The Effects of Early Grade Retention on Student Outcomes over Time: Regression Discontinuity Evidence from Florida *

Guido Schwerdt†
Ifo Institute, IZA and CESifo

Martin R. West‡
Harvard Graduate School of Education

Preliminary and incomplete. Please do not cite or circulate without authors' permission.

June 18, 2012

Abstract

Several school districts and states require students to meet basic performance standards in core academic subjects in order to be promoted to the next grade. We exploit a discontinuity in the probability of third grade retention under Florida's test-based promotion policy to study the effects of retention on future student outcomes over time. Although conventional OLS estimates indicate negative effects of grade retention on achievement, our regression discontinuity analysis confirms large positive effects on achievement and a reduced probability of retention in subsequent years. The initial achievement gains from retention fade out over time, however, and are statistically insignificant by eighth grade.

JEL Codes: H52, I21, I28

Keywords: Educational production, public schools, grade retention

---

* We are grateful to the Florida Department of Education for providing the primary dataset for this study and to seminar participants at CESifo, Harvard University, and Mathematica Policy Research for helpful comments. Any errors are our own.

† Ifo Institute, Poschingerstraße 5, 81679 München, Germany. E-Mail: schwerdt@ifo.de,
Tel.: +49 89-9224-1310 Fax.: +49 89-9224-14606

‡ Appian Way, Gutman 454, Cambridge, MA 02138, USA. Email: martin_west@gse.harvard.edu,
Tel: +1 617-496-4803 Fax: +1 617-496-3095

1. Introduction

Should students who fail to meet basic performance standards in core academic subjects be retained in the same grade? Recent estimates from the U.S. Department of Education's Office for Civil Rights indicate that 2.3 percent of American public school students are retained each year, with the incidence of retention concentrated among traditionally disadvantaged minorities (Adams et al. 2012). Retention is costly in terms of additional per pupil spending and foregone earnings, if students spend more time in full-time public education as a result of being retained. Yet consensus is lacking as to whether retention yields benefits for students that could offset these costs and, if so, under what conditions.

In recent years, several large school districts and states have enacted policies requiring that students meet explicit performance targets in certain grades in order to be promoted. Since the 2002-03 school year, for example, Florida has required that schools retain third grade students failing to demonstrate basic proficiency on the state reading test unless the student is eligible for one of a specified set of exemptions. Arizona, Indiana, and Oklahoma enacted policies modeled after Florida's in 2011, and similar bills were introduced in the legislatures of at least four other states. Although these policies aim primarily to provide incentives for teachers, parents, and students to ensure that students are meeting performance expectations, they are also expected to increase retention rates in targeted grades. The number of Florida third graders retained jumped sharply the year the policy was introduced to 21,799 (13.5 percent), up from 4,819 (2.8 percent) the previous year.

Proponents of test-based promotion policies contend that students who are retained as a result will benefit from additional instruction and an improved matching of ability to curriculum and peers. Critics, meanwhile, warn that retained students may be harmed by stigmatization, reduced expectations, and challenges of adjusting to a new peer group. In fact, a large literature in education suggests that retained students achieve at lower levels, complete fewer years of school, and have worse socio-emotional outcomes than observably similar students who are promoted.[1] Because retention decisions typically reflect student

[1] Influential studies in this area include Jimerson (1999) and Jimerson et al. (2000, 2002), and McCoy and Reynolds (1999). A survey of 47 empirical studies conducted between by Holmes (1989) concluded that retained students performed 0.19 to 0.31 standard deviations worse on various measures of academic achievement than similar students who were not retained. A meta-analysis of post-1990 studies by Allen et al. (2009) concluded that, although most studies indicated negative effects of retention, a subset with more rigorous designs yielded more positive effects.

characteristics unobserved by the researcher, however, these studies are likely to suffer from severe selection issues.

In this paper, we use statewide administrative data covering Florida public schools in grade 3 to 9 to study the causal effect of third grade retention on future student outcomes up to 6 years later. The Florida database has two key advantages for studying the consequences of grade retention. First, the state's test-based promotion policy generates a clear discontinuity in the probability of retention that allows us to overcome the selection issues that plague most existing research on this topic. By comparing our regression discontinuity estimates to those obtained with the conventional selection-on-observables estimation strategies most commonly used to study the effects of retention, we provide direct evidence that the latter do not deal adequately with selection issues. And although our preferred estimates of the effects of retention are local to students with third grade reading test scores at the cutoff for promotion, variation in these students' math achievement enables us to provide suggestive evidence on the extent to which the effects we identify are generalizable to a broader population. Second, the database contains vertically scaled test scores in reading and math that make it possible to compare the achievement of students tested in different grades. The ability to make this comparison is essential because the counter-factual condition for retained students is to have been immediately promoted to the next grade. While often reported in the literature, same-grade comparisons conflate any effect of retention on achievement with the effect of being a year older at the time the relevant test is administered.

It is important to note that the Florida policy also required that retained students be given the opportunity to attend a summer reading camp prior to the following school year and that they be assigned to a "high-performing" teacher and receive intensive reading interventions during that year. Although our estimates will capture the combined effect of retention and these additional measures, such requirements are typical of test-based promotion policies currently in use and under consideration.

Our analysis confirms that Florida students at the promotion cutoff who were retained in third grade experienced substantial short-term gains in both math and reading achievement. These positive effects fade out over time, however, becoming statistically insignificant in both subjects within five years. We also find that retention reduced the probability that a student would be retained in each of the four subsequent years. In contrast, we find no effects of third grade retention on student absences, special education placement, and attrition from the Florida public schools.

Our findings contribute to an emerging literature using quasi-experimental research designs to study the effects of retention in U.S. public schools.[2] Jacob and Lefgren (2004, 2009) exploit a non-linearity in the relationship between test scores and retention in third, sixth, and eighth grades to study the impact of retention on achievement and high school completion of Chicago students. They find that retention had a positive short-term effect on the achievement of third graders but not of sixth graders. They also find that retention increases dropout rates for eighth graders but not for sixth graders. In a previous study of the Florida policy, Greene and Winters (2007) find that third grade retention improved student achievement after two years.[3] Taken as a whole, this evidence suggests that retention in higher grade levels may have detrimental effects on future student outcomes, but that early grade retention may be beneficial. While we confirm that early grade retention positively affects student achievement in the short-run, we find that these initial effects fade out within five years.

Our evidence that early grade retention reduces the probability of later grade retention helps to clarify potential tradeoffs associated with policies that increase retention rates in early grades. Specifically, these results suggest that many of the students retained as third graders as a result of Florida's test-based promotion policy would otherwise have been retained in a subsequent grade. To the extent that later grade retention is in fact a less beneficial treatment, students who were retained earlier rather than later may benefit as a result of the policy's introduction. Overall, our results indicate that six years after potential grade 3 retention retained students are on average only 0.74 grade levels behind their non-retained peers. This implies that the cost of early grade retention for the individual student in terms of foregone earnings is likely to be substantially less than a full year.

The paper proceeds as follows. In Section 2 we describe the institutional setting for our analysis and our data. Section 3 presents our identification strategy and provides graphical evidence supporting its validity, while Section 4 presents our main findings concerning the effects of grade retention in third grade on future student outcomes. Section 5 concludes.

---

[2] In addition to the studies discussed in the text, Eide and Showalter (2001) use variation in kindergarten entry ages across states as an instrument for retention and conclude that retention increases high school completion and earnings for white students, although their results are not statistically significant. In a comparative setting, Manacorda's (2010) regression discontinuity analysis finds that retention in junior high school increases dropout rates for Uruguayan students.

[3] Greene and Winters (2012) present additional evidence of the impact of the Florida policy that is based on same-grade comparisons and does not provide a causal estimate of the impact of retention.

2. Institutional setting and data

In 2002, Florida's legislature mandated that third grade students scoring below level two (of five performance levels) on the Florida Comprehensive Assessment Test (FCAT) reading test be retained unless they qualify for one of six "good cause exemptions."[4] The Florida policy's exclusive focus on third grade reading distinguishes it from test-based promotion measures in Chicago and New York City, which include retention gates based on reading and math achievement at multiple grade levels. This focus reflects a common belief among educators that acquiring basic reading proficiency by third grade is essential for subsequent performance in multiple subjects, as well as the fact that third is the lowest grade included in the state testing program.

Students scoring below the level two cutoff can be granted an exemption from the policy if they fall into any of the following categories: Limited English proficiency (LEP) students with less than two years of instruction in English; students with disabilities whose Individualized Education Plan indicates that the third grade test is inappropriate; students with disabilities who were previous retained in third grade; students scoring above the 51st percentile nationally on another standardized reading test; students who demonstrated proficiency through a portfolio of work; or students who had been retained twice previously. Since the 2004-05 school year, retained students have also been given the opportunity for a midyear promotion to fourth grade if they demonstrate mastery of necessary skills.

As noted above, the policy includes several provisions intended to ensure that retained students acquire the reading skills needed to be promoted the following year. First, retained students must be given the opportunity to participate in their district's summer reading camp. Schools must also develop an academic improvement plan for each retained student and assign them to a "high-performing teacher," as determined by student performance data and satisfactory performance appraisals. Finally, during their retained year, retained students must receive intensive reading interventions, including ninety uninterrupted minutes daily of research-based reading instruction.[5] A lack of detailed information on take-up and implementation makes it impossible to disentangle the separate effects of these additional requirements.

[4] The description of the Florida program in this section is based on Office of Program Policy Analysis & Government Accountability (2006).

[5] Since 2004-05, the uninterrupted ninety minute reading block has been mandatory for all K-5 students.

The data for our analysis are drawn from the Florida Department of Education's PK-20 Education Data Warehouse and contain information on all Florida students attending public schools in grades 3 to 10 from the 2000–01 through 2008–09 school years. We identify retained students based on the grade level of the state tests taken in adjacent years.[6] Our data extract includes the school each student attends and its location; student characteristics such as ethnicity, gender, special education classification, English proficiency, and free lunch eligibility; annual measures of absences; and annual FCAT math and reading test scores.

Table 1 documents the structure of our data on student cohorts impacted by the test-based promotion policy. The first relevant cohort (which we will refer to as the 2003 cohort) entered third grade in the 2002–03 school year and can be followed for an additional six years after potential grade 3 retention, at which point promoted students who were not retained in a later grade should have reached ninth grade. The right-most column indicates that roughly 13 percent of the 2003 cohort were retained as third graders; six years later, the vast majority of these students were enrolled in eighth grade, but some were in grade seven (indicating that they had been retained a second time) or in grade nine (indicating that they had subsequently skipped a grade level). Among students not retained in third grade, most had progressed to ninth grade but a substantial number (five percent of the original cohort) were in eighth grade.[7] The five additional student cohorts that we include in our analysis enter third grade and can therefore be tracked for progressively shorter time period. The left-most column shows that, on average, 8 percent of all students in our sample were retained in grade 3.

Table 2 provides summary statistics for our pooled sample including the 2003-2008 cohorts. The first column reports mean characteristics (measured in third grade) for all students, columns 2 and 3 in turn include all students scoring below the cutoff and all students who were retained in third grade, and column 4 includes students who were retained in third grade despite exceeding the cutoff. Naturally, students scoring below the cutoff and retained students perform at low levels. For example, retained students score 1.43 standard deviations below the average student in reading and 1.22 standard deviations below the average student in math. There are almost twice as many students scoring below the cutoff as students actually retained, indicating that a large number of students were exempted from the retention requirement. Nonetheless, students scoring below the cutoff and retained students are quite

[6] Students receiving mid-year promotions after 2004-05 will therefore be recorded as not being retained.

[7] Below we examine the extent to which these differing patterns of grade progression for students retained and promoted as third graders reflect a causal impact of the initial retention decision.

similar with respect to their observable characteristics. In contrast, the relatively few voluntarily retained students are better performing on average, more likely to be white, and substantially younger than the average retained student. They are also more absent more frequently as third graders, perhaps suggesting the importance of behavioral indicators to voluntary retention decisions.

In addition to raw test scores that we standardize by grade, subject, and year, our data extract includes vertically equated Developmental Scale Scores (DSS) intended to support comparisons of student achievement across grade levels. During the 2000-01 school year, when the FCAT testing program was expanded to include reading and math in all grades three through ten, a special data collection scheme incorporated the use of common items administered to students across multiple grades. Specifically, operational items from each grade's test were also included on the test administered to the higher and lower adjacent grade. These common items provide a basis to use Item Response Theory (IRT) methods to place results from each grade's test on a common scale.[8]

Figure 1 plots average DSS scores in reading and math by grade for all students in the pooled dataset. The DSS scores have an across-grade, student-level standard deviation of 364 points in reading and 305 points in math. The jagged trajectory evident in both subjects indicates that average achievement gains as measured by developmental scale scores vary considerably by grade. For example, math gains are very small in grade six while reading gains are particularly pronounced in grade four. This variation could reflect imperfections in the vertical scaling process. Alternatively, it could reflect true differences in the average rate of learning in Florida public schools across grades. For example, the small math gains in grade six likely reflects the fact that most Florida students transition into a middle school in grade six, which Schwerdt and West (2011) show has a causal impact on their achievement growth. To the extent that retention simply delays students from experiencing a grade in which their own achievement growth is likely to be smaller, policymakers arguably would want to incorporate this information into the metric used to compare their achievement to that of promoted students.

The variation in achievement gains by grade motivates our construction of an alternative vertical scaling of reading and math achievement, which is also plotted in Figure 1. Specifically, we subtract from each student's DSS the grade-specific mean DSS score and

[8] See Hoffman et al. 2001 for technical details on the construction of the developmental scale scores.

then add the predicted value for each grade from a linear regression of average DSS scores on grade level. By construction, these rescaled scores increase linearly from grades three to ten, with the rate of achievement growth determined by the coefficient on grade level in the regression model. We use the rescaled scores as an alternative outcome measure when estimating the impact of retention on student achievement in our empirical analysis below.

3. Empirical strategy

Empirical strategies that rest on a selection-on-observable assumption will fail to provide unbiased estimates of the effect of early grade retention on future student outcomes if students are selected for retention on the basis of factors that are unobservable to the researcher and influence educational outcomes. We address this concern by taking advantage of Florida's test-based promotion policy, which led to a discontinuous relationship between third grade reading test scores and the probability of grade retention. This discontinuity generates plausibly exogenous variation in retention, which we leverage to identify the causal effect of grade retention on the future outcomes.

3.1 Graphical Evidence of the Discontinuity

Our identification strategy hinges on the assumption that Florida's test-based promotion policy generates exogenous variation in third grade retention which we can use standard regression discontinuity methods to exploit. We therefore first present graphical evidence of the existence of a discontinuity in the relationship between a student's third grade reading test scores and the probability of grade retention. We then discuss potential threats to the validity of regression discontinuity studies and provide additional graphical evidence demonstrating that these threats are not applicable in this setting (c.f., Lee and Lemieux, 2010). Unless otherwise noted, all figures are based on the pooled data set of students in the 2003-2008 cohorts.[9]

Figure 2, which plots the share of students retained as a function of third grade reading scores (measured relative to the test score cutoff), provides visual evidence of the discontinuity in retention probabilities. The data points represent the share of students retained for each possible score on the third grade reading test, with the size of each marker proportional to the number of students receiving that score. The solid line represents predicted values from separate local linear regressions on either side of the cutoff. For students 35 or

[9] Cohort-specific graphs are available from the authors upon request.

more points (>0.1 standard deviations) below the cutoff, retention probabilities are relatively stable at just above 50 percent. The probability of retention then declines as test scores increase, with retention probabilities immediately to the left of the cutoff approaching 30 percent. Retention probabilities drop sharply to less than 5 percent at the cutoff, however, and approach zero 50 points about the cutoff.

Figure 3 displays the same relationship for the two cohorts of students in our data extract entering third grade immediately prior to the introduction of the test-based promotion policy. Note that the probability of retention for students in these cohorts rarely exceeded 20 percent, even for very low-scoring students. More importantly, the probability of retention is essentially continuous around the cutoff, indicating that the discontinuity evident in Figure 2 was in fact generated by the policy change.

While Figure 2 is based on the full distribution of third grade reading test scores, we base our regression discontinuity analysis of the causal effects of retention on a narrower sample of students within a bandwidth of 10 test score points on either side of the cutoff.[10] Figure 4 illustrates the discontinuity within this more restricted sample, again plotting the fraction of students retained by third grade reading test scores measured relative to the cutoff. Local linear regressions on either side of the cutoff suggest an approximately linear relationship between test scores and retention probabilities in the cutoff region. However, the slope of this relationship clearly differs for students below and above the cutoff. We make use of this observation below when specifying the functional relationship between the forcing variable (reading test scores) and the treatment indicator in our empirical model.

A common concern with the regression discontinuity approach is the possibility of precise manipulation of the forcing variable around the cutoff (c.f., Urquiola and Verhoogen 2009). In this setting, for example one might, for example, worry that teachers were able to manipulate reading scores of students to push their scores just above the cutoff. The fact that the FCAT reading test is objectively scored without teacher input makes this possibility unlikely, however, and Figure 5 confirms that the overall distribution of reading test scores shows no evidence of a heaping of observations around the cutoff.

Finally, the regression discontinuity identification strategy also assumes that there are not discontinuities in other characteristics associated with student outcomes at the cutoff.

[10] We select this bandwidth based on the optimal bandwidth algorithm developed by Imbens and Kalyanaraman (2009).

Figure 6 addresses this issue by plotting the mean value of the observable student characteristics available in our data against third grade reading test scores. In addition to examining each characteristic individually, we also generate a predicted retention probability for each student based on all available background characteristics (except for reading scores). The figure confirms the absence of discontinuities in student observed characteristics at the test-score cutoff used to inform retention decisions.

3.2 Estimation

Because only a subset of students scoring below the cutoff in reading test scores were actually retained, our empirical analysis takes the form of a fuzzy regression discontinuity design which can be implemented via instrumental variables (IV) estimation. Thus, in our preferred specification we estimate the causal effect of early grade retention on future student outcomes in a two-stage least squares model. The first stage is given by the following equation:

(1) $retain = \beta_1 * \text{below} + \beta_2 * \text{below} * \text{forcevar} + \beta_3 * \text{forcevar} + \text{BX} + \varepsilon,$

where *retain* indicates retention in grade 3, *below* indicates that the student scored below the promotion cutoff on the grade 3 reading test, *forcevar* measures student achievement on the grade 3 reading test, *X* is a vector of student demographic characteristics including the student's math score in grade 3, and ε is a standard zero-mean error term. Note that we model the relationship between reading scores and the retention indicator as a linear relationship with a break in trend at the cutoff, because of the graphical evidence of this relationship in Figure 4.

The corresponding second stage of our 2SLS model is given by:

(2) $y = \gamma_1 * \text{retained} + \gamma_2 * \text{below} * \text{forcevar} + \gamma_3 * \text{forcevar} + \Gamma\text{X} + \eta,$

where *y* denotes the student outcome of interest. Note that we achieve identification of γ_1 by instrumenting for grade retention in grade 3 (*retained*) with the indicator for being below the cutoff for promotion to grade 4 (*below*). As explained above, we estimate the 2SLS model for the sample of students within ten test score points on either side of this cutoff.[11] To compare our preferred IV results with commonly reported estimates of the effects of retention based on a selection-on-observables assumption, we additionally estimate Equation (2) based on OLS.

[11] Results for alternative bandwidths are qualitatively identical to the results presented below and are available from the authors upon request.

To maximize comparability across designs, we also limit the OLS analysis to the regression discontinuity sample.

4 Results

Table 3 reports results from estimating the first-stage model in Equation (1) separately for each cohort of students separately and for the pooled sample. For purposes of comparison, we also present results for the two cohorts of students in our data who were not impacted by the policy. Note that all estimations are based on our preferred discontinuity sample with a bandwidth of 10 test score points around the cutoff. Despite this narrow bandwidth, we still have between 9,981 and 15,687 students in each post-2002 cohort and a total of nearly 75,000 students in the pooled sample.

The first row of Table 3 presents estimates of the jump in the probability of retention at the promotion cutoff. Consistent with the evidence in Figure 3, the first two columns show that there was essentially no such jump in the two years immediately preceding the policy's introduction.[12] In contrast, each of the cohort-specific estimates for students impacted by the policy is positive and highly statistically significant, with F-statistics on the excluded instrument exceeding 400 in four of six cases. Point estimates of the size of the jump in retention probabilities at the cutoff range from 0.20 to 0.37, with the largest estimate observed for the initial 2003 cohort and the smallest two estimates observed for the 2007 and 2008 cohorts. This suggests that compliance with the retention requirement was relatively low (a pattern which is arguably consistent with the availability of good cause exemptions) and appears to have declined over time. The overall first stage effect for the pooled sample indicates an increase of 28 percentage points in the probability of students scoring immediately below the cutoff, relative to students scoring one point higher.

4.1 The effect of early grade retention on student achievement

We begin our discussion of the effects of grade retention on student outcomes with a graphical investigation of the reduced form relationship between students' third grade reading test scores and their future achievement. Figures 7 and 8 depict the relationship between predicted student achievement from local linear regressions in math and reading up to six years after potential third grade retention and students' third grade reading test scores relative

[12] The results for the 2002 cohort show a statistically significant, but substantively small (<2 percentage point) increase in the probability of retention for students scoring below the cutoff. The cohort-specific estimates while the policy was in place are all more than ten times as large.

to the cutoff. In both reading and math, we observe students scoring below the promotion cutoff performing at higher levels in the first three years after potential third grade retention. However, these differences dissipate in later years and, in some cases, appear to turn slightly negative.

Table 4 presents estimates of the effects of grade three retention on reading and math achievement over time. Columns 1 and 2 report OLS estimates from Equation (2) with and without covariates, while columns 3 and 4 report results from our preferred IV model exploiting the discontinuity. As expected, the inclusion of covariates does not notably influence the IV point estimates (although it modestly improves their precision) but does substantially alter the OLS results.[13]

Our IV estimates indicate that third grade retention substantially improves students reading and math achievement in the short run. Measured relative to the statewide standard deviation in reading DSS scores, reading achievement improves by 23 percent of a standard deviation after one year and by as much as 48 percent of a standard deviation after two years. The estimated impact of retention on math scores is 30 percent of a standard deviation after one year and grows to 36 percent of a standard deviation after three years. These initial benefits fade out strongly in subsequent years, however. The effects of third grade retention on reading scores are reduced in years three and four and become statistically insignificant in years five and six. In the case of math achievement, the estimated effects become slightly negative in years four and five but are statistically insignificant after six years.

In contrast to our preferred IV estimates, OLS estimates of the effects of third grade retention are always more negative and would suggest a statistically significant negative impact on reading and math achievement after 6 years. We interpret this difference as evidence that OLS estimates fail to adequately control for unobserved confounding factors and, thus, will understate any benefits (and exaggerate any harms) of grade retention.

One unusual aspect of the results in Table 4 is the non-monotonic relationship between the size of the estimated impacts of retention and the time elapsed since the student was retained. The estimated impact is largest after two years in the case of reading achievement and after three years in math. Especially given the overall pattern of fade out, one would expect the impact of retention to be largest in the year the student was retained. This pattern

[13] Appendix Table A-1 provides estimates of the impact of third grade retention on reading and math achievement results by cohort.

likely stems in part from the grade-to-grade variation in the average achievement gains of Florida public school students as measured by DSS scores. For example, Figure 1 shows that Florida students statewide experience large gains in DSS reading achievement in fourth grade, which promoted students enter immediately and (most) retained students enter one year later. This difference in timing could explain the unexpected growth from year one to year two in the estimated impact of retention on DSS reading achievement. As discussed above, our alternative scaling of the DSS scores eliminates variation in average achievement grade across grades and arguably provides a clearer basis for evaluating the retention policy's impact.

Table 5 presents OLS and IV estimates of Equation (2) based on these rescaled DSS scores. In both reading and math, the magnitude of the estimated impacts now decreases monotonically with distance from treatment. In reading, the impacts are as large as 59 percent of a standard deviation after one year but fade to 12 percent of a standard deviation by year four and are statistically insignificant thereafter. In math, the impacts start at 44 percent of a standard deviation but are statistically insignificant by year four and become modestly negative after six years. Qualitatively, however, the results concerning achievement impacts of third grade retention do not depend on the test scaling. Both approaches show large positive initial impacts of retention that fade out fully over time.[14]

4.2 The effect of early grade retention on grade progression

We next present estimates of the effect of third grade retention on the probability that students will be retained in subsequent grades and on the number of grades the student has completed. This is an important issue to consider when evaluating test-based promotion policies for at least two reasons. First, it has direct implications for retention's costs to both the individual and society. If early grade retention influences the probability that students are retained at higher grade levels, the cost of early grade retention in terms of foregone earnings and additional educational expenditures could be well below a full school year. Second, the effects of retention on outcomes such as student achievement and attainment could vary according to the grade level at which the student is retained. To the extent that retention in early grades is more beneficial to students than later retention, test-based promotion policies targeting early grades could benefit students who would eventually be retained by ensuring that they are retained at a younger age.

[14] Appendix Tables A-3, A-5, and A-6 provide additional estimates of the impact of third grade retention on reading and math achievement for various student subgroups.

Figure 9 provides graphical evidence on the reduced form relationship between third grade reading test scores and retention probabilities in each of the next six years after their initial third grade year. The figure indicates that students below the promotion cutoff are substantially less likely to be retained each year from two to five years after potential third grade retention.

Table 6 shows the corresponding estimates of the effect of third grade retention on future retention probabilities.[15] The IV estimates confirm that third grade retention reduces the probability that the student will be retained two years later by 11 percentage points. The effect is smaller in subsequent years, but remains statistically significant and ranges from 3-4 percentage points in magnitude in years three to five. Table 7 uses grade level as the outcome variable in Equation (2), thereby providing direct evidence on the differences in the grade progression of retained and promoted students. The IV estimates show that 6 years after being retained in third grade, students impacted by Florida's test-based promotion policy are only 0.74 grade levels behind comparable peers who were promoted.

The evidence in Tables 6 and 7 confirm that third grade retention substantially reduced the probability that Florida students at the promotion cutoff would be retained in future grades. Could these differences in the subsequent grade progression of retained and promoted students explain the fade out of test score impacts evident in Tables 4 and 5? To evaluate this possibility, we assume that (1) the effects of retention on student achievement after one year are in fact fully persistent and (2) that students retained in subsequent grades experience the same benefits, regardless of the grade in which they were retained. We then ask how much of the overall fade out in test score impacts would be explained by the additional gains made by students retained in subsequent years. The results suggest that differences in subsequent retention could account for no more than 34 percent of the fade out in reading effects after two years and 22 percent of the fade out in math effects. We also confirm that the fade out in test score impacts in both subjects remains when students who were subsequently retained are excluded from the analysis.[16]

4.3 The effect of early grade retention on other student outcomes

Tables 8, 9, and 10 report estimates of the effect of third grade retention on three other student outcomes: absences, special education placement, and attrition from Florida public

[15] Appendix Table A-2 provides estimates of the impact of third grade retention on future retention by cohort. Appendix Table A-4 provides parallel estimates for various student subgroups.

[16] These results are available from the authors upon request.

schools. The results confirm that retention had no impact on these outcomes for students with third grade reading scores at the promotion cutoff.

5. Conclusion

Our analysis exploits a discontinuity in the probability of grade retention under Florida's test-based promotion policy to study the effects of third grade retention on student outcomes up to six years later. We find evidence of substantial short-term gains in both math and reading achievement. However, these positive effects fade out over time and become statistically insignificant within five years. We also find substantial negative effects of grade three retention on the probability of being retained in later grades. We do not find any significant effects of grade three retention on student absences, special education placement and attrition from Florida public schools.

In sum, our analysis provides more favorable evidence on the effects of early grade retention than found in many previous studies – in particular those which do not rely on credible quasi-experimental methods to address unobserved selection into the retention treatment. We show that early grade retention has substantial positive effects on reading and math achievement in the short run, has no detrimental effects on those student outcomes we can measure, and generates educational and opportunity costs well below a full year when later grade retention is taken into account. To the extent that early grade retention is more beneficial than later grade retention (as suggested by the results of Jacob and Lefgren 2005, 2009), students who were retained in third grade and would have been retained later may have benefited from the introduction of the Florida policy. However, we also do not find clear evidence that early grade retention is beneficial for students in the long run.

The fade out of test score impacts is a common pattern in the literature on educational interventions, including those which have been shown to generate lasting impacts on adult outcomes. For example, Chetty et al. (2011) show that kindergarten classroom quality improves adult earnings despite test-score fade out. The same appears to be true of early childhood interventions such as the Perry and Abecederian preschool demonstration projects and the Head Start program (see Almond and Currie [2011] for a review). Does the retention of low-performing students yield similar long-term benefits? If so, do these benefits exceed the costs? Whatever their merits as a matter of public policy, the spread of test-based promotion policies providing credibly exogenous variation in retention probabilities will soon provide a welcome opportunity for research addressing these questions.

References

Adams, C.J., E.W. Robelen, and N. Shah. 2012. "Civil Rights Data Show Retention Disparities," *Education Week*, March 6, http://www.edweek.org/ew/articles/2012/03/07/23data_ep.h31.html.

Allen, C.S., Q. Chen, V.L. Willson and J.N. Hughes. 2009. "Quality of Research Design Moderates Effects of Grade Retention on Achievement: A Meta-Analytic, Multilevel Analysis," *Educational Evaluation and Policy Analysis*, 31(4): 480-499.

Almond, D. and J. Currie 2011. "Human Capital Development Before Age 5." In *Handbook of Labor Economics*, vol. 4b, ed. O. Ashenfelter and D. Card, 1315-1486. Amsterdam: Elsevier.

Chetty, R., J.N. Friedman, N. Hilger, E. Saez, D.W. Schanzenbach, D. Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR," *Quarterly Journal of Economics*, 126(4): 1593-1660.

Eide, E.R. and M.H. Showalter. 2001. "The effect of grade retention on educational and labor market outcomes," *Economics of Education Review*, 20(6): 563-576.

Greene, J.P. and M.A. Winters. "Revisiting Grade Retention: An Evaluation of Florida's Test-Based Promotion Policy," *Education Finance and Policy*, 2(4): 319-340.

Greene, J.P. and M.A. Winters. 2012. "The Medium-Run Effects of Florida's Test-Based Promotion Policy," *Education Finance and Policy*, forthcoming.

Hoffman, R.G., L.L. Wise and A.A. Thacker. 2001. "FLORIDA COMPREHENSIVE ASSESSMENT TEST: Technical Report on Vertical Scaling for Reading and Mathematics." San Antonio, TX: Harcourt Educational Measurement.

Imbens, G. and K. Kalyanaraman. 2009. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," NBER Working Paper 14726. Cambridge, MA: National Bureau of Economic Research.

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

Jacob, B., A. and L. Lefgren. 2009. "The Effect of Grade Retention on High School Completion," *American Economic Journal: Applied Economics*, 1(3): 33-58.

Jimerson, S., B. Egeland, L.A. Sroufe and B. Carlson. 2000. "A Prospective Longitudinal Study of High School Dropouts Examining Multiple Predictors Across Development," *Journal of School Psychology*, 38(6): 525-549.

Jimerson, S.R. 1999. "On the Failure of Failure: Examining the Association between Early Grade Retention and Education and Employment Outcomes During Late Adolescence," *Journal of School Psychology*, 37(3): 243-272.

Jimerson, S.R., G.E. Anderson and A.D. Whipple. 2002. "Winning the Battle and Losing the War: Examining the Relation between Grade Retention and Dropping Out of High School," *Psychology in the Schools*, 39(4): 441-457.

Lee, D.S. and T. Lemieux. 2010. "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48(2): 281-355.

Manacorda, M. 2012. "The Cost of Grade Retention," *Review of Economics and Statistics*, 94(2): 596-606.

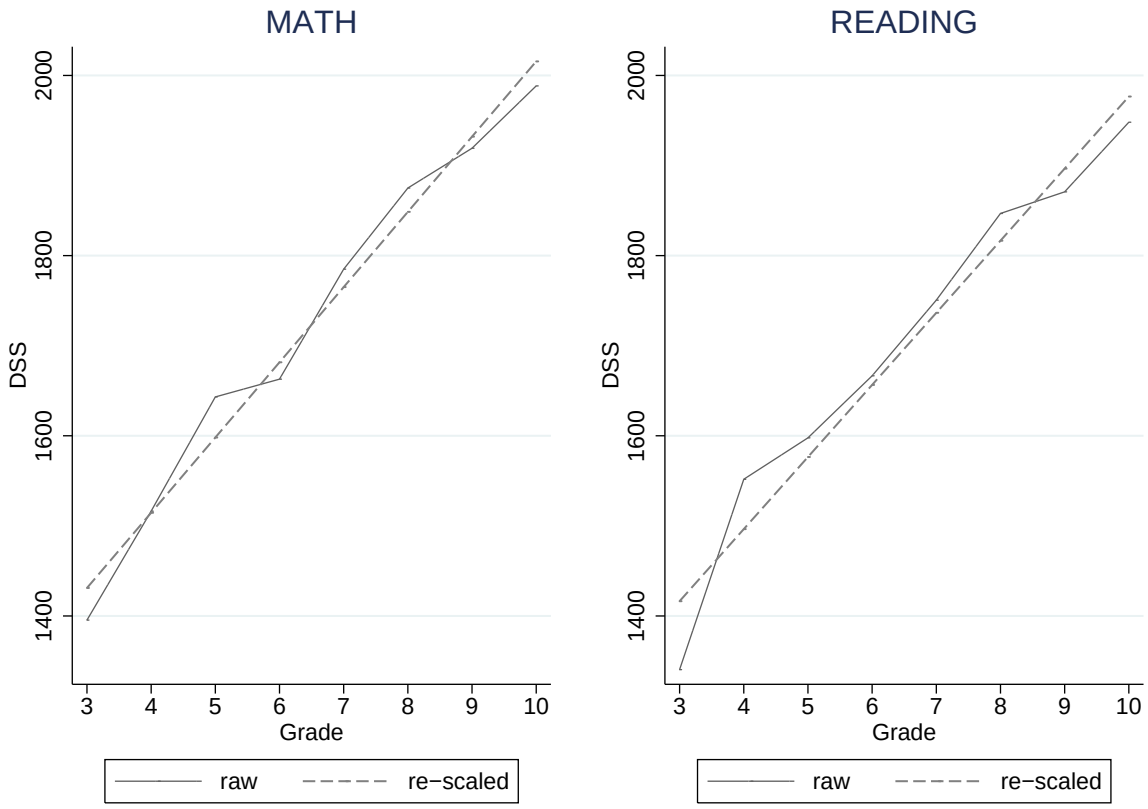
McCoy, A.R. and A.J. Reynolds. 1999. "Grade Retention and School Performance: An Extended Investigation," *Journal of School Psychology*, 37(3): 273-298.

Office of Program Policy Analysis & Government Accountability. 2006. "Third Grade Retention Leading to Better Student Performance Statewide." OPPAGA Report No. 06-66, http://www.oppaga.state.fl.us/reports/pdf/0666rpt.pdf.

Schwerdt, G. and M.R. West. 2011. “The Impact of Alternative Grade Configurations on Student Outcomes Through Middle and High School.” PEPG Working Paper 11-02. Cambridge, MA: Program on Education Policy and Governance.

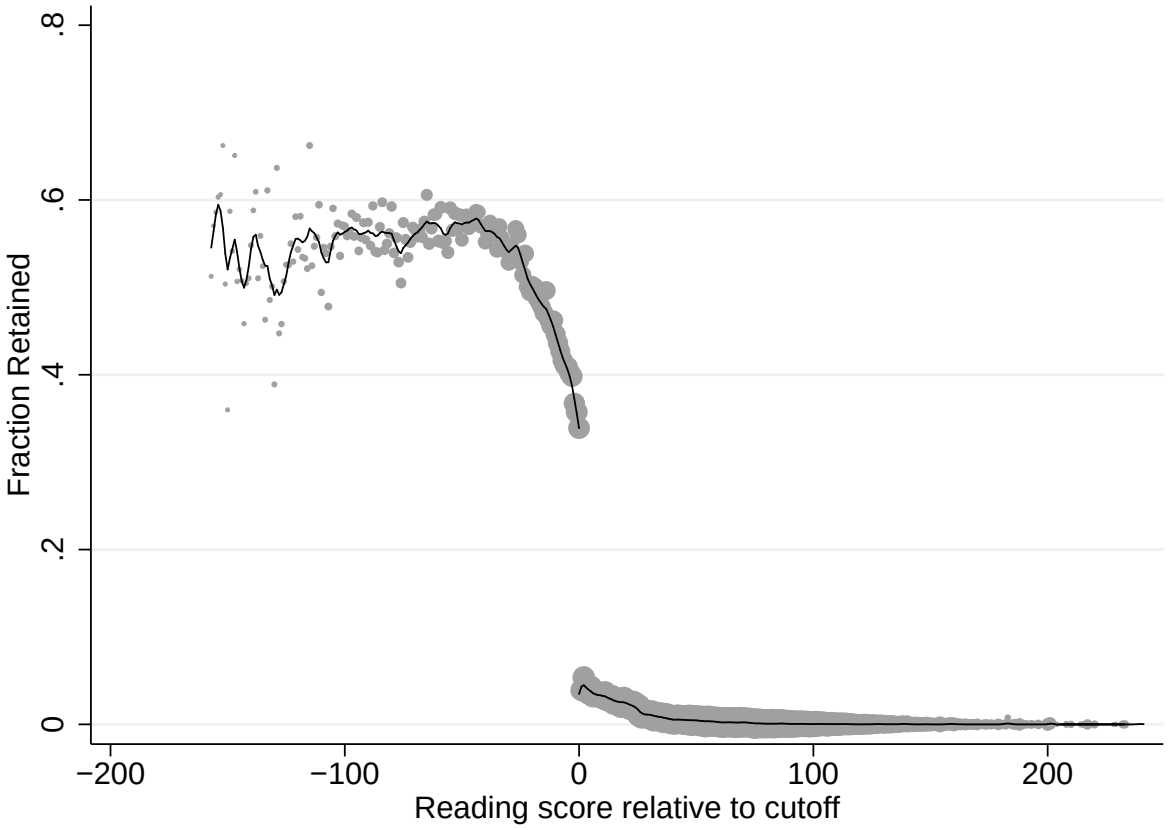
Urquiola, M. and E. Verhoogen. 2011. “Class-Size Caps, Sorting, and the Regression Discontinuity Design,” *American Economic Review*, 99(1): 179-215.

Figure 1: Average Developmental Scale Scores by Subject and Grade



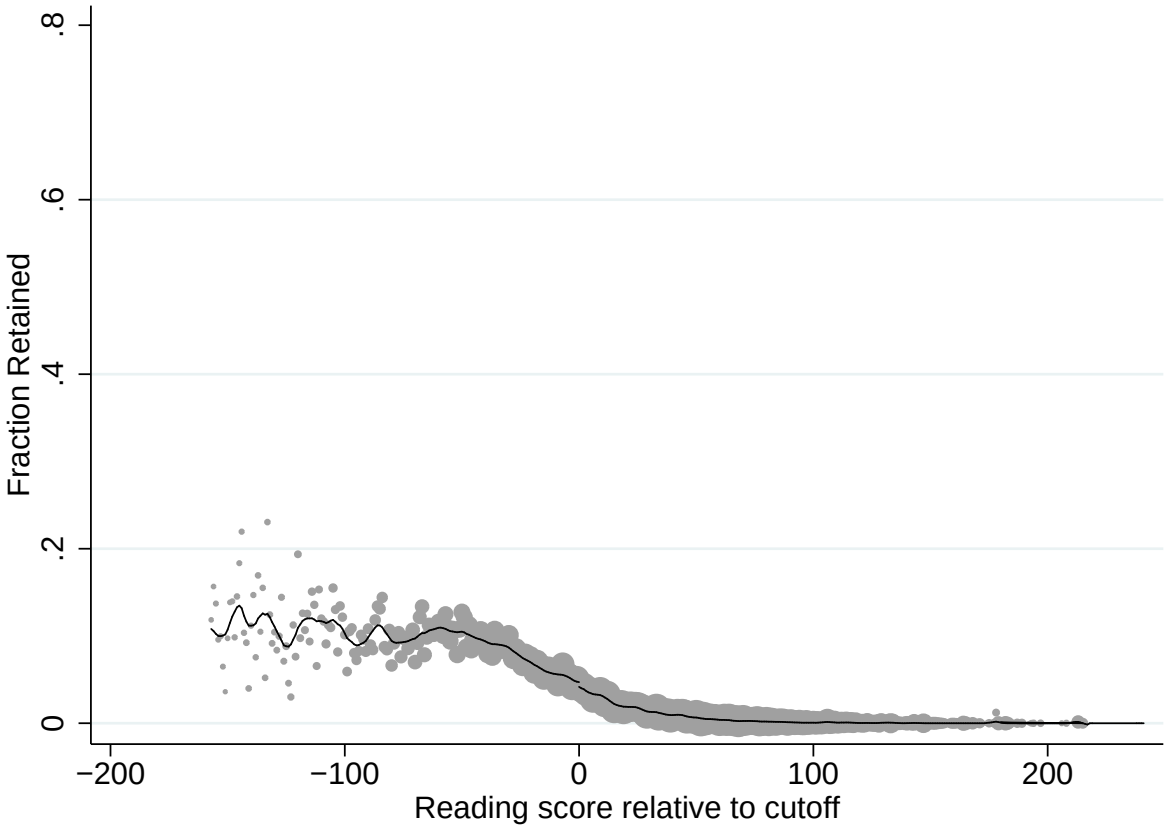
Note: Based on all students in grades 3 to 10 between 2002 and 2009. Re-scaled scores stem from predicted values of a linear regression of developmental scale scores on grade levels.

Figure 2: The Relationship between Reading Scores in Grade 3 and the Probability of being retained in Grade 3 after 2003



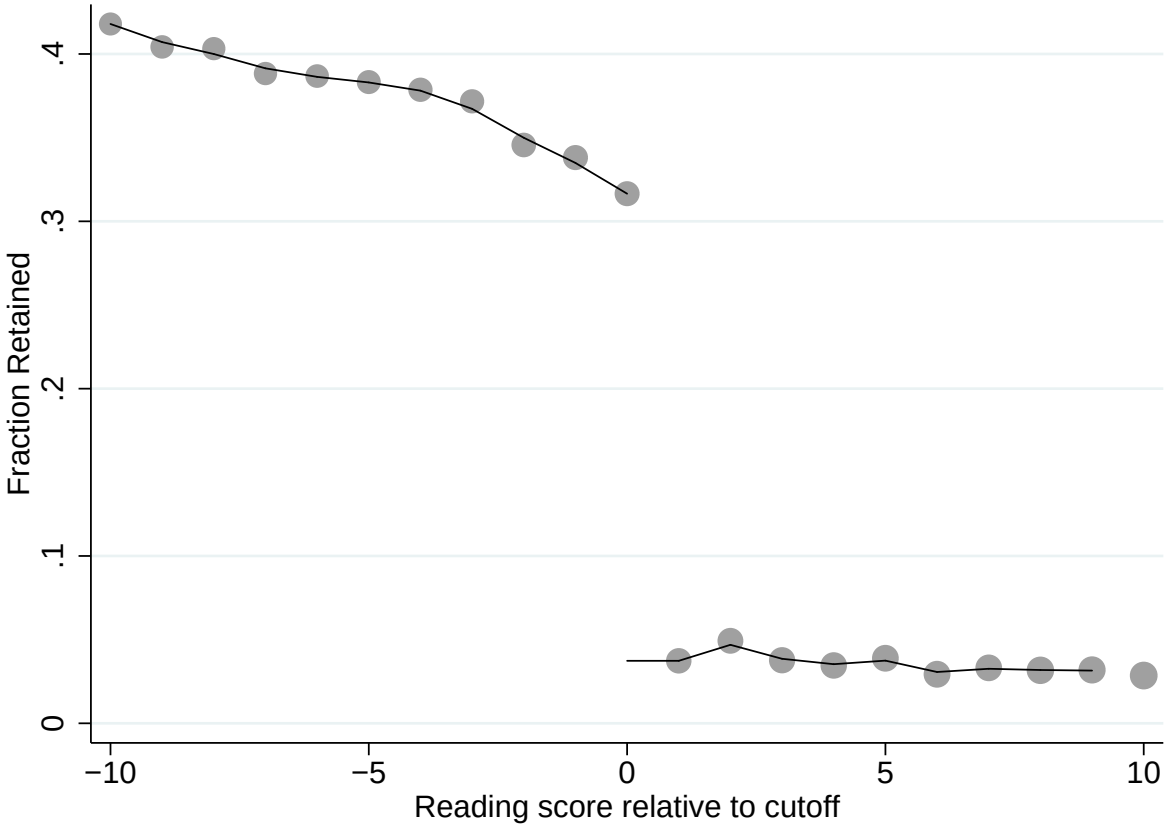
Note: Based on cohorts 2003-2008. Entire sample. Solid line represents predicted values from local linear regressions on both sides of the cutoff. Marker size represents relative group size.

Figure 3: The Relationship between Reading Scores in Grade 3 and the Probability of being retained in Grade 3 before 2003



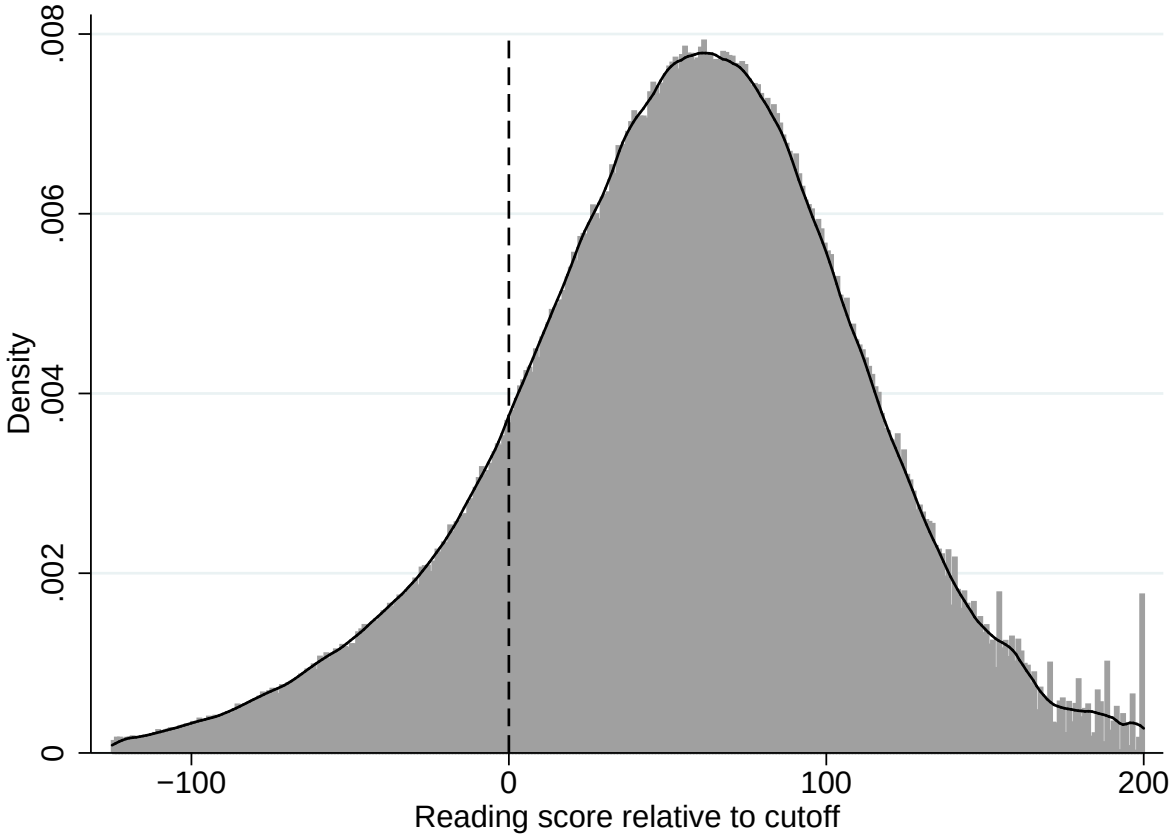
Note: Based on cohorts 2001-2002. Entire sample. Solid line represents predicted values from local linear regressions on both sides of the cutoff. Marker size represents relative group size.

Figure 4: The Relationship between Reading Scores and Grade Retention around the Cutoff [Bandwidth = 10]



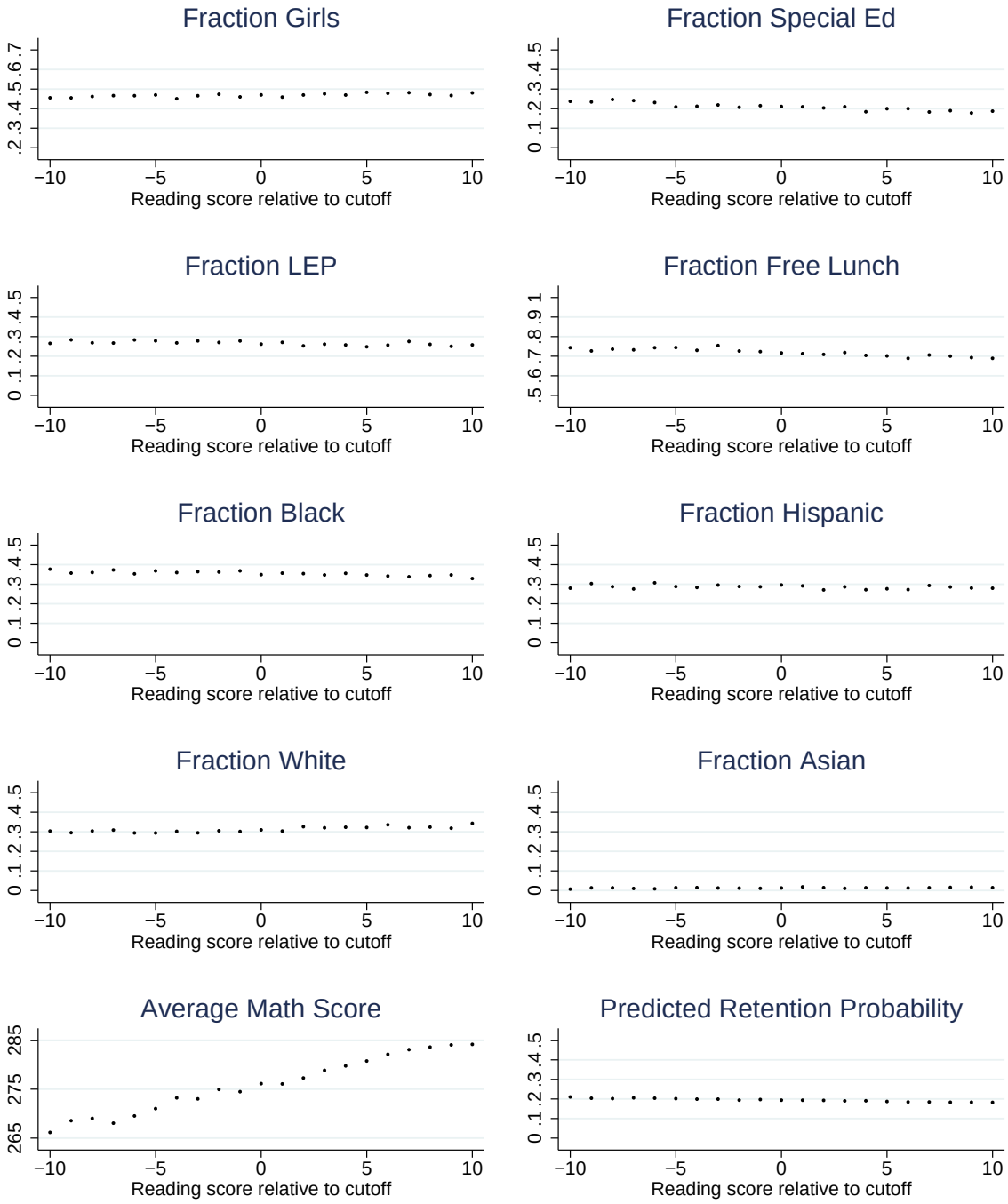
Note: Based on cohorts 2003-2008. Discontinuity sample with 10 test score points bandwidth. Solid line represents predicted values from local linear regressions on both sides of the cutoff. Marker size represents relative group size.

Figure 5: Distribution of Reading Scores in Grade 3



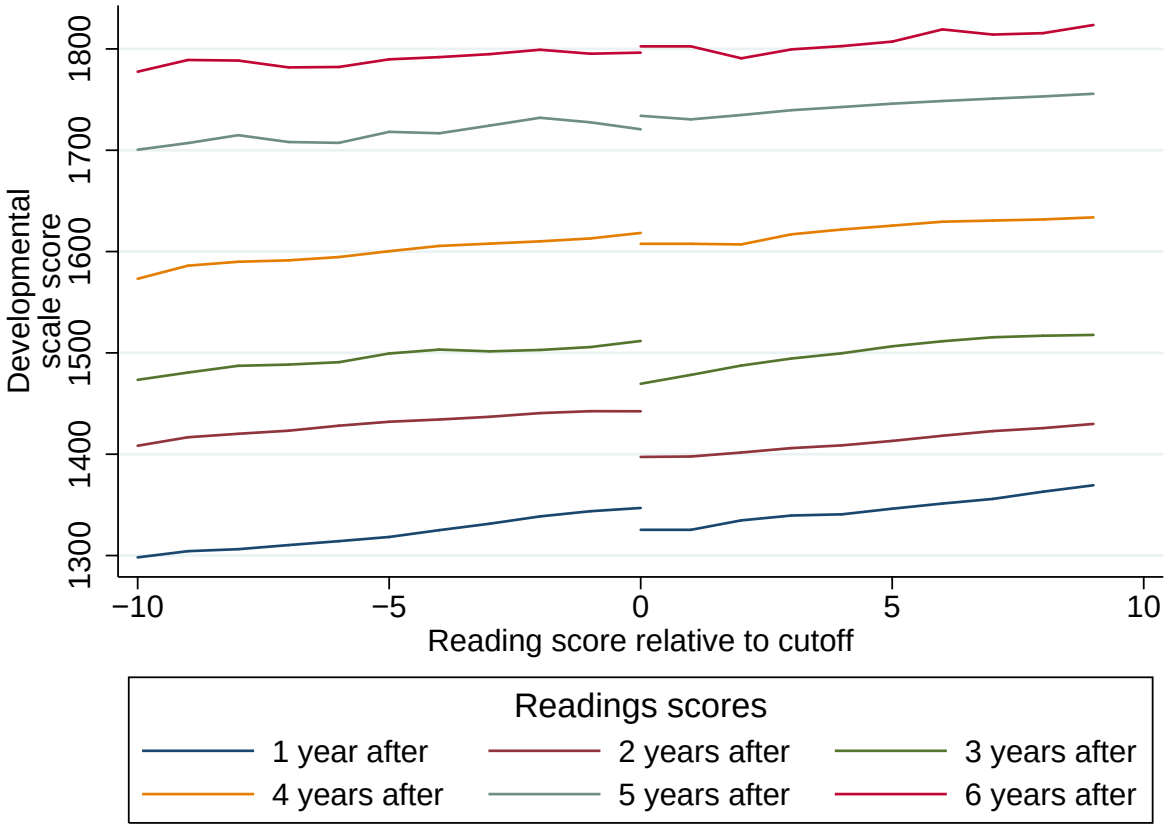
Note: Based on cohorts 2003-2008. Entire sample. Solid line represents kernel density estimates.

Figure 6: The Relationship between Reading Scores in Grade 3 and Student Background Characteristics



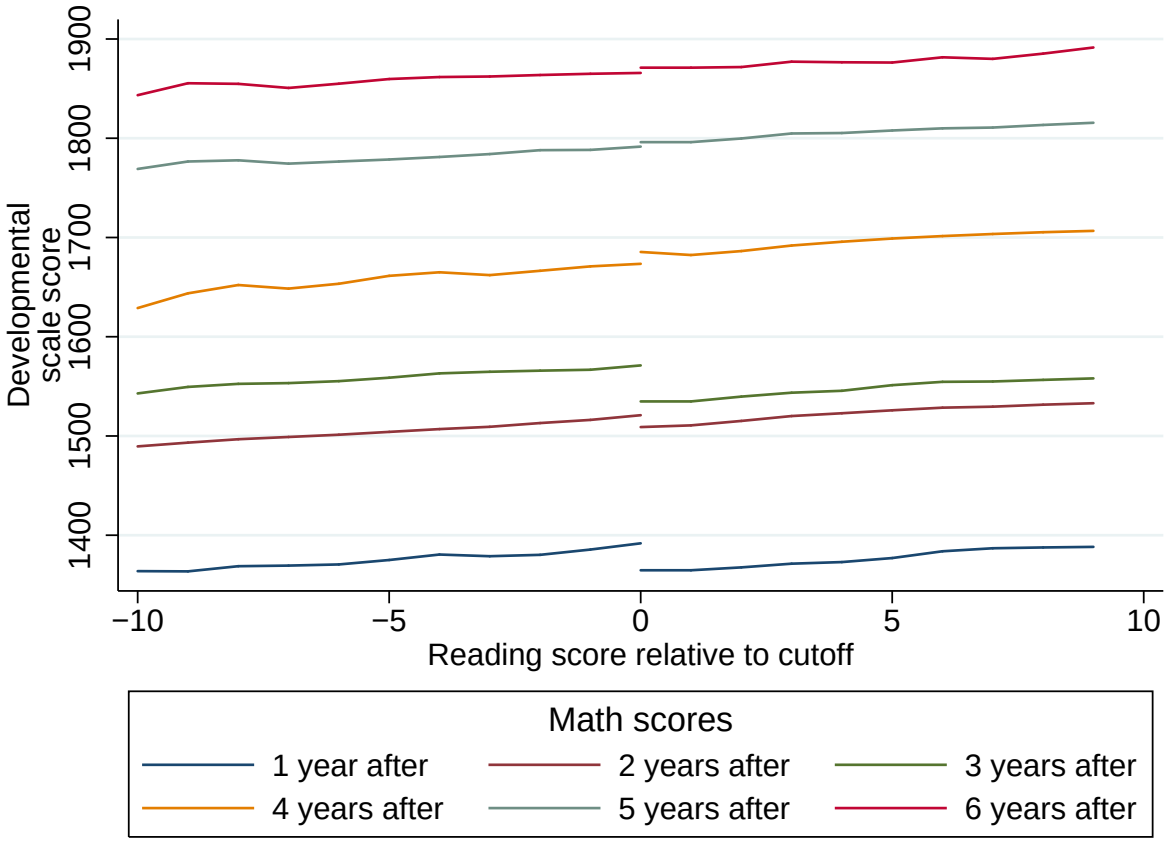
Note: Based on cohorts 2003-2008. Discontinuity sample with 10 test score points bandwidth. Predicted retention probability displays predicted values after estimating a probit model that includes all student background variables except for reading scores as explanatory variables.

Figure 7: The Relationship between Reading Scores in Grade 3 and Future Reading Achievement around the Cutoff



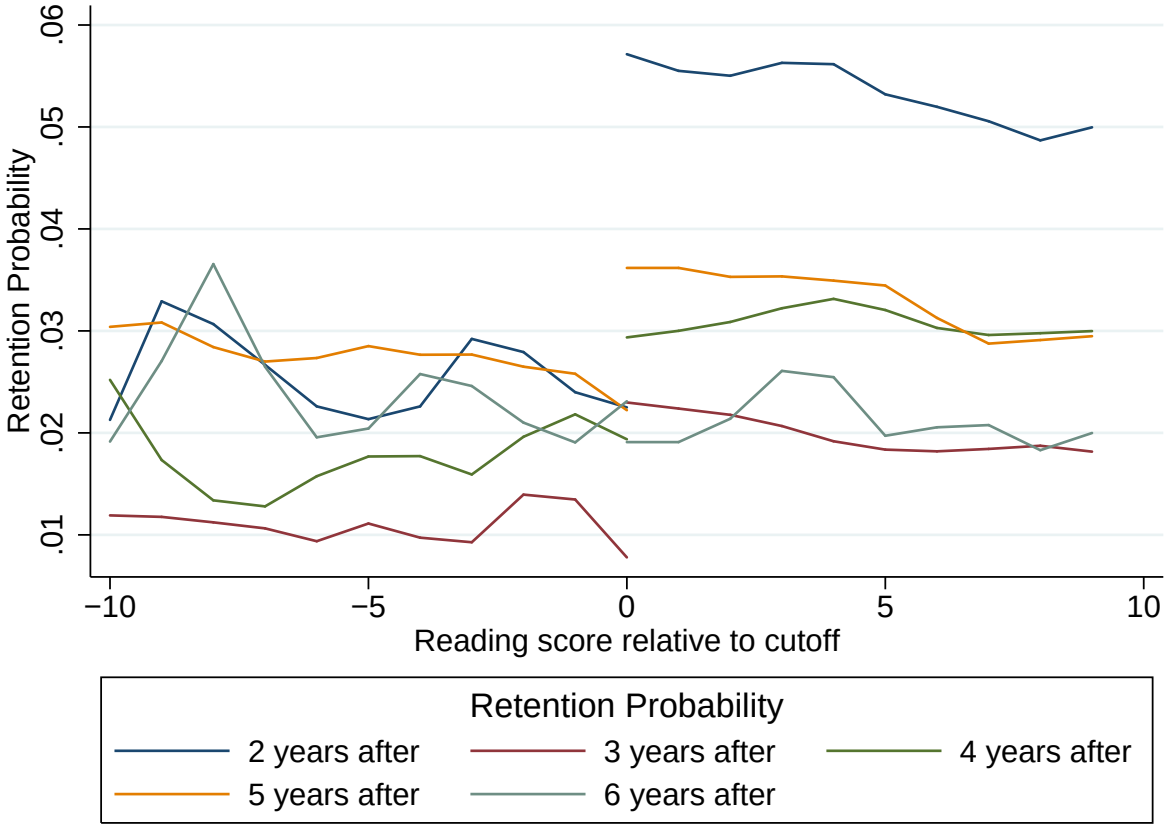
Note: Based on cohorts 2003-2008. Discontinuity sample with 10 test score points bandwidth. Solid line represents predicted values from local linear regressions on both sides of the cutoff.

Figure 8: The Relationship between Reading Scores in Grade 3 and Future Math Achievement around the Cutoff



Note: Based on cohorts 2003-2008. Discontinuity sample with 10 test score points bandwidth. Solid line represents predicted values from local linear regressions on both sides of the cutoff.

Figure 9: The Relationship between Reading Scores in Grade 3 and Future Grade Retention around the Cutoff



Note: Based on cohorts 2003-2008. Discontinuity sample with 10 test score points bandwidth. Solid line represents predicted values from local linear regressions on both sides of the cutoff.

Table 1: Observations by Year and Grade

	Years after potential treatment (= retention in grade 3)					
	1	2	3	4	5	6
Grade 3						
T=1	**0.08**	0.00	-	-	-	-
T=0	0.00	0.00	-	-	-	-
Grade 4						
T=1	0.00	**0.09**	0.00	-	-	-
T=0	**0.92**	0.01	0.00	-	-	-
Grade 5						
T=1	0.00	0.00	**0.09**	0.00	-	-
T=0	0.00	**0.90**	0.02	0.00	-	-
Grade 6						
T=1	-	0.00	0.00	**0.10**	0.01	-
T=0	-	0.00	**0.88**	0.04	0.00	-
Grade 7						
T=1	-	-	0.00	0.00	**0.10**	0.01
T=0	-	-	0.00	**0.86**	0.05	0.00
Grade 8						
T=1	-	-	-	0.00	0.01	**0.11**
T=0	-	-	-	0.00	**0.83**	0.05
Grade 9						
T=1	-	-	-	-	0.00	0.01
T=0	-	-	-	-	0.00	**0.81**
Cohorts	2003-2008	2003-2007	2003-2006	2003-2005	2003-2004	2003
Students	983,308	768,593	578,387	418,680	275,194	134,284

Table 2: Summary Statistics

	Total	Failed Promotion Cuttoff	Retained	Retained, but above Cuttoff
FCAT Math	0.06	-1.13	-1.22	-0.83
FCAT Reading	0.07	-1.46	-1.43	-0.38
Female	0.49	0.42	0.42	0.46
Age	8.84	9.06	8.89	8.77
White	0.48	0.28	0.28	0.50
Black	0.22	0.38	0.40	0.29
Hispanic	0.24	0.31	0.29	0.15
Asian	0.02	0.01	0.01	0.01
Other	0.04	0.03	0.03	0.04
Free or reduced lunch	0.52	0.78	0.79	0.65
Limited English proficiency	0.19	0.30	0.29	0.11
Special Education	0.16	0.37	0.29	0.15
Days absent	7.46	9.10	9.28	10.13
Number of students	983,308	159,866	81,357	4,959

Note: Based on cohorts 2003-2008. Entire sample. Test scores in math and reading are normalized by subject, year, and grade to have a mean of zero and a standard deviation of one.

Table 3: The Effect of Reading Performance on the Probability of Grade Retention in Grade 3

	Policy off		Policy on						
Cohorts	2001	2002	2003	2004	2005	2006	2007	2008	2003-2008
Below cutoff	0.006	0.019***	0.373***	0.268***	0.295***	0.338***	0.198***	0.217***	0.283***
	[0.006]	[0.007]	[0.012]	[0.013]	[0.013]	[0.016]	[0.012]	[0.013]	[0.005]
Reading	–0.000	–0.001	–0.001	–0.000	–0.001	–0.001	–0.001	0.001	–0.001*
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.000]
Reading	0.000	0.000	–0.010***	–0.009***	–0.006***	–0.005**	–0.007***	–0.008***	–0.008***
× Below cutoff	[0.001]	[0.001]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.001]
Math	–0.000***	–0.001***	–0.001***	–0.001***	–0.001***	–0.001***	–0.001***	–0.001***	–0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Female	–0.009***	–0.012***	–0.026***	–0.041***	–0.022***	–0.023***	–0.035***	–0.040***	–0.031***
	[0.003]	[0.003]	[0.006]	[0.007]	[0.006]	[0.007]	[0.006]	[0.006]	[0.003]
Age	–0.014***	–0.011***	–0.033***	–0.042***	–0.037***	–0.057***	–0.031***	–0.038***	–0.039***
	[0.002]	[0.003]	[0.005]	[0.005]	[0.005]	[0.005]	[0.004]	[0.004]	[0.002]
Black	–0.029***	–0.025***	–0.027***	–0.030***	–0.010	–0.001	–0.023***	–0.032***	–0.021***
	[0.004]	[0.005]	[0.008]	[0.009]	[0.009]	[0.010]	[0.008]	[0.008]	[0.003]
Hispanic	–0.017***	–0.024***	0.002	–0.007	–0.032***	–0.000	–0.026***	–0.026***	–0.015***
	[0.005]	[0.005]	[0.010]	[0.011]	[0.010]	[0.012]	[0.009]	[0.010]	[0.004]
Asian	–0.028***	–0.009	–0.026	–0.009	–0.083***	–0.005	–0.026	–0.072***	–0.035***
	[0.009]	[0.013]	[0.024]	[0.027]	[0.025]	[0.028]	[0.022]	[0.021]	[0.010]
Other race	–0.037***	–0.002	–0.018	–0.047**	–0.050***	0.020	–0.025	–0.009	–0.022***
	[0.007]	[0.012]	[0.018]	[0.019]	[0.017]	[0.022]	[0.015]	[0.016]	[0.007]
Free lunch	0.017***	0.014***	0.047***	0.023***	0.030***	–0.001	0.024***	0.017**	0.027***
	[0.003]	[0.004]	[0.007]	[0.008]	[0.008]	[0.009]	[0.007]	[0.007]	[0.003]
Special Education	–0.023***	–0.020***	–0.060***	–0.065***	–0.066***	–0.090***	–0.046***	–0.045***	–0.061***
	[0.004]	[0.005]	[0.008]	[0.008]	[0.008]	[0.009]	[0.007]	[0.007]	[0.003]
LEP	–0.013***	–0.014***	–0.049***	–0.060***	–0.004	–0.013	–0.026***	–0.007	–0.027***
	[0.004]	[0.005]	[0.009]	[0.010]	[0.009]	[0.011]	[0.008]	[0.009]	[0.004]
Year FE	No	No	No	No	No	No	No	No	Yes
Students	17,676	16,516	15,687	12,040	12,435	9,981	12,995	11,536	74,674
R^2	0.018	0.020	0.297	0.207	0.229	0.253	0.158	0.169	0.227
F-statistic on instrument	0.92	8.01	895.72	402.03	497.35	473.78	289.23	292.30	2,778.43
Pr > F	0.34	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

* p<0.10, ** p<0.05, *** p<0.01

Note: Based on discontinuity sample with 10 test score points bandwidth. Dependent variable is a dummy indicating retention in grade 3 in all columns. Robust standard errors in parentheses.

Table 4: The Effect of Grade Retention on Student Achievement

	Specification			
	OLS		IV	
Dependent Variable	(1)	(2)	(3)	(4)
Reading				
1 year (n = 74443)	–56.42***	–41.19***	82.62***	83.61***
	(2.077)	(2.058)	(11.072)	(10.515)
2 years (n = 59554)	61.12***	76.43***	174.7***	173.8***
	(2.295)	(2.263)	(12.281)	(11.722)
3 years (n = 45175)	–2.443	14.09***	91.37***	91.26***
	(2.702)	(2.650)	(13.079)	(12.511)
4 years (n = 35001)	–52.30***	–35.85***	41.83***	41.02***
	(2.983)	(2.934)	(14.235)	(13.546)
5 years (n = 23568)	–70.54***	–55.23***	–7.767	–8.182
	(3.198)	(3.135)	(14.466)	(13.596)
6 years (n = 12912)	–30.30***	–14.74***	14.43	14.87
	(3.878)	(3.779)	(16.325)	(15.287)
Math				
1 year (n = 74327)	1.971	47.85***	91.96***	92.48***
	(2.105)	(1.729)	(10.894)	(8.495)
2 years (n = 59354)	–55.52***	–15.26***	26.10**	23.76***
	(2.116)	(1.789)	(10.735)	(8.499)
3 years (n = 45093)	35.65***	73.82***	110.8***	110.3***
	(2.500)	(2.155)	(12.610)	(10.529)
4 years (n = 34987)	–114.2***	–76.95***	–21.90	–25.58**
	(2.888)	(2.561)	(13.475)	(11.341)
5 years (n = 23563)	–77.34***	–48.60***	–23.28*	–25.47**
	(2.818)	(2.473)	(12.767)	(10.701)
6 years (n = 12905)	–56.97***	–31.37***	–4.666	–8.586
	(3.182)	(2.796)	(13.089)	(11.073)
Performance and demographic covariates	No	Yes	No	Yes

Note: Based on discontinuity sample with 10 test score points bandwidth. Dependent variables are developmental scales scores in reading and math. Performance and demographic covariates include grade 3 reading and math scores, gender, age, race, special education status in grade 3, free- or reduced lunch in grade 3, LEP status in Grade 3 and cohort dummies. Robust standard errors in parentheses.

Table 5: The Effect of Grade Retention on Student Achievement (re-scaled)

	Specification			
	OLS		IV	
Dependent Variable	(1)	(2)	(3)	(4)
Reading				
1 year (n = 74443)	74.84***	90.07***	213.9***	214.9***
	(2.077)	(2.058)	(11.072)	(10.515)
2 years (n = 59554)	29.00***	44.76***	144.9***	143.9***
	(2.294)	(2.261)	(12.288)	(11.718)
3 years (n = 45175)	–12.90***	3.861	82.03***	81.90***
	(2.703)	(2.651)	(13.092)	(12.518)
4 years (n = 35001)	–48.92***	–32.57***	44.68***	43.90***
	(2.983)	(2.934)	(14.223)	(13.536)
5 years (n = 23568)	–57.39***	–42.45***	3.171	2.774
	(3.188)	(3.125)	(14.366)	(13.519)
6 years (n = 12912)	–73.81***	–56.91***	–22.37	–22.39
	(3.885)	(3.798)	(16.568)	(15.476)
Math				
1 year (n = 74327)	39.43***	85.30***	129.4***	129.9***
	(2.105)	(1.729)	(10.894)	(8.495)
2 years (n = 59354)	–14.83***	24.89***	64.18***	61.90***
	(2.113)	(1.788)	(10.683)	(8.475)
3 years (n = 45093)	–21.81***	17.44***	59.49***	58.87***
	(2.492)	(2.143)	(12.644)	(10.503)
4 years (n = 34987)	–81.51***	–45.10***	6.592	3.093
	(2.875)	(2.542)	(13.278)	(11.175)
5 years (n = 23563)	–70.73***	–42.27***	–17.65	–19.88*
	(2.789)	(2.447)	(12.658)	(10.605)
6 years (n = 12905)	–87.71***	–61.14***	–30.71**	–34.99***
	(3.211)	(2.836)	(13.401)	(11.346)
Performance and demographic covariates	No	Yes	No	Yes

Note: Based on discontinuity sample with 10 test score points bandwidth. Dependent variables are re-scaled developmental scales scores in reading and math. Performance and demographic covariates include grade 3 reading and math scores, gender, age, race, special education status in grade 3, free- or reduced lunch in grade 3, LEP status in Grade 3 and cohort dummies. Robust standard errors in parentheses.

Table 6: The Effect of Grade Retention in Grade 3 on Future Grade Retention

	Specification			
	OLS		IV	
Dependent Variable	(1)	(2)	(3)	(4)
Retention Probability				
2 years (n = 59493)	–.0488***	–.0615***	–.114***	–.112***
	(.001)	(.002)	(.011)	(.011)
3 years (n = 44028)	–.00767***	–.0124***	–.0325***	–.0320***
	(.001)	(.001)	(.008)	(.007)
4 years (n = 33718)	–.0234***	–.0291***	–.0380***	–.0381***
	(.002)	(.002)	(.011)	(.011)
5 years (n = 22511)	–.00129	–.00657**	–.0421***	–.0434***
	(.003)	(.003)	(.015)	(.014)
6 years (n = 12238)	.00921**	.00577	–.00205	–.00239
	(.004)	(.004)	(.014)	(.014)
Performance and demographic covariates	No	Yes	No	Yes

Note: Based on discontinuity sample with 10 test score points bandwidth. Dependent variable is a dummy indicating grade retention. Performance and demographic covariates include grade 3 reading and math scores, gender, age, race, special education status in grade 3, free- or reduced lunch in grade 3, LEP status in Grade 3 and cohort dummies. Robust standard errors in parentheses.

Table 7: The Effect of Grade Retention in Grade 3 on Grade Level

	Specification			
	OLS		IV	
Dependent Variable	(1)	(2)	(3)	(4)
Grade Level				
2 years (n = 59684)	–.945***	–.932***	–.874***	–.877***
	(.002)	(.002)	(.011)	(.011)
3 years (n = 45299)	–.921***	–.900***	–.816***	–.819***
	(.003)	(.003)	(.015)	(.015)
4 years (n = 35126)	–.884***	–.857***	–.747***	–.754***
	(.004)	(.004)	(.021)	(.020)
5 years (n = 23681)	–.860***	–.829***	–.682***	–.695***
	(.006)	(.006)	(.030)	(.029)
6 years (n = 13000)	–.854***	–.822***	–.729***	–.740***
	(.009)	(.009)	(.038)	(.037)
Performance and demographic covariates	No	Yes	No	Yes

Note: Based on discontinuity sample with 10 test score points bandwidth. Dependent variable is the grade level that students are observed in. Performance and demographic covariates include grade 3 reading and math scores, gender, age, race, special education status in grade 3, free- or reduced lunch in grade 3, LEP status in Grade 3 and cohort dummies. Robust standard errors in parentheses.

Table 8: The Effect of Grade Retention on Student Absence

	Specification			
	OLS		IV	
Dependent Variable	(1)	(2)	(3)	(4)
Days of absence				
1 year (n = 74599)	.502***	.304***	–.337	–.348
	(.081)	(.081)	(.412)	(.403)
2 years (n = 59597)	.501***	.326***	–.211	–.230
	(.093)	(.092)	(.456)	(.444)
3 years (n = 45267)	–.343***	–.485***	–1.284**	–1.350***
	(.110)	(.110)	(.522)	(.508)
4 years (n = 35101)	.274*	.0313	–.826	–.821
	(.149)	(.148)	(.714)	(.693)
5 years (n = 23659)	1.050***	.831***	.905	.634
	(.208)	(.207)	(.968)	(.936)
6 years (n = 12985)	1.785***	1.406***	–.890	–1.147
	(.304)	(.302)	(1.191)	(1.160)
Performance and demographic covariates	No	Yes	No	Yes

Note: Based on discontinuity sample with 10 test score points bandwidth. Dependent variable are the days of absence in a school year. Performance and demographic covariates include grade 3 reading and math scores, gender, age, race, special education status in grade 3, free- or reduced lunch in grade 3, LEP status in Grade 3 and cohort dummies. Robust standard errors in parentheses.

Table 9: The Effect of Grade Retention on Special Education Placement

	Specification			
	OLS		IV	
Dependent Variable	(1)	(2)	(3)	(4)
Special Ed Placement				
1 year (n = 74674)	–.0505***	.0122***	.00290	.0134
	(.004)	(.003)	(.022)	(.012)
2 years (n = 59684)	–.0367***	.0172***	.00622	.0168
	(.005)	(.004)	(.024)	(.016)
3 years (n = 45299)	–.0361***	.0114***	.0268	.0260
	(.005)	(.004)	(.025)	(.018)
4 years (n = 35126)	–.0354***	.00139	.0374	.0344*
	(.006)	(.005)	(.028)	(.021)
5 years (n = 23681)	–.0327***	–.00649	.0422	.0209
	(.007)	(.006)	(.033)	(.025)
6 years (n = 13000)	–.0252***	–.00751	.0395	.0227
	(.009)	(.007)	(.037)	(.028)
Performance and demographic covariates	No	Yes	No	Yes

Note: Based on discontinuity sample with 10 test score points bandwidth. Dependent variable is a dummy indicating special education placement. Performance and demographic covariates include grade 3 reading and math scores, gender, age, race, special education status in grade 3, free- or reduced lunch in grade 3, LEP status in Grade 3 and cohort dummies. Robust standard errors in parentheses.

Table 10: The Effect of Grade Retention on Attrition from Florida Public Schools

	Specification			
	OLS		IV	
Dependent Variable	(1)	(2)	(3)	(4)
Dropout Probability				
2 years (n = 63144)	.00531**	.00390	.00546	.00589
	(.003)	(.003)	(.012)	(.012)
3 years (n = 50149)	–.0120***	–.0133***	–.0213	–.0205
	(.003)	(.004)	(.017)	(.016)
4 years (n = 40168)	.00584	.00447	.00552	.00683
	(.004)	(.005)	(.021)	(.021)
5 years (n = 27733)	.0200***	.0193***	–.0153	–.0168
	(.006)	(.006)	(.026)	(.026)
6 years (n = 15699)	.00482	.00685	–.00642	–.0153
	(.008)	(.008)	(.033)	(.032)
Performance and demographic covariates	No	Yes	No	Yes

Note: Based on discontinuity sample with 10 test score points bandwidth. Dependent variable is a dummy indicating that a students in not observed in the Florida data in the specific school year. Performance and demographic covariates include grade 3 reading and math scores, gender, age, race, special education status in grade 3, free- or reduced lunch in grade 3, LEP status in Grade 3 and cohort dummies. Robust standard errors in parentheses.

Figure A-1: The Distribution of School-specific Shares of Retained Students in Grade 3

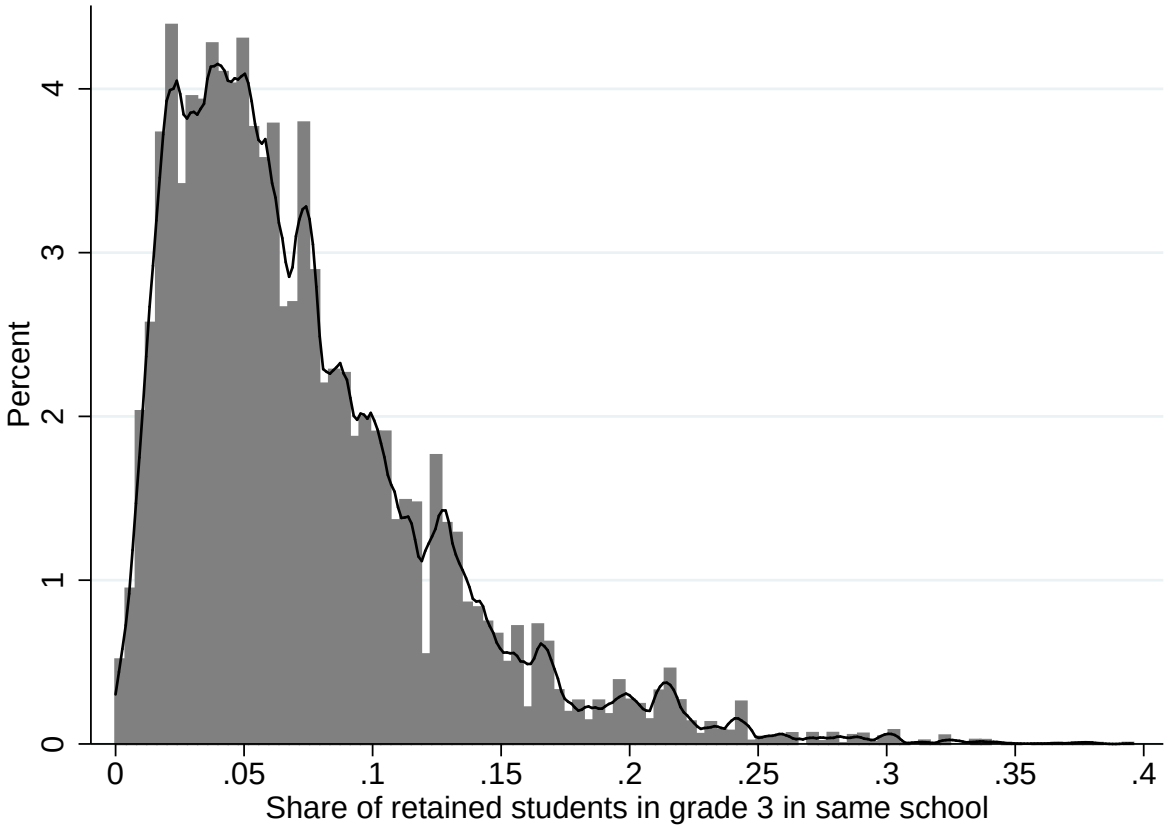


Table A-1: Achievement Results by Cohort

Cohort	2003	2004	2005	2006	2007	2008
First Stage	.373***	.268***	.295***	.338***	.198***	.217***
	(.012)	(.013)	(.013)	(.016)	(.012)	(.013)
Reading						
1 year	51.833***	85.018***	108.999***	86.842***	165.844***	19.837
	(15.827)	(27.489)	(23.463)	(27.651)	(39.213)	(32.179)
2 years	215.219***	156.784***	116.768***	158.166***	199.172***	
	(18.121)	(30.538)	(27.185)	(26.008)	(36.741)	
3 years	76.722***	145.506***	71.932***	84.065***		
	(18.222)	(32.422)	(27.415)	(26.675)		
4 years	32.625*	48.157	50.111**			
	(18.536)	(30.426)	(25.045)			
5 years	–30.827*	33.297				
	(15.788)	(25.259)				
6 years	14.867					
	(15.287)					
Math						
1 year	63.988***	114.708***	83.188***	105.680***	123.010***	85.951***
	(13.452)	(21.358)	(20.011)	(21.377)	(29.697)	(27.949)
2 years	–4.456	19.932	14.924	43.732**	77.955***	
	(12.847)	(21.129)	(19.052)	(20.215)	(28.099)	
3 years	104.492***	106.417***	109.324***	126.743***		
	(15.555)	(28.874)	(21.791)	(21.945)		
4 years	–41.076**	–36.278	4.889			
	(16.846)	(22.899)	(20.613)			
5 years	–30.928**	–16.839				
	(12.576)	(19.596)				
6 years	–8.586					
	(11.073)					
Students	15,687	12,040	12,435	9,981	12,995	11,536

Note: Based on discontinuity sample with 10 test score points bandwidth. The table displays IV estimation results with performance and demographic covariates by cohort of students. A cohort is defined by the school year students attended third grade for the first time. The first row shows first stage estimates. The last row indicates the number of students by cohort in the first stage regression.

Table A-2: Retention Results by Cohort

Cohort	2003	2004	2005	2006	2007
First Stage	.373***	.268***	.295***	.338***	.198***
	(.012)	(.013)	(.013)	(.016)	(.012)
Retention probability					
2 years	–.097***	–.176***	–.100***	–.095***	–.104***
	(.018)	(.031)	(.024)	(.024)	(.028)
3 years	–.029**	–.045**	–.039***	–.018	
	(.013)	(.021)	(.014)	(.012)	
4 years	–.048***	–.070***	.005		
	(.015)	(.025)	(.019)		
5 years	–.039**	–.049*			
	(.017)	(.026)			
6 years	–.002				
	(.014)				
Students	15,687	12,040	12,435	9,981	12,995

Note: Based on discontinuity sample with 10 test score points bandwidth. The table displays IV estimation results with performance and demographic covariates by cohort of students. A cohort is defined by the school year students attended third grade for the first time. The first row shows first stage estimates. The last row indicates the number of students by cohort in the first stage regression.

Table A-3: Achievement Results by Gender and Race

Subgroup	Girls	Boys	Whites	Blacks	Hispanics	Asian
First Stage	.279***	.288***	.248***	.314***	.287***	.175***
	(.008)	(.007)	(.010)	(.009)	(.010)	(.047)
Reading						
1 year	88.690***	78.905***	110.384***	66.408***	77.289***	167.536
	(14.799)	(14.813)	(22.426)	(15.759)	(18.805)	(137.624)
2 years	178.091***	169.678***	175.333***	169.885***	178.023***	271.012*
	(16.216)	(16.788)	(25.314)	(17.459)	(20.927)	(140.171)
3 years	107.041***	76.760***	126.553***	69.772***	83.568***	210.701
	(17.299)	(17.887)	(27.361)	(18.509)	(22.092)	(174.978)
4 years	53.623***	29.713	82.429***	13.627	34.469	89.770
	(18.848)	(19.334)	(29.272)	(20.481)	(23.313)	(268.098)
5 years	7.063	–20.562	–4.944	–47.565**	32.002	175.269
	(18.784)	(19.532)	(28.737)	(20.883)	(24.000)	(234.760)
6 years	13.876	16.189	52.323	13.227	–13.038	133.456
	(21.238)	(21.787)	(32.599)	(22.538)	(28.387)	(179.719)
Math						
1 year	102.824***	83.031***	108.085***	84.134***	83.746***	323.304**
	(12.547)	(11.527)	(16.886)	(13.450)	(15.172)	(135.165)
2 years	32.425***	15.215	20.336	42.615***	–1.400	.217
	(12.083)	(11.888)	(17.208)	(13.456)	(15.163)	(89.119)
3 years	87.661***	128.845***	133.777***	93.622***	116.005***	138.945
	(14.932)	(14.782)	(22.151)	(16.364)	(18.269)	(137.852)
4 years	–19.625	–32.878**	–20.751	–25.548	–30.594*	–36.916
	(15.637)	(16.257)	(23.159)	(18.781)	(18.495)	(204.009)
5 years	–5.947	–43.600***	–6.946	–50.821***	–11.019	–67.386
	(14.690)	(15.379)	(21.103)	(17.764)	(18.456)	(174.342)
6 years	–20.278	.233	–1.946	–6.050	–7.097	–119.075
	(14.892)	(16.193)	(23.725)	(16.834)	(19.504)	(168.593)
Students	35,023	39,651	23,507	26,510	21,151	976

Note: Based on discontinuity sample with 10 test score points bandwidth. The table displays IV estimation results with performance and demographic covariates by gender and race. The first row reports first stage estimates. The last row indicates the number of students by gender and race in the first stage regression.

Table A-4: Retention Results by Gender and Race

Subgroup	Girls	Boys	Whites	Blacks	Hispanics	Asian
First Stage	.279***	.288***	.248***	.314***	.287***	.175***
	(.008)	(.007)	(.010)	(.009)	(.010)	(.047)
Retention probability						
2 years	–.122***	–.103***	–.126***	–.125***	–.075***	–.227**
	(.015)	(.015)	(.022)	(.018)	(.017)	(.110)
3 years	–.034***	–.030***	–.045***	–.040***	–.003	–.087
	(.010)	(.011)	(.015)	(.013)	(.011)	(.060)
4 years	–.030**	–.045**	–.033	–.044**	–.038**	.107
	(.012)	(.018)	(.021)	(.019)	(.016)	(.077)
5 years	–.035**	–.050**	–.056**	–.049*	–.022	.173
	(.017)	(.022)	(.028)	(.026)	(.022)	(.237)
6 years	–.012	.007	–.032	.042*	–.026	–.025
	(.017)	(.022)	(.029)	(.024)	(.022)	(.041)
Students	35,023	39,651	23,507	26,510	21,151	976

Note: Based on discontinuity sample with 10 test score points bandwidth. The table displays IV estimation results with performance and demographic covariates by gender and race. The first row reports first stage estimates. The last row indicates the number of students by gender and race in the first stage regression.

Table A-5: Achievement Results by Math Level in Grade 3

Math Level	1	2	> 3
First Stage	.385***	.305***	.189***
	(.010)	(.009)	(.009)
Reading			
1 year	76.834***	72.878***	113.739***
	(15.863)	(16.404)	(25.477)
2 years	130.464***	189.614***	210.637***
	(17.174)	(18.135)	(29.102)
3 years	64.170***	116.371***	90.907***
	(18.286)	(19.981)	(30.224)
4 years	20.494	61.381***	16.974
	(20.326)	(21.019)	(31.890)
5 years	–7.689	–1.180	–28.230
	(20.723)	(21.034)	(32.399)
6 years	–3.987	27.339	28.637
	(22.968)	(23.578)	(38.943)
Math			
1 year	84.126***	109.235***	78.833***
	(14.545)	(12.824)	(18.280)
2 years	16.338	36.455***	7.441
	(14.633)	(12.579)	(17.795)
3 years	97.882***	137.146***	81.029***
	(17.424)	(16.418)	(22.659)
4 years	–57.259***	–21.185	5.285
	(19.341)	(17.218)	(23.411)
5 years	–36.424**	–2.601	–52.406**
	(18.100)	(16.235)	(22.388)
6 years	–5.283	–10.438	–14.632
	(18.623)	(16.029)	(25.590)
Students	20,200	25,834	28,640

Note: Based on discontinuity sample with 10 test score points bandwidth. The table displays IV estimation results with performance and demographic covariates by math achievement levels in grade 3. The first row reports first stage estimates. The last row indicates the number of students by gender and race in the first stage regression.

Table A-6: Retention Results by Math Level in Grade 3

Math Level	1	2	> 3
First Stage	.385***	.305***	.189***
	(.010)	(.009)	(.009)
Retention probability			
2 years	–.153***	–.097***	–.077***
	(.020)	(.017)	(.020)
3 years	–.031**	–.037***	–.020
	(.014)	(.011)	(.014)
4 years	–.026	–.055***	–.036*
	(.018)	(.018)	(.021)
5 years	–.024	–.038*	–.074**
	(.024)	(.023)	(.030)
6 years	.004	–.001	.006
	(.023)	(.023)	(.032)
Students	20,200	25,834	28,640

Note: Based on discontinuity sample with 10 test score points bandwidth. The table displays IV estimation results with performance and demographic covariates by math achievement levels in grade 3. The first row reports first stage estimates. The last row indicates the number of students by gender and race in the first stage regression.

Table A-7: Achievement Results by School-specific Shares of Retained Students

Share of retained students in school	1st quartile	2nd quartile	3rd quartile	4th quartile
First Stage	.189***	.241***	.279***	.343***
	(.014)	(.012)	(.010)	(.009)
Reading				
1 year	156.129***	99.433***	88.354***	57.322***
	(44.097)	(27.420)	(20.351)	(13.964)
2 years	227.664***	181.663***	191.716***	149.351***
	(49.761)	(31.017)	(23.352)	(15.345)
3 years	108.218**	86.636**	130.228***	64.455***
	(52.089)	(33.996)	(24.501)	(16.369)
4 years	36.682	41.742	65.635**	25.735
	(52.186)	(35.240)	(26.482)	(18.117)
5 years	14.336	19.769	–18.528	–18.225
	(45.647)	(38.222)	(24.373)	(19.281)
6 years	56.650	6.649	3.801	7.988
	(50.290)	(39.025)	(27.931)	(21.984)
Math				
1 year	138.255***	83.712***	90.286***	87.599***
	(34.602)	(21.419)	(16.104)	(11.757)
2 years	80.293**	–24.188	22.317	30.953***
	(35.091)	(22.468)	(16.554)	(11.484)
3 years	179.121***	39.298	109.817***	119.720***
	(42.226)	(28.466)	(20.583)	(14.106)
4 years	–14.115	–69.214**	–13.218	–18.819
	(42.471)	(28.831)	(22.059)	(15.526)
5 years	15.083	–61.810**	–9.374	–33.648**
	(36.816)	(29.059)	(18.929)	(15.459)
6 years	–8.880	–38.362	–11.881	.913
	(35.740)	(27.843)	(20.483)	(16.129)
Students	10,326	15,607	20,687	27,901

Note: Based on discontinuity sample with 10 test score points bandwidth. The table displays IV estimation results with performance and demographic covariates by gender and race. The first row reports first stage estimates. The last row indicates the number of students by gender and race in the first stage regression.

Table A-8: Retention Results by School-specific Shares of Retained Students

Share of retained students in school	1st quartile	2nd quartile	3rd quartile	4th quartile
First Stage	.189***	.241***	.279***	.343***
	(.014)	(.012)	(.010)	(.009)
Retention probability				
2 years	–.083**	–.144***	–.097***	–.118***
	(.037)	(.026)	(.022)	(.015)
3 years	.004	–.074***	–.013	–.035***
	(.023)	(.022)	(.014)	(.011)
4 years	–.008	–.075***	–.044*	–.029*
	(.037)	(.026)	(.023)	(.015)
5 years	.012	–.065*	–.034	–.057***
	(.048)	(.035)	(.028)	(.021)
6 years	.046	–.015	–.046*	.019
	(.036)	(.032)	(.028)	(.023)
Students	10,326	15,607	20,687	27,901

Note: Based on discontinuity sample with 10 test score points bandwidth. The table displays IV estimation results with performance and demographic covariates by gender and race. The first row reports first stage estimates. The last row indicates the number of students by gender and race in the first stage regression.