749	1721
Suburban and Town	-0.021^{+}	[-0.085,0.005]	194	662	0.043^{*}	[0.007,0.087]	1167	2963
Rural	0.015	[-0.012,0.034]	372	1342	0.041^{**}	[0.019,0.075]	1887	5027

*Note: Note: Regression discontinuity estimates of the impact of retention-eligibility on the probability of retention. Estimates are from local linear models with triangular kernels. The running variable is students' ELA M-STEP scale score. Students are retention eligible if they score 1252 or below. Each row represents a separate model estimate on a given subpopulation. In each model the bandwidth is 12 scale score points below and 12 scale score points above the cut-off. Column (1) – (4) show models estimated for only non-economically disadvantaged students. Columns (5) – (8) are estimated with only economically disadvantaged students. +p<0.1 *p<0.05 **p<0.01 ***p<0.001*

Appendix Table A5. Regression Discontinuity Estimates of the Impact of Retention Eligibility on the Probability of Retention (+/- 7 Scale Score Point Bandwidth)

	(1)	(2)	(3)	(4)
	RD Estimate - Impact of Retention Eligibility	Confidence Interval	N - Below Cut-off	N - Above Cut-off
Overall	0.033**	[0.012,0.047]	1964	4192
STUDENT CHARACTERISTICS				
Econ. Dis.	0.044***	[0.023,0.063]	1629	3299
Not Econ. Dis.	-0.019*	[-0.070,-0.003]	335	893
White	0.021*	[0.001,0.030]	1026	2270
Black	0.045***	[0.025,0.067]	586	1140
Female	0.023	[-0.010,0.039]	850	1872
Male	0.042***	[0.025,0.060]	1114	2320
MATH PERFORMANCE				
Math Score <=1252	0.050***	[0.038,0.089]	835	1430
Math Score 1253 to 1272	0.019	[-0.047,0.051]	669	1525
Math Score >1272	0.024	[-0.034,0.059]	393	1119
GCE QUALIFICATION				
Student with Disabilities	0.025***	[0.016,0.052]	559	1161
Not Student with Disabilities	0.036+	[-0.002,0.057]	1405	3031
English Learner	0.015	[-0.053,0.067]	263	526
Not English Learner	0.036***	[0.020,0.047]	1701	3666
Enrolled in District < 2 Years	-0.020	[-0.093,0.016]	184	411
Not Enrolled in District < 2 Years	0.039***	[0.019,0.054]	1780	3781
Previously Retained	-0.015*	[-0.047,-0.002]	328	682
Not Previously Retained	0.044***	[0.025,0.062]	1636	3510
DISTRICT CHARACTERISTICS				
Charter School	0.039*	[0.010,0.080]	427	830
Traditional Public School	0.032*	[0.006,0.047]	1537	3362
DISTRICT'S QUARTILE FOR ELA IN 2019				
Lowest	0.039+	[-0.006,0.083]	597	1053
Mid-Low	0.065***	[0.047,0.096]	489	1053
Mid-High	0.024	[-0.009,0.036]	468	1047
High	-0.013+	[-0.066,0.005]	381	990
Urban	0.045***	[0.023,0.073]	591	1071
Suburban and Town	0.033**	[0.012,0.049]	978	1929
Rural	0.037***	[0.015,0.057]	1603	3370

*Note: Regression discontinuity estimates of the impact of retention-eligibility on the probability of retention. Estimates are from local linear models with triangular kernels. The running variable is students' ELA M-STEP scale score. Students are retention eligible if they score 1252 or below. Each row represents a separate model estimate on a given subpopulation. In each model the bandwidth is 7 scale score points above and 7 scale score points below the cut-off. Column (1) shows the LATE estimate of the impact of being just below the retention-eligibility threshold on the probability of retention. Column (2) shows robust confidence intervals. Columns (3) and (4) show the number of observations within the bandwidth below and above the cut-off. +p<0.1 *p<0.05 **p<0.01 ***p<0.001*

Appendix Table A6. Regression Discontinuity Estimates by Economic Status (+/- 7 Scale Score Point Bandwidth)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Non-Economically Disadvantaged				Economically Disadvantaged			
	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF
Overall	-0.019*	[-0.069,-0.004]	335	893	0.044*	[0.007,0.079]	1629	3299
Student Characteristics								
White	-0.021+	[-0.078,0.003]	243	652	0.034	[-0.018,0.081]	783	1618
Black	0.007	[-0.028,0.037]	46	102	0.048	[-0.020,0.119]	540	1038
Female	-0.055*	[-0.157,-0.011]	118	346	0.034	[-0.025,0.082]	732	1526
Male	0.001	[-0.039,0.013]	217	547	0.052*	[0.008,0.104]	897	1773
Math Performance								
Math Score <=1252	-0.013	[-0.089,0.021]	92	180	0.059*	[0.014,0.139]	743	1250
Math Score 1253 to 1272	-0.046+	[-0.111,0.001]	113	308	0.033	[-0.039,0.067]	556	1217
Math Score >1272	0.003	[-0.060,0.047]	123	389	0.032	[-0.030,0.069]	270	730
GCE Qualification								
Student with Disabilities	-0.001	[-0.058,0.011]	126	267	0.036	[-0.016,0.123]	433	894
Not Student with Disabilities	-0.028+	[-0.091,0.005]	209	626	0.046+	[-0.003,0.080]	1196	2405
English Learner	-0.024	[-0.149,0.044]	27	65	0.021	[-0.044,0.075]	236	461
Not English Learner	-0.019*	[-0.070,-0.001]	308	828	0.048*	[0.008,0.089]	1393	2838
Enrolled in District < 2 Yrs	-0.027	[-0.222,0.072]	20	70	-0.020	[-0.135,0.063]	164	341
Not Enrolled in District < 2 Yrs	-0.018*	[-0.067,-0.001]	315	823	0.051**	[0.014,0.090]	1465	2958
Previously Retained	0.000	[0.000,0.000]	39	93	-0.016+	[-0.057,0.003]	289	589
Not Previously Retained	-0.020*	[-0.077,-0.003]	296	800	0.058**	[0.017,0.103]	1340	2710
District Characteristics								
Charter School	-0.084	[-0.177,0.078]	34	74	0.044	[-0.043,0.136]	393	756
Traditional Public School	-0.015*	[-0.071,-0.003]	301	819	0.043*	[0.004,0.080]	1236	2543

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Non-Economically Disadvantaged				Economically Disadvantaged			
	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF
District ELA Quartile 2018-19								
Lowest	-0.041	[-0.130,0.042]	40	72	0.044	[-0.018,0.107]	557	981
Mid-Low	-0.006	[-0.141,0.067]	56	166	0.073*	[0.007,0.163]	433	887
Mid-High	-0.029	[-0.105,0.026]	106	242	0.040	[-0.046,0.103]	362	805
High	-0.016*	[-0.086,0.000]	131	408	-0.012	[-0.074,0.027]	250	582
Urban	0.002	[-0.046,0.022]	60	125	0.051+	[-0.007,0.120]	531	946
Suburban and Town	-0.045*	[-0.122,-0.001]	135	335	0.046+	[-0.007,0.099]	843	1594
Rural	-0.001	[-0.043,0.012]	263	687	0.044*	[0.009,0.084]	1340	2683

*Note: Regression discontinuity estimates of the impact of retention-eligibility on the probability of retention. Estimates are from local linear models with triangular kernels. The running variable is students' ELA M-STEP scale score. Students are retention eligible if they score 1252 or below. Each row represents a separate model estimate on a given subpopulation. In each model the bandwidth is 7 scale score points above and 7 scale score points below the cut-off. Column (1) – (4) show models estimated for only non-economically disadvantaged students. Columns (5) – (8) are estimated with only economically disadvantaged students. +p<0.1 *p<0.05 **p<0.01 ***p<0.001*

Appendix Table A7: Regression Discontinuity Estimates of Differential Attrition				
	RD Estimate - Impact of Retention Eligibility	Confidence Interval	N - Below Cut-off	N - Above Cut-off
Overall	-0.002	[-0.008,0.000]	2,766	22,657
Student Characteristics				
Econ. Dis.	-0.001	[-0.007,0.003]	2,289	15,990
Not Econ. Dis.	-0.008*	[-0.018,-0.002]	477	6,666
White	-0.007***	[-0.016,-0.005]	1,409	13,969
Black	0.005	[-0.017,0.027]	836	4,527
Female	-0.005*	[-0.018,-0.001]	1,193	10,454
Male	0.000	[-0.004,0.007]	1,573	12,202
Math Performance				
Math Score <=1252	-0.007	[-0.013,0.002]	1,199	4,304
Math Score 1253 to 1272	0.006	[-0.011,0.020]	941	6,971
Math Score >1272	-0.006**	[-0.017,-0.003]	531	10,929
GCE Qualification				
Student with Disabilities	-0.008*	[-0.012,0.000]	806	5,356
Not Student with Disabilities	0.000	[-0.008,0.003]	1,960	17,300
English Learner	-0.017	[-0.040,-0.007]	358	2,695
Not English Learner	0.000	[-0.005,0.004]	2,408	19,961
Enrolled in District < 2 Years	-0.002	[-0.038,0.025]	268	2,021
Not Enrolled in District < 2 Years	-0.002	[-0.009,0.002]	2,498	20,636
Previously Retained	0.001	[-0.015,0.013]	483	3,277
Not Previously Retained	-0.003	[-0.010,0.002]	2,283	19,380
District Characteristics				
Charter School	-0.006	[-0.032,0.018]	596	3,590
Traditional Public School	-0.001	[-0.008,0.003]	2,170	19,067
District's Quartile for ELA in 2019				
Lowest	0.001	[-0.012,0.013]	839	4,031
Mid-Low	-0.008**	[-0.019,-0.004]	689	5,262
Mid-High	-0.001	[-0.013,0.009]	655	6,474
High	-0.001	[-0.014,0.012]	541	6,676
Rural	-0.005	[-0.019,0.002]	840	4,826
Suburban and Town	-0.005*	[-0.014,-0.002]	1,369	9,742
Rural	-0.001	[-0.007,0.003]	2,274	17,907

*Note: Regression discontinuity estimates of the impact of retention-eligibility on the probability of attrition from the data by Fall 2022. Estimates are from local linear models with triangular kernels. The running variable is students' ELA M-STEP scale score. Students are retention eligible if they score 1252 or below. Each row represents a separate model estimate on a given subpopulation. In each model the bandwidth is 12 scale score points below and 30 scale score points above the cut-off. Column (1) shows the LATE estimate of the impact of being just below the retention-eligibility threshold on the probability of switching schools. Column (2) shows robust confidence intervals. Columns (3) and (4) show the number of observations within the bandwidth below and above the cut-off. +p<0.1 *p<0.05 **p<0.01 ***p<0.001*

Appendix Table A8: Attrition Correction, Lower and Upper Bound Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Lower Bound Estimates				Upper Bound Estimates (Worst-Case Scenario)			
	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF
Overall	0.031**	[0.015,0.052]	2,766	22,657	0.044***	[0.027,0.064]	2,766	22,657
Student Characteristics								
Econ. Dis.	0.037***	[0.025,0.061]	2,289	15,990	0.052***	[0.038,0.075]	2,289	15,990
Not Econ. Dis.	-0.002	[-0.043,0.016]	477	6,666	0.003	[-0.041,0.020]	477	6,666
White	0.030**	[0.014,0.051]	1,409	13,969	0.032***	[0.016,0.049]	1,409	13,969
Black	0.029*	[0.006,0.064]	836	4,527	0.060***	[0.039,0.096]	836	4,527
Female	0.032*	[0.005,0.060]	1,193	10,454	0.043*	[0.009,0.070]	1,193	10,454
Male	0.030***	[0.018,0.051]	1,573	12,202	0.045***	[0.037,0.064]	1,573	12,202
Math Performance								
Math Score <=1252	0.027	[-0.002,0.069]	1,199	4,304	0.038*	[0.004,0.082]	1,199	4,304
Math Score 1253 to 1272	0.034	[-0.022,0.090]	941	6,971	0.052	[-0.016,0.117]	941	6,971
Math Score >1272	0.026	[-0.019,0.079]	531	10,929	0.034	[-0.013,0.084]	531	10,929
GCE Qualification								
Student with Disabilities	0.015	[-0.004,0.043]	806	5,356	0.020*	[0.004,0.049]	806	5,356
Not Student with Disabilities	0.037*	[0.006,0.070]	1,960	17,300	0.052**	[0.019,0.085]	1,960	17,300
English Learner	0.005	[-0.053,0.050]	358	2,695	0.013	[-0.041,0.040]	358	2,695
Not English Learner	0.035***	[0.023,0.055]	2,408	19,961	0.048*	[0.036,0.069]	2,408	19,961
Enrolled in District < 2 Yrs	-0.009	[-0.072,0.053]	268	2,021	0.020	[-0.045,0.080]	268	2,021
Not Enrolled in District < 2 Yrs	0.036***	[0.022,0.055]	2,498	20,636	0.046***	[0.029,0.068]	2,498	20,636
Previously Retained	-0.007	[-0.043,0.007]	483	3,277	0.005	[-0.030,0.017]	483	3,277
Not Previously Retained	0.040***	[0.027,0.063]	2,283	19,380	0.053***	[0.036,0.078]	2,283	19,380
District Characteristics								
Charter School	0.013	[-0.008,0.057]	596	3,590	0.043**	[0.014,0.093]	596	3,590
Traditional Public School	0.035**	[0.016,0.056]	2,170	19,067	0.043***	[0.027,0.060]	2,170	19,067
District ELA Quartile 2018-19								
Lowest	0.019	[-0.015,0.062]	839	4,031	0.045*	[0.006,0.094]	839	4,031
Mid-Low	0.056***	[0.036,0.074]	689	5,262	0.058***	[0.031,0.075]	689	5,262
Mid-High	0.034**	[0.018,0.065]	655	6,474	0.046***	[0.029,0.078]	655	6,474
High	0.009	[-0.037,0.046]	541	6,676	0.017	[-0.036,0.065]	541	6,676
Urban	0.026*	[0.005,0.043]	840	4,826	0.040**	[0.012,0.056]	840	4,826
Suburban and Town	0.032**	[0.012,0.050]	1,369	9,742	0.041**	[0.014,0.059]	1,369	9,742
Rural	0.027**	[0.010,0.051]	2,274	17,907	0.043***	[0.025,0.064]	2,274	17,907

*Note: Regression discontinuity estimates of the impact of retention-eligibility on the probability of retention. Each row represents a separate model estimate on a given subpopulation. +p<0.1 *p<0.05 **p<0.01 ***p<0.001*

Appendix Table A9: Attrition Correction, Lower and Upper Bound Estimates, Non-Economically Disadvantaged Students								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Lower Bound Estimates				Upper Bound Estimates (Worst-Case Scenario)			
	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF
Overall	-0.002	[-0.043,0.016]	477	6,666	0.003	[-0.041,0.020]	477	6,666
Student Characteristics								
White	-0.004	[-0.042,0.009]	343	5,223	-0.001	[-0.036,0.011]	343	5,223
Black	-0.002	[-0.059,0.058]	62	531	0.003	[-0.076,0.045]	62	531
Female	-0.015*	[-0.087,-0.008]	183	2,904	-0.011+	[-0.094,0.002]	183	2,904
Male	0.009	[-0.032,0.049]	294	3,762	0.014	[-0.027,0.055]	294	3,762
Math Performance								
Math Score <=1252	0.004	[-0.057,0.038]	137	661	0.007	[-0.067,0.035]	137	661
Math Score 1253 to 1272	-0.027***	[-0.056,-0.014]	164	1,565	-0.026**	[-0.054,-0.010]	164	1,565
Math Score >1272	0.008	[-0.038,0.046]	167	4,371	0.015	[-0.029,0.055]	167	4,371
GCE Qualification								
Student with Disabilities	0.007	[-0.044,0.041]	178	1,515	0.013	[-0.037,0.047]	178	1,515
Not Student with Disabilities	-0.007	[-0.048,0.008]	299	5,151	-0.004	[-0.051,0.012]	299	5,151
English Learner	-0.003	[-0.115,0.083]	39	455	0.016	[-0.097,0.101]	39	455
Not English Learner	-0.002	[-0.040,0.012]	438	6,211	0.001	[-0.039,0.015]	438	6,211
Enrolled in District < 2 Yrs	-0.004	[-0.159,0.137]	33	468	0.024	[-0.130,0.167]	33	468
Not Enrolled in District < 2 Yrs	-0.001	[-0.037,0.013]	444	6,198	0.002	[-0.037,0.015]	444	6,198
Previously Retained	-0.001	[-0.006,0.005]	64	681	-0.002	[-0.007,0.002]	64	681
Not Previously Retained	-0.002	[-0.049,0.019]	413	5,985	0.003	[-0.047,0.023]	413	5,985

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Lower Bound Estimates				Upper Bound Estimates (Worst-Case Scenario)			
	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF	RD ESTIMATE - IMPACT OF RETENTION ELIGIBILITY	CONFIDENCE INTERVAL	N - BELOW CUT-OFF	N - ABOVE CUT-OFF
District Characteristics								
Charter School	-0.103*	[-0.216,0.000]	41	482	-0.080	[-0.191,0.027]	41	482
Traditional Public School	0.006	[-0.040,0.027]	436	6,184	0.009	[-0.039,0.028]	436	6,184
District ELA Quartile 2018-19								
Lowest	-0.014	[-0.049,0.018]	57	351	-0.010	[-0.078,0.017]	57	351
Mid-Low	0.019	[-0.076,0.083]	80	1,041	0.028	[-0.065,0.090]	80	1,041
Mid-High	-0.026**	[-0.080,-0.012]	146	2,011	-0.022**	[-0.073,-0.011]	146	2,011
High	0.015	[-0.039,0.053]	190	3,219	0.017	[-0.034,0.054]	190	3,219
Urban	-0.007	[-0.051,0.031]	86	837	0.007	[-0.050,0.034]	86	837
Suburban and Town	-0.024+	[-0.077,-0.009]	195	2,289	-0.015*	[-0.078,-0.001]	195	2,289
Rural	0.006	[-0.034,0.034]	374	5,253	0.010	[-0.032,0.036]	374	5,253

*Note: Regression discontinuity estimates of the impact of retention-eligibility on the probability of retention. Each row represents a separate model estimate on a given subpopulation. +p<0.1 *p<0.05 **p<0.01 ***p<0.001*

ENDNOTES

[1] We examine the potential for attrition of this nature in response to retention-eligibility in Section 6.3 below. Moreover, only 21 retention-eligible students are missing fall 2021-22 grade information, limiting the impact of any potential bias.

[2] Economically disadvantaged status is defined in Michigan as students who are eligible for free- or reduced-price lunch, are in households receiving food (SNAP) or cash (TANF) assistance, are homeless, are migrant, and/or are in foster care.

[3] The mapping of item responses to scale scores on the M-STEP is determined by the test vendor after all students have completed the exam. It is not possible for students or educators to know a priori what a student's scale score will be based solely on the proportion of correct item responses.

[4] When examining race, we focus on differences between Black and White students because this was the most significant disparity by racial group shown above.

[5] Appendix Figure A1 contains graphical evidence of these results.

[6] Appendix Figure A2 shows these results graphically.

[7] Notably, retention eligible students enrolled in charter schools are more likely to move than their peers in traditional public schools. There are other substantive differences between groups of students, but none are statistically significant.