

Early Grade Retention Harms Adult Earnings

Jiee Zhong *

Texas A&M University

Abstract

This paper provides the first causal evidence of the effects of grade retention on the labor market outcomes using Texas' policy of retaining third graders who fail a reading test. The fuzzy regression discontinuity design estimates show that third-grade retention significantly reduces adult earnings. Although the policy aims to improve academic achievement, the results demonstrate that third-grade retention lowers high school graduation rates without improving college outcomes and aggravates absenteeism and violent behavior.

Keywords: Grade Retention, Labor Market Outcomes, Earnings Inequality

*Zhong: Department of Economics, Texas A&M University (email: jieezhong@tamu.edu). I am grateful to Andrew Barr and Jason Lindo for their invaluable guidance and insightful advice during the development of this paper. I extend my sincere appreciation to Yonghong An, Joanna Lahey, and the attendees of the 2022 Midwest Economics Association, 2022 Western Economic Association International, 2022 Association for Public Policy Analysis and Management, 2022 Southern Economics Association conferences, 2023 Eastern Economic Association Conferences, and 2023 LBJ School Policy Research Workshop, for their constructive comments and helpful suggestions. Any errors that remain in this paper are solely my responsibility.

1 Introduction

One in ten public school students in the United States has undergone grade retention from kindergarten to grade 12, with even higher rates among Black (21%) and Hispanic (12%) students (National center for education statistics, 2010). The implementation of grade retention has been the subject of debate, with proponents citing its potential to improve the academic performance of low-achieving students through additional time to acquire essential knowledge. Conversely, opponents argue that retention can negatively impact student outcomes by creating a sense of punishment and stigma. Given the prevalence and growth of grade retention practices and their unequal impact on disadvantaged populations, it is imperative to examine the effect of this policy.

The current body of evidence presents a mixed picture of grade retention's effects on short-term outcomes. While previous studies have demonstrated that retention in early grades can lead to improved educational attainment, particularly in English literacy ([Jacob and Lefgren, 2004](#); [Greene and Winters, 2007](#); [Figlio and Özek, 2020](#)), other studies have highlighted its negative impact on disciplinary issues and dropout rates ([Jacob and Lefgren, 2009](#); [Manacorda, 2012](#); [Özek, 2015](#)).

Furthermore, the short-term and long-term effects of grade retention diverge. Works that extend the short-term outcomes find that the positive effect on English literacy disappears in later grades and has no effect on high school completion ([Schwerdt et al., 2017](#)).¹ Meanwhile, the findings of [Eren et al. \(2017\)](#) indicate a negligible impact of grade retention on juvenile crime convictions, while [Eren et al. \(2022\)](#) reveals a significant increase in adult criminal convictions. In light of these conflicting results, the overall impact of this policy on student well-being remains to be determined. The mixed effects on cognitive and non-cognitive outcomes, combined with the diverging short-term and long-term consequences, highlight the need for more research to determine

¹Schwerdt et al. (2017) have examined the impact of a combination of third-grade retention and English remediation on reading scores and have found a positive effect. However, the researchers did not observe any significant effect on high school completion. It is plausible that the negative consequences of retention offset the positive impact of English remediation. Nonetheless, as English remediation was only offered to students who failed the test, the researchers found it challenging to disentangle the two effects. In contrast, my research leverages Texas's test-based retention policy to isolate the effect of retention from that of English remediation. My findings indicate a detrimental effect of third-grade retention on high school graduation, which enriches the findings of grade retention on high school completion.

the true impact of grade retention on students.

This study examines the long-term impacts of grade retention on labor market outcomes using a fuzzy regression discontinuity design (RDD). The design leverages the sharp discontinuity in the likelihood of retention generated by the Texas test-based policy, where students who barely fail their third attempt at a standardized reading test are more likely to be retained than those who barely pass. This paper provides novel perspectives on the effects of grade retention on labor market outcomes and the mechanisms driving these effects by combining administrative education and labor income data.

The results show that grade retention leads to a substantial reduction in earnings 15 years after the third grade, amounting to \$4,513 (35% relative to the control group's mean and statistically significant at the 1% level). An examination of the underlying mechanisms reveals that while grade retention decreases the risk of repeating a grade and may result in a temporary improvement in reading scores, it lowers the probability of high school graduation. There exists suggestive evidence that part of the explanation for these paradoxical effects is the increase in behavioral issues associated with grade retention. Evidence shows that retention jeopardizes the development of non-cognitive skills, including increases in absenteeism and violent behavior that persist over time. A growing body of research indicates that non-cognitive skills, rather than cognitive test scores, are stronger predictors of long-term life and career success, especially for low-achieving students ([Chetty et al., 2011](#); [Lindqvist and Vestman, 2011](#); [Heckman et al., 2006](#); [Heckman and Rubinstein, 2001](#)).

These findings suggest that, rather than helping struggling students to catch up academically, grade retention leads to more behavioral issues and lower educational attainment, leaving them even further behind as they enter the labor market. The results also reveal that the retention rates are disproportionately high for Black and Hispanic students relative to White students, even after controlling for test scores. This disparity, coupled with the significant reduction in earnings associated with grade retention, raises concerns about the potential role of this practice in exacerbating racial earnings inequality.

This study sheds new light on the effects of grade retention policies on labor market outcomes, offering valuable contributions to the existing body of literature. Previous studies have investigated the short-term impacts of third-grade retention on test scores, disciplinary issues, and high school graduation in Florida ([Schwerdt et al., 2017](#); [Özek, 2015](#); [Greene and Winters, 2007](#)). However, this paper extends these findings by exploring post-secondary education and labor market outcomes. This study aligns with the work of [Jacob and Lefgren \(2009\)](#) and [Eren et al. \(2022\)](#), which focuses on the effect of holding students back in a grade relative to promoting them. These prior studies have indicated that holding back low-achieving students in eighth grade can lead to increased dropout and crime convictions. This study supplements these findings by providing unique causal evidence of the long-term effects of early grade retention on labor market outcomes.

The findings of this paper also add to the growing body of research that emphasizes the significant impacts of early interventions on educational achievement and labor market outcomes ([Barr and Gibbs, 2022](#); [García et al., 2020](#); [Currie and Almond, 2011](#); [Chetty et al., 2011](#); [Deming, 2009](#)). These findings indicate that early childhood intervention has lasting effects. This paper furthers these findings by demonstrating the substantial impact of early grade retention on adult earnings.

2 Background

Following the No Child Left Behind Act of 2001, Texas implemented a test-based retention policy for third graders in the 2002-03 school year, replacing its prior policy of social promotion. This change aimed to improve student performance through standardized testing and minimum performance standards. Third grade marks a crucial transition from learning to read to reading to learn. Policymakers establish the grade retention policy to ensure students have a solid foundation in reading before moving forward to more advanced academic subjects.

The implementation of this policy in Texas sparked a highly debated topic among educators, policymakers, and parents, as opinions were divided on its effectiveness. Supporters of the policy believed that repeating the third grade would give additional time to develop the necessary

reading skills for success in later grades. On the other hand, opponents argued that retention negatively impacted students, including decreased academic achievement, increased dropout rates, and behavioral issues. Despite the ongoing debate, the third-grade retention policy remained in place.

Under the new policy, third-grade students must pass the Texas Assessment of Knowledge and Skills (TAKS) reading test to advance to fourth grade. The TAKS test is a standardized assessment used to measure student performance in Texas public schools. Suppose a student fails the TAKS reading test on their third attempt. In that case, they will be subject to automatic retention in the same grade for the following academic year, regardless of their performance in other subjects.

However, exceptions apply. The school would notify the student's parent or guardian of this failure and retention decision. An appeals process is available by submitting a request to the Grade Placement Committee (GPC), which typically consists of the principal, an English teacher, and a parent or guardian. The GPC members promote the student based on various academic factors, such as teacher recommendation and TAKS reading score. Parents might also request their child to repeat the third grade, even if the student had achieved passing scores in the reading assessment. These institutional details are relevant to the research design and are discussed in [Section 4](#).

3 Data

This study assesses the long-run impact of grade retention on earnings by leveraging administrative data on education and labor market outcomes from the Texas Education Research Center (ERC) located at the University of Houston. The utilization of Texas ERC data is paramount to the empirical analysis of this paper, as it offers multiple benefits to the empirical examination.

The study provides the first causal examination of the impact of grade retention on labor market outcomes by leveraging the unique data linkage between administrative education records and quarterly wage data from the Texas Workforce Commission (TWC). The primary outcome of interest is annual earnings 15 years after third grade, or around age 23. The study also considers alternative earnings measures, including 14 years after third grade and average earnings between

2018 and 2020, to assess the robustness of the estimated effect of grade retention on earnings outcomes.² Note that the TWC data only captures wages from formal employment within Texas, excluding government jobs and employment outside the state. In Section 6, I address the potential implications of these omissions and demonstrate that they do not significantly impact the estimated effect of grade retention on earnings outcomes.

Moreover, the Texas Education Research Center (ERC) has developed an exceptional data integration system, linking primary and secondary education records with post-secondary education records from Texas and other states' universities and colleges. Drawing upon this innovative linkage, this paper contributes rigorous empirical evidence on the causal relationship between grade retention and post-secondary enrollment.

The Texas ERC data also includes a comprehensive set of cognitive measures such as test scores, grade retention after grade 3, high school graduation, college enrollment, and non-cognitive measures such as absenteeism, violence, and crime.³ The latter measures are derived from disciplinary records and categorize incidents as crimes or acts of violence. Through the analysis of these rich data sources, this study investigates the effect of grade retention on intermediate outcomes, illuminating the connection between short-term effects and long-term earnings outcomes.

This paper leverages the strengths of the Texas ERC data to provide new causal evidence of the impact of third-grade retention on the labor market and to analyze the underlying mechanisms. The study focuses on three cohorts of first-time third-graders enrolled between 2002-03 and 2004-05, with the first cohort exposed to the newly implemented third-grade retention policy. The sample is restricted to students who failed the reading test twice and took the third attempt, resulting in a sample size of 37,201 students, representing approximately 4% of the third-graders enrolled in 2002-03 to 2004-05 school years.

Table 1 summarizes the characteristics of the students in the sample who scored within 8 points

²I adjusted the wages each year to the value in 2020.

³It is important to note that students who do not graduate from high school are not necessarily dropouts. Among the students who took the third attempt reading test and did not graduate from high school (43%), 6% continued their education by attending grade 12 without obtaining their high school diploma, 17% dropped out, and 16% left the Texas public school system and enrolled in private schools or schools outside of Texas.

Table 1: Summary statistics

	All		Black		Hispanic		White	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Below Cutoff	Yes	No	Yes	No	Yes	No	Yes	No
Student Characteristics								
Third Grade Reading Score	-4.15	3.41	-4.20	3.42	-4.13	3.38	-4.12	3.53
Age	8.29	8.24	8.29	8.23	8.30	8.25	8.32	8.24
Male	0.56	0.52	0.58	0.54	0.54	0.51	0.56	0.51
Eligible for Free Meals	0.65	0.62	0.74	0.70	0.65	0.64	0.45	0.38
Bilingual Program	0.17	0.17	0.00	0.00	0.29	0.28	0.01	0.00
Migrant	0.04	0.04	0.00	0.00	0.06	0.06	0.00	0.00
Limited English Proficiency	0.36	0.35	0.01	0.01	0.56	0.55	0.03	0.02
Special Education	0.07	0.06	0.07	0.05	0.07	0.05	0.13	0.12
Education Outcomes								
Third Grade Retention	0.52	0.05	0.50	0.05	0.54	0.05	0.41	0.04
Fifth Grade Reading Scores	-0.94	-0.71	-1.00	-0.79	-0.90	-0.68	-1.00	-0.74
Retention After Grade Three	0.51	0.65	0.56	0.72	0.51	0.67	0.37	0.51
High School Graduation	0.58	0.64	0.55	0.61	0.59	0.65	0.55	0.62
Dropout	0.16	0.13	0.18	0.15	0.16	0.14	0.11	0.09
Leave Texas Public Schools	0.18	0.16	0.16	0.16	0.18	0.16	0.24	0.19
Enroll Any University	0.30	0.37	0.34	0.41	0.29	0.35	0.28	0.34
Enroll Community College	0.29	0.35	0.31	0.37	0.27	0.33	0.26	0.32
Enroll Public University	0.06	0.09	0.07	0.11	0.05	0.08	0.06	0.07
Enrolled Community College Age	19.01	18.60	19.14	18.79	18.96	18.51	18.89	18.65
Enrolled Public University Age	19.96	19.75	19.76	19.23	19.98	20.05	20.18	19.75
Behavioral Outcomes								
Days absent within 9 years after grade 3	84.98	81.55	87.36	82.75	85.56	83.47	80.82	76.42
Violent behaviors within 9 years after grade 3	0.93	0.78	1.39	1.15	0.82	0.72	0.46	0.45
Crimes within 9 years after grade 3	0.38	0.34	0.35	0.35	0.42	0.38	0.23	0.20
Labor Market Outcomes								
Experience 6-15 Years Later(Quarter)	15.27	16.14	15.61	16.76	15.03	15.96	16.14	16.29
Earnings 15 Years Later(\$)	11902	13021	9558	10583	12578	13665	13974	14642
Earnings 14 Years Later(\$)	10562	11450	8213	9030	11248	12274	12726	12396
Avg Earnings 2018-20(\$)	11525	12644	9125	10307	12258	13374	13509	13769
Observations	11529	12339	3108	3011	6924	7418	1297	1664

Notes: The sample includes three cohorts of first-time third-grade students who took the third-attempt reading test between the 2002-03 and 2004-05 school years and scored 0-8 points to the promotion cutoff. School districts record students' characteristics while they enroll in a grade, which would not be affected by the reading test.

of the third-attempt reading test cutoff. I chose the bandwidth of 8 points to the cutoff based on the maximum bandwidth from all regressions in the main analysis. This sample encompasses 60% of Hispanic students, 26% of Black students, and 12% of White students. While students around the cutoff are similar in pre-determined characteristics, those who score below the cutoff are significantly more likely to be held back in third grade and to earn lower wages in their twenties. The subgroup statistics reveal that the retention rates in third grade are higher for Hispanic and Black students than White students, even when controlling for test scores.

4 Empirical Strategy

4.1 Fuzzy Regression Discontinuity Design

The test-based retention policy in Texas, which allows students who fail the third-attempt reading test to avoid retention through parental appeal and agreement by the GPC members, results in a selection bias in retention associated with parental involvement. However, the promotion cutoff generates a quasi-random variation in the likelihood of third-grade retention, given that the cutoff score for passing the reading test fluctuates based on the difficulty level. This random variation generates a natural experiment that allows this paper to address the selection bias and estimate the causal effects of grade retention on student outcomes. To accomplish this, I employ a fuzzy regression discontinuity framework that leverages the quasi-random variation in the probability of being retained around the promotion cutoff.

First, I present the reduced-form model, which estimates the effect of falling below the cutoff of the third-attempt reading test on students' outcomes. The reduced-form model is expressed as follows:

$$Y_i = \alpha_0 + \alpha_1 FailTest_i + \alpha_2 score_i + \alpha_3 FailTest_i * score_i + \lambda X_i + \epsilon_i \quad (1)$$

where the dependent variable Y represents the educational and earnings outcomes. The key explanatory variable, $FailTest$, is an indicator that takes a value of 1 if the student fails the third-

attempt reading test. The primary parameter of interest, α_1 , captures the estimated effect of failing the test on the students' outcomes. The running variable, *score*, denotes the third-attempt reading test scores minus the cutoff. The vector of control variables, X , includes demographic characteristics of the student such as gender, eligibility for free meals, and cohort fixed effects.

Next, I estimate the causal effect of grade retention on students' educational and earnings outcomes using a fuzzy regression discontinuity design. In this framework, failing the third attempt reading test is an instrumental variable for being retained in third grade. The estimation is performed through a two-stage least squares (2SLS) format, as described below:

$$Retained_i = \theta_0 + \theta_1 FailTest_i + \theta_2 score_i + \theta_3 FailTest_i * score_i + \gamma X_i + \varepsilon_i \quad (2)$$

$$Y_i = \beta_0 + \beta_1 \widehat{Retained}_i + \beta_2 score_i + \beta_3 FailTest_i * score_i + \eta X_i + \mu_i \quad (3)$$

where *Retained* is an indicator variable with a value of 1 indicating retention in third grade. *FailTest*, which captures the failure of the third-attempt reading test, serves as an instrumental variable for *Retained*. Y denotes the outcomes of interest, such as educational and earnings outcomes. $\widehat{Retained}$ is the estimate of *Retained* in equation 1. The parameter of interest, β_1 , measures the estimated effect of third-grade retention on earnings or educational outcomes. The definitions of other variables are similar to those in equation 1. The standard errors are clustered at the school level to account for similarities between students within the same school. This clustering follows the approaches in [Eren et al. \(2022\)](#) and [Kolesár and Rothe \(2018\)](#). The bandwidth for each regression is selected based on optimizing the mean squared error, as in [Calonico et al. \(2017\)](#). The results are robust to alternative bandwidth choices, as demonstrated in section 6.

The primary assumption for identifying β_1 is the fundamental premise of the regression discontinuity design, which states that the expected outcomes for students close to the promotion cutoff are continuous in the absence of third-grade retention. This assumption implies that students' characteristics move smoothly across the promotion cutoff. While testing this assumption

directly without knowledge of the outcomes in the absence of grade retention is impossible, the continuity of observed characteristics across the cutoff can serve as a validity check for this assumption. An additional assumption is that the impact of failing the third attempt reading test on earnings outcomes is solely through retention. Furthermore, failing the third attempt reading test should correlate with third-grade retention. In Section 4.2, I provide a comprehensive examination of the validity of these assumptions.

4.2 Validity checks

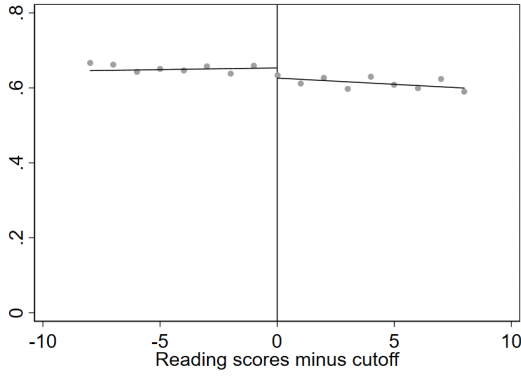
Figure 1 lends credible visual support to the first assumption that, in the absence of third-grade retention, the expectation of potential outcomes for students near the promotion cutoff is continuous. The figure demonstrates that students who scored below the cutoff have similar observed characteristics, such as free meals, gender, race, special education, and English proficiency, to those who scored above the cutoff within a bandwidth of eight points from the cutoff. Table 2 provides statistical evidence to support the graphical evidence, with small and insignificant estimates at the 5% level.

Table 2: Balance test

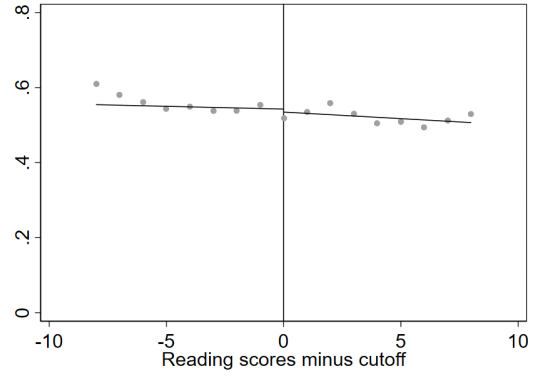
	(1)	(2)	(3)	(4)	(5)
	Free Meal	Male	Special Edu.	LEP	Exit Public School
RD Estimate	0.029	0.050*	0.004	0.004	0.021
	(0.019)	(0.027)	(0.010)	(0.019)	(0.017)
Control group mean	[0.62]	[0.52]	[0.06]	[0.35]	[0.16]
Observations	17,404	11,646	14,536	19,909	11,646

Note: This table shows the reduced-form estimate (α_1 in equation 1) of failing the third-attempt reading test on predetermined characteristics. These predetermined variables are observed before the first-attempt reading test. LEP represents limited English proficiency. Students exiting public schools might enroll in private schools or schools in other states. The sample includes three cohorts of first-time third-grade students who took the third-attempt reading test from 2002-03 to 2004-05 school years. The bandwidth for each regression is chosen by optimizing the mean squared error, as introduced in [Calonico et al. \(2017\)](#). The standard errors are clustered at the school level to account for similarities between students within the same school and are reported in parentheses.

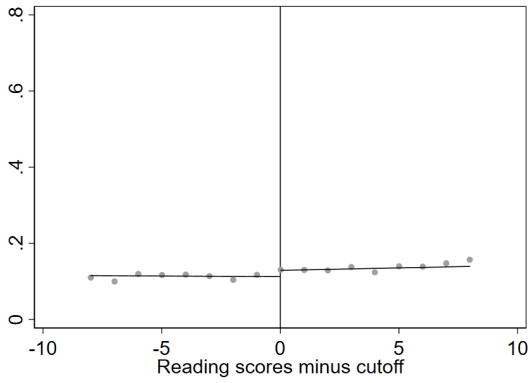
Figure 1: Balance Test



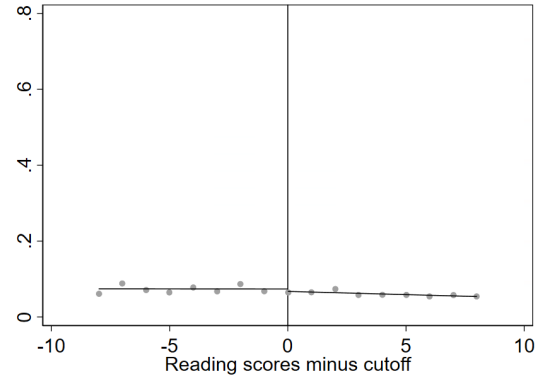
(a) Free meals



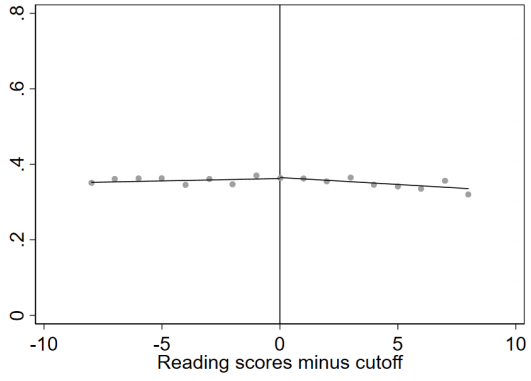
(b) Male



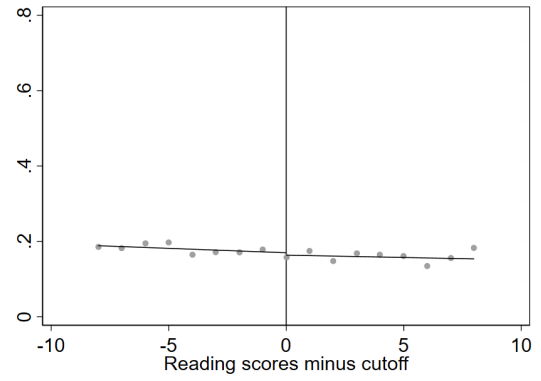
(c) White students



(d) Special Education



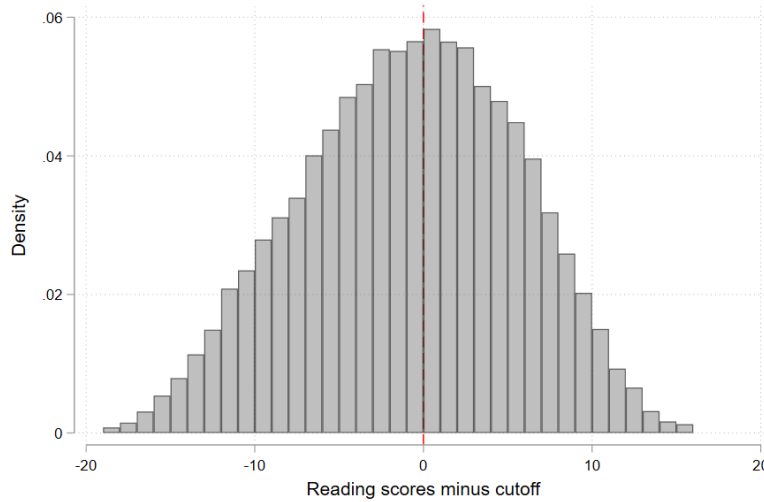
(e) Limited English Proficiency



(f) Exit Texas Public School

Notes: These figures show the reduced-form estimate of failing the third-attempt reading test on predetermined characteristics using α_1 from equation (1).

Figure 2: The distribution of the third-attempt reading test scores



Notes: This figure displays the distribution of the third-attempt reading test scores.

Non-random sorting around the cutoff is a potential threat that could challenge the first assumption. However, the likelihood of sorting is low for several reasons. Firstly, students are less likely to manipulate their scores to match the cutoff since it varies with test difficulty. Secondly, TAKS is a state-standardized test that is less likely to be manipulated by teachers. Thirdly, the distribution of third-attempt reading test scores across the cutoff, as displayed in Figure 2, shows a smooth progression without any signs of heaping that would constitute evidence of sorting (McCrary, 2008; Barreca et al., 2016). Taken together, Figures 1 and 2 provide strong evidence to support the validity of the first assumption.

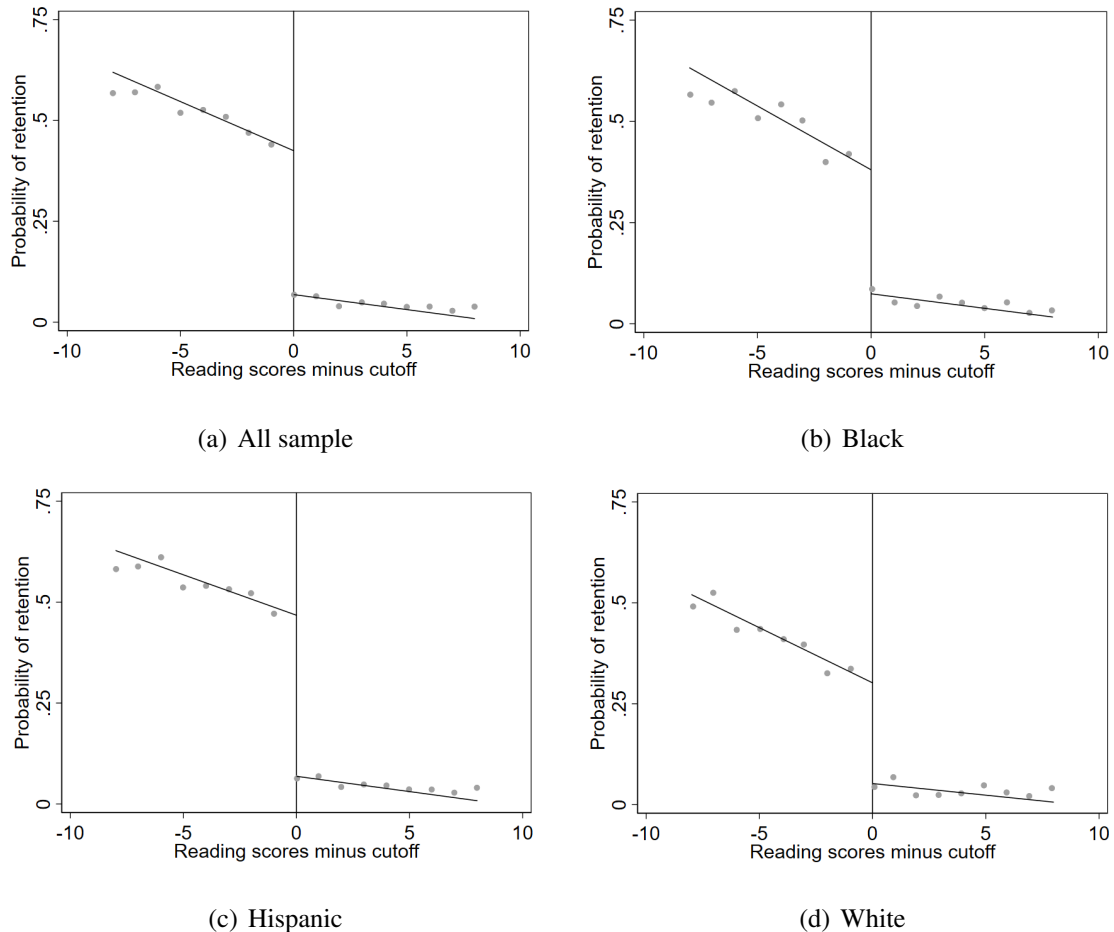
The validity of the second assumption, which posits that the instrumental variable, failing the third-attempt reading test, only impacts earnings and educational outcomes via retention in the third grade, is subject to potential challenges. One such challenge is that failing the test may disproportionately increase the likelihood of students leaving Texas public schools. They could thus improve their outcomes by attending private schools within Texas or other states. However, Figure 1 demonstrates that failing the third-attempt reading test has no significant impact on the likelihood of leaving Texas public schools.

The validity of the third assumption that failing the third-attempt reading test is strongly associated with third-grade retention is confirmed by Table 3. This table shows that the first-stage F-statistic, measuring the predictive power of the instrumental variable, is large, with a value exceeding 100. This indicates that failing the third-attempt reading test strongly predicts being retained in the third grade.

5 Results

5.1 Effects on Labor Market Outcomes

Figure 3: The probability of being retained in grade 3



Notes: This figure displays the likelihood of repeating grade 3 overall and by race and ethnicity as a function of the third-attempt reading test scores around the promotion cutoff. The bandwidth is 8, which is the maximum bandwidth from all the regressions in the main results. Each regression controls for gender, eligibility for free meals, and cohort fixed effect.

Figure 3 depicts the relationship between the likelihood of retention and reading scores at the third-attempt test in the vicinity of the promotion cutoff. This figure reveals two key insights. Firstly, there is a substantial rise in the probability of retention for students who scored below the promotion cutoff. This conclusion is substantiated by the statistical analysis provided in Table 3,

where column (1) shows that the failure of the third-attempt reading test increases the likelihood of repeating third grade by 34.2 percentage points, estimated at a 1% significance level. This effect size accounts for 66% of the average retention rates among students who scored 1 to 8 points below the promotion cutoff.

Table 3: Failing the third-attempt reading test on third-grade retention

	(1)	(2)	(3)	(4)
	Outcome: Retention in grade 3			
	All	Black	Hispanic	White
Fail the third-attempt reading test	0.342*** (0.018)	0.285*** (0.035)	0.398*** (0.020)	0.246*** (0.041)
Mean	[0.52]	[0.50]	[0.54]	[0.41]
F-statistics	1354	235	1037	100
Observations	22,070	5,620	13,315	2,727

Notes: This table shows the first-stage result of failing the reading test on the third attempt (the instrumental variable) on the likelihood of being retained in grade 3 (endogenous variable) overall and by race and ethnicity. Each regression controls for gender, eligibility for free school meals, and cohort fixed effects. The bandwidths are 0–8 test-score points to the cutoff. The standard errors are clustered at the school level to account for similarities between students within the same school and are reported in parentheses.

The second insight from Figure 3 is the disparity in retention rates among racial and ethnic minority groups. The results, presented in Table 3, columns (2) to (4), suggest that failing the third-attempt reading test has a larger impact on the likelihood of retention for Hispanic and Black students compared to White students, even when controlling for test scores. The most significant effect is observed for Hispanic students at 39.8 percentage points, Black students at 28.5 percentage points, and White students at 24.6 percentage points, all estimated at a significance level of 1%.

Significantly, holding students back in third grade substantially reduces earnings. Table 4 displays the fuzzy RDD estimates of grade retention’s effects on earnings outcomes. The baseline estimate in column (1) shows that grade retention substantially reduces earnings 15 years after grade 3 (around age 23) by \$ 4,384. Column (2), the preferred model, adds demographic characteristics, including gender, eligibility for free meals, and cohort fixed effects. The result in column (2) demonstrates that grade retention significantly reduces earnings 15 years after grade 3 by \$ 4,513 at 1% level. This level effect accounts for 35% of the average earnings 15 years after grade

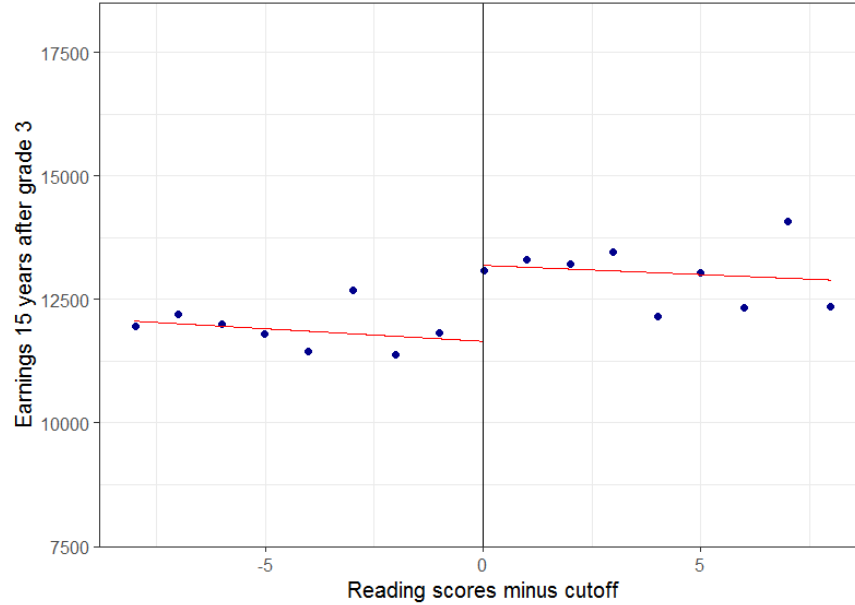
3 among students who scored 0 to 8 points above the cutoff of the third-attempt reading test.

Table 4: Grade Retention Reduces Earnings

Outcomes	Earnings 15 years later		Earnings 14 years later		Ave.Earnings 2018-20	
	(1)	(2)	(3)	(4)	(5)	(6)
Third-grade retention	-4384*** (1426)	-4513*** (1401)	-3,667*** (1,207)	-3,826*** (1,221)	-3,125** (1,224)	-3,200*** (1,217)
Control group mean	[13,021]	[13,021]	[11,450]	[11,450]	[12,644]	[12,644]
Observations	19909	19909	22070	22070	19909	19909
Demographic controls	No	Yes	No	Yes	No	Yes

Notes: This table displays the fuzzy RDD estimate of the effect of grade retention on earnings outcomes. The sample includes three cohorts of first-time third-grade students who took the third-attempt reading test from 2002–03 to 2004–05 school years. Odd-numbered columns report baseline estimates. Even-numbered columns control for gender, eligibility for free meals, and cohort fixed effects. Values with square brackets represent the mean outcomes among students whose scores are 0–8 points above the cutoff (or the control group). The bandwidth for each regression is chosen by optimizing the mean squared error, as introduced in [Calonico et al. \(2017\)](#). The bandwidth for earnings 15 years later differs from that for earnings 14 years later, which explains the difference in the observations. The standard errors are clustered at the school level to account for similarities between students within the same school and are reported in parentheses.

Figure 4: Failing the third attempt reading test reduces earnings 15 years later



Notes: This figure shows the reduced-form estimate (α_1 from equation 1) of failing the third-attempt reading test on average earnings 15 years after grade 3 (around age 23) after controlling for gender, free meals, and cohort fixed effect. The blue dots are cell means, and the lines are fitted values from a first-order polynomial regression. The top of each sub-figure indicates the estimates from the fuzzy RDD (β_1 from equation 3) and reduced-form model (α_1 from equation 1).

The visual representation of the relationship between third-attempt reading test scores and average earnings 15 years after grade 3, as depicted in Figure 4, provides compelling evidence of the significant reductions in earnings. A notable decline in earnings is observed for individuals who narrowly failed the reading test on their third attempt, as indicated on the left of the cutoff. The results from Table 4 and Figure 4 jointly illustrate the substantial and detrimental effect of grade retention on earnings 15 years later, highlighting the imperative need for a reevaluation of current retention policies.

The observed reduction in earnings is consistent across multiple measures of earning outcomes, as evidenced in columns (3) to (6) of Table 4. The findings indicate that grade retention significantly negatively impacts earnings not only 15 years after grade 3 but also 14 years after grade 3 and the average earnings between 2018 and 2020. Although the magnitude of these effects may be smaller than that reported in column (2), substantial reductions in earnings are still observed.

Table 5 provides evidence of the detrimental impact of grade retention on earnings 15 years after grade 3, differentiating results by gender and race/ethnicity. Results show that for Black and Hispanic students, grade retention leads to a statistically significant reduction in earnings at the 5% level, with reductions of \$5,846 and \$3,189, respectively. Evidence for a similar effect among White students is inconclusive, with a large reduction of \$-6,322 that is not statistically significant at the 10% level. Additionally, male and female students experience statistically significant reductions in earnings due to grade retention, with male students experiencing a slightly larger effect (\$5,361) compared to female students (\$3,121).

Table 5: Heterogeneous Effect of Grade Retention on Earnings 15 Years after Grade 3

	(1) All	(2) Black	(3) Hispanic	(4) White	(5) Female	(6) Male
Baseline Estimate	-4384*** (1426) [13,021]	-5846** (2607) [10,583]	-3189** (1610) [13,665]	-6322 (6445) [14,642]	-3121** (1573) [10,302]	-5361** (2288) [15,503]
Observations	19909	5620	13315	2727	8092	11827

Notes: This table displays the baseline estimate of the effect of grade retention on earnings outcomes by race and ethnicity as well as by gender. The sample includes three cohorts of first-time third-grade students who took the third-attempt reading test from 2002–03 to 2004–05 school years. The bandwidth for each regression is chosen by optimizing the mean squared error, as introduced in [Calonico et al. \(2017\)](#). The standard errors are clustered at the school level to account for similarities between students within the same school and are reported in parentheses.

In conclusion, the findings in this study highlight the harmful effects of grade retention on future earnings and disproportionate retention rates among minority students, which further exacerbate racial and ethnic earnings inequality. These results suggest the need for a reconsideration of current grade retention policies.

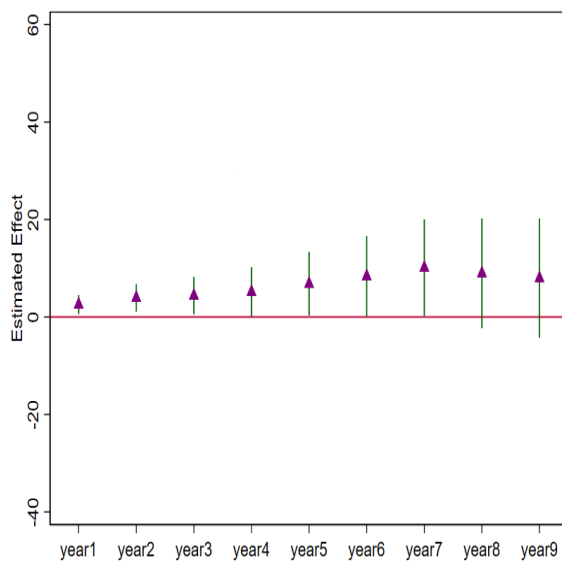
5.2 Effects on Educational Attainments

The impacts of grade retention on earnings outcomes are documented above, but the underlying mechanism remains unclear. Advocates claim retention helps by giving students more time to gain

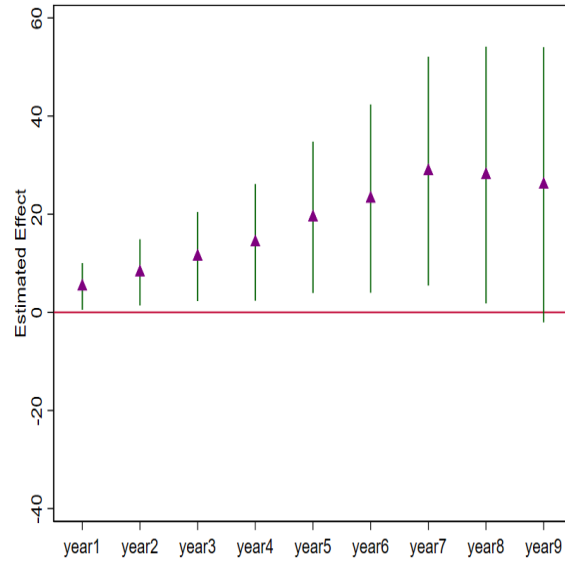
the cognitive skills needed to succeed in later grades. At the same time, critics argue that retention can be a penalty and result in a stigmatized perception among students. This section examines the effect of grade retention on intermediate outcomes and identifies how these intermediate effects contribute to reductions in earnings.

This study finds that grade retention adversely affects non-cognitive skills, such as school attendance and violent behavior. As seen in Figure 5, holding students back in the third-grade results in a significant increase in cumulative days absent from class in subsequent years, with 2.6 days increase observed one year after grade 3 (at a 5% significance level). This adverse effect persists over the seven years after grade 3, with the total number of days absent reaching its maximum of 10.1 days. Moreover, the subgroup analysis shows clear evidence of such adverse effects for Black students.

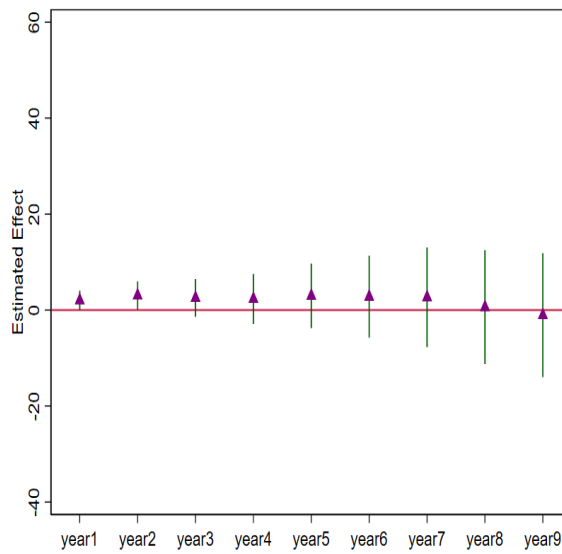
Figure 5: Effect of grade retention on days absent



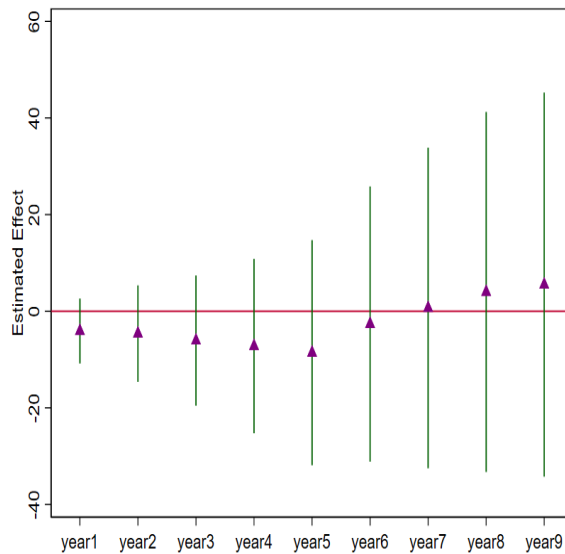
(a) Full sample



(b) Black



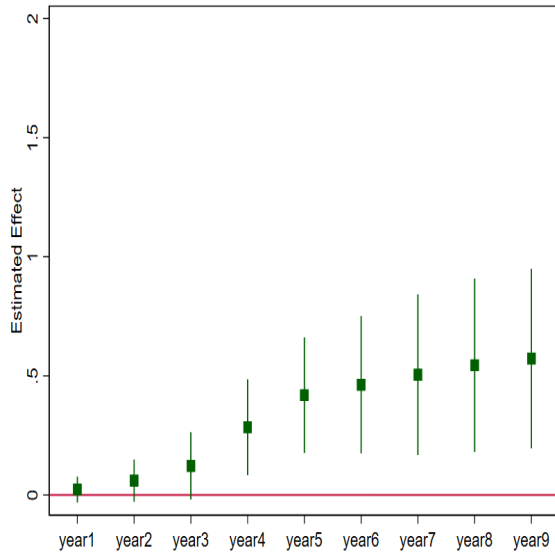
(c) Hispanic



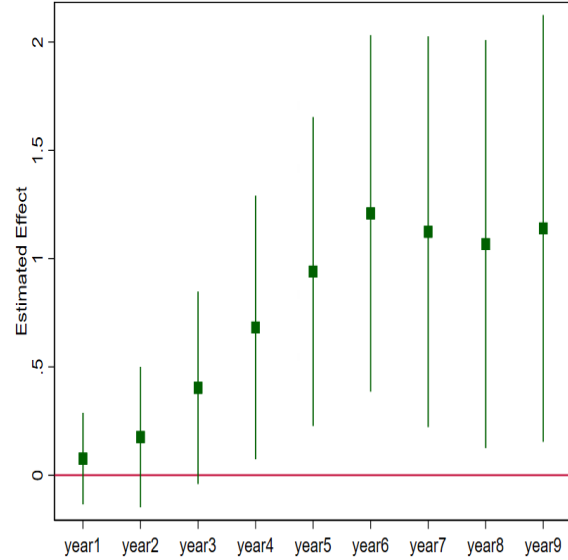
(d) White

Notes: These figures display the fuzzy RDD estimates of the effect of grade retention on the cumulative days of absence in the years after grade 3 and by race/ethnicity.

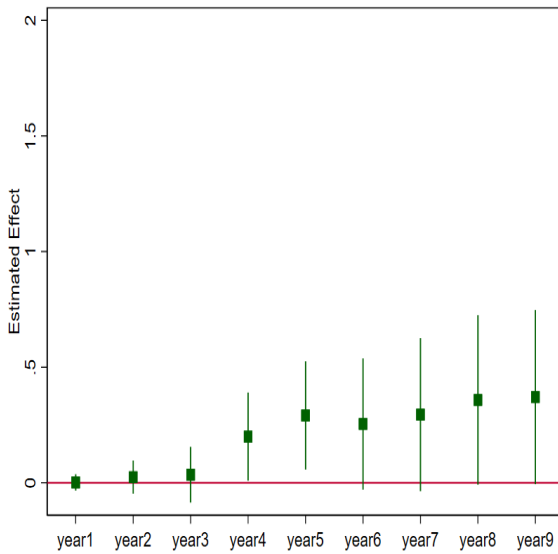
Figure 6: Effect of grade retention on violent behaviors



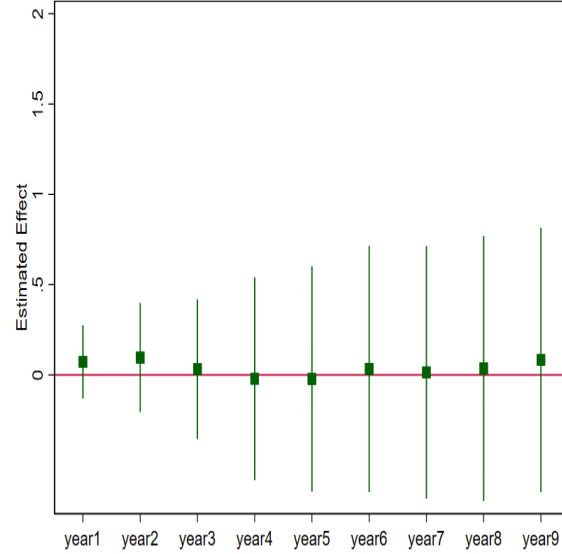
(a) Full sample



(b) Black



(c) Hispanic



(d) White

Notes: These figures show the fuzzy RDD estimate of how grade retention affects cumulative violent behaviors in the years after the third grade and by race/ethnicity.

Additionally, Figure 6 displays the effect of grade retention on violent behaviors in subsequent years after grade 3. Cumulative violent behaviors increase by 0.3 times starting four years after grade 3, representing a 100% increase relative to those who scored 0 to 8 points above the cut-

off. In addition, this adverse effect persists over later years, doubling in nine years after grade 3. In addition, the evidence of such adverse effects on violent behaviors is clear among Black and Hispanic students.

On the other hand, grade retention positively impacts some cognitive skills, including reducing the likelihood of being retained after grade 3 and increasing reading scores. Figure 7 demonstrates that retaining students in third grade significantly decreases the likelihood of being held back again in the future, reducing it by 0.552 (85%) at a 1% significance level, as supported by Table 6.

Table 6: Effect of retention on education attainments after grade 3

	(1) Future retention	(2) Grade 12	(3) HS graduation	(4) Any enrollment	(5) Community college.	(6) Public University
All	-0.552*** (0.059)	-0.088** (0.041)	-0.081* (0.043)	0.029 (0.041)	0.012 (0.040)	-0.013 (0.021)
Observations	19909	19909	19909	19909	19909	22070
Black	-0.588*** (0.132)	-0.166* (0.100)	-0.165 (0.101)	-0.089 (0.096)	-0.082 (0.093)	-0.060 (0.058)
Observations	5052	5052	5052	5052	5052	5052
Hispanic	-0.551*** (0.066)	-0.045 (0.045)	-0.003 (0.047)	0.044 (0.044)	0.031 (0.044)	-0.004 (0.024)
Observations	13315	13315	13315	13315	13315	12022
White	-0.578*** (0.195)	-0.154 (0.156)	-0.351** (0.162)	0.260* (0.153)	0.161 (0.146)	0.104 (0.083)
Observations	2727	2961	2961	2961	2961	2727

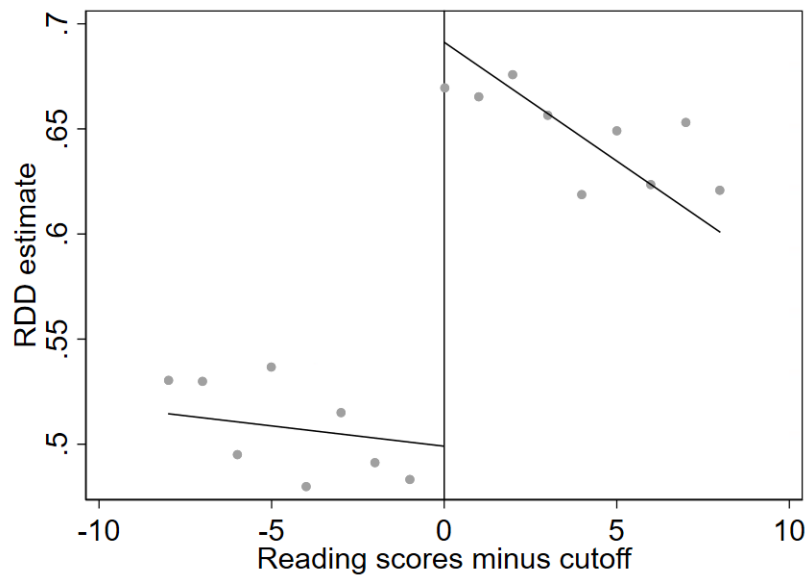
Notes: This table shows the fuzzy RDD estimate of grade retention's effect on education attainment after grade 3. The sample includes three cohorts of first-time third-grade students who took the third-attempt reading test from 2002–03 to 2004–05 school years. Each regression controls for gender, free lunch eligibility, and cohort fixed effect. The bandwidth for each regression is chosen by optimizing the mean squared error, as introduced in Calonico et al. (2017). The standard errors are clustered at the school level to account for similarities between students within the same school and are reported in parentheses.

Additionally, Figure 8 suggests that grade retention positively affects reading scores when both the retained and promoted students reach the same grade. However, this effect tends to diminish over time.⁴ Moreover, it's essential to consider that the results may be imprecise due to variations in the test difficulty and peer composition among cohorts. Furthermore, the potential gain in reading scores could reflect the age of testing rather than an actual improvement in understanding, as

⁴I standardize the reading scores with a mean of zero and a standard deviation of one by cohort.

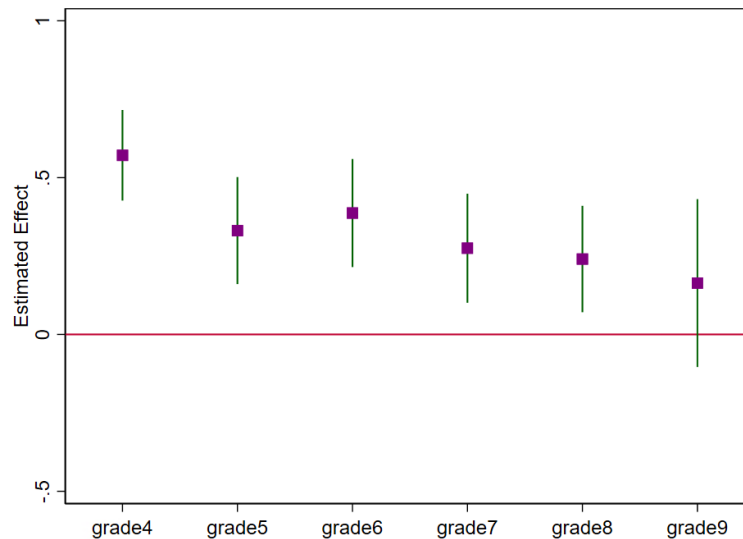
older students tend to score higher on tests (Elder and Lubotsky, 2009).

Figure 7: The effect of retention on total times of retention after grade 3



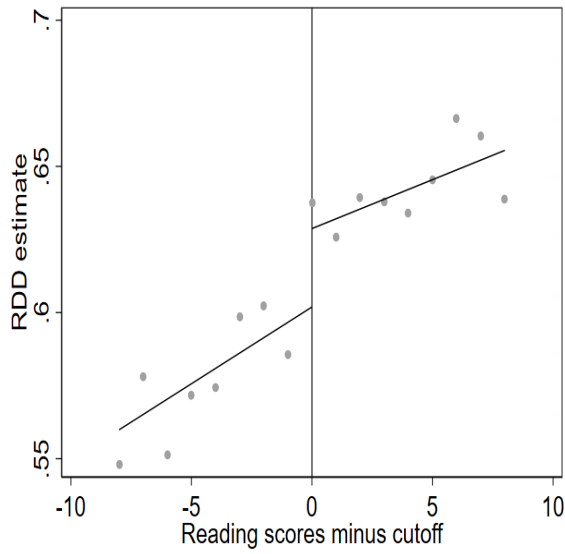
Notes: This figure displays the reduced-form estimate (α_1 from equation 1) of failing the third-attempt reading test on the times of being retained after grade 3.

Figure 8: The effect of retention on reading scores

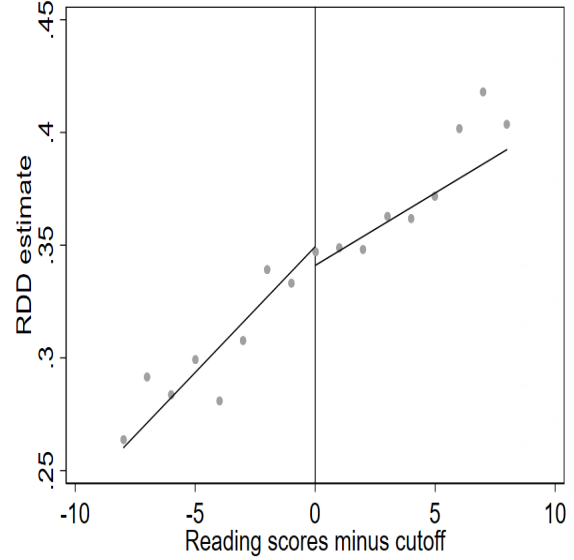


Notes: This figure displays the reduced-form estimate (α_1 from equation 1) of failing the third-attempt reading test on reading scores in subsequent grades after grade 3.

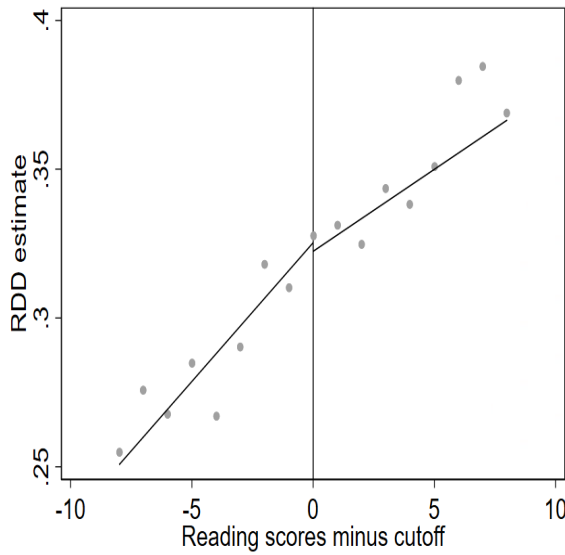
Figure 9: Effect of grade retention on education attainments



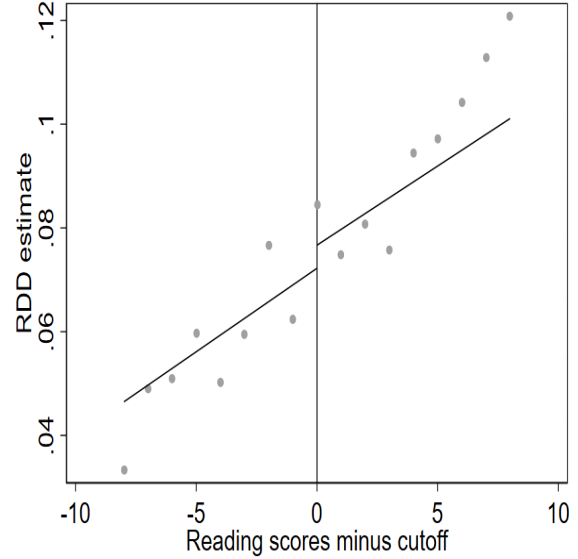
(a) High school graduation



(b) Enrollment in community or public university



(c) Enrollment in community college



(d) Enrollment in public university

Notes: These figures show the reduced-form estimates of the effect of third-grade retention on high school graduation and college enrollment.

Nonetheless, any initial cognitive improvement appears to diminish over time, resulting in lower educational attainment as evidenced by the negative impact on high-school graduation, as depicted in Figure 9. Table 6 shows that third-grade retention significantly decreases the proba-

bility of high school graduation by 8.1 percentage points, with a significance level of 10%. While post-secondary education has been found to have a critical effect on earnings, grade retention does not reduce earnings outcomes by discouraging retained students from enrolling in college. Figure 9 shows that third-grade retention does not significantly affect community college or public university enrollment.⁵

I use back-of-the-envelope calculations to provide suggestive evidence on the contributions of different intermediate outcomes to the observed reductions in earnings. To do so, I multiply the causal effect of grade retention on each intermediate outcome by the conditional correlation between that intermediate outcome and earnings 15 years after grade 3 in the control group (i.e., individuals who scored 0 to 8 points from the cutoff of the third-attempt reading test). The results, shown in online Appendix Table A7 (Panel C), indicate that the adverse effects of grade retention on violent behaviors and high school graduation together explain about 21% of the reductions in earnings 15 years after grade 3.

In addition to the intermediate outcomes estimated in our previous analysis, grade retention may also have a consequential impact on earnings through various means. Although grade retention does not notably affect the timing of initial earnings and cumulative work experience, it significantly diminishes the first annual earnings following the age of 18, which is the legal age for dropping out of school in Texas. Table 7 provides statistical evidence for the statement above, which defines the first wages earned after the age of 18, provided they surpass the amount of \$200 per year. Figures 10 to 12 provide visual evidence of these statistics.

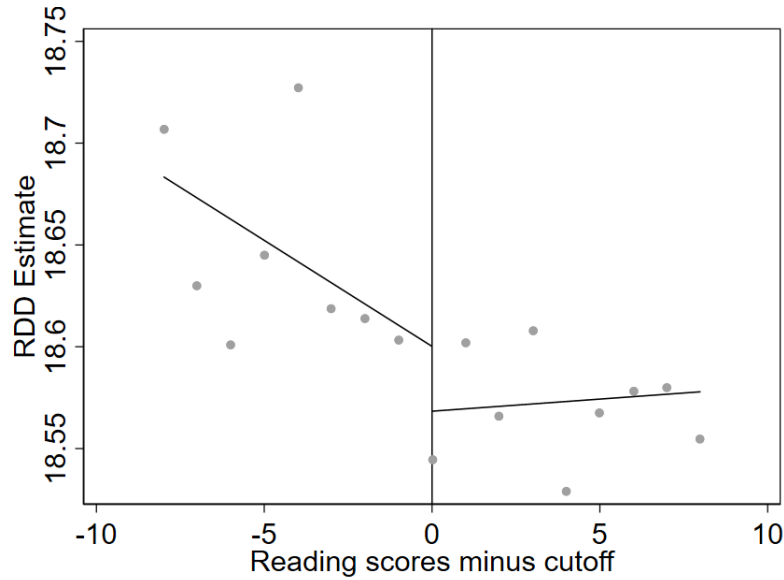
⁵About 30% students enrolled in a community college or public university, with over 90% enrolling in community colleges. The average age of enrollment was 19. The enrollment age, combined with the null effect on college enrollment, suggests that college enrollment does not disproportionately increase missing earnings for students who scored below the cutoff.

Table 7: Effect of Grade Retention on First Earnings and Work Experience

	Age at the first wages (>18)		The first wages (age>18)		Experience 6-15 years later	
	(1)	(2)	(3)	(4)	(5)	(6)
Fuzzy RDD Estimate	0.02 (0.13)	0.05 (0.13)	-1573.96** (730.91)	-1654.16** (730.17)	-1.32 (0.95)	-1.29 (0.96)
Control group's mean	[18.57]	[18.57]	[7194]	[7194]	[16.14]	[16.14]
Observations	16738	16738	19909	19909	19909	19909
Add Covariates	No	Yes	No	Yes	No	Yes

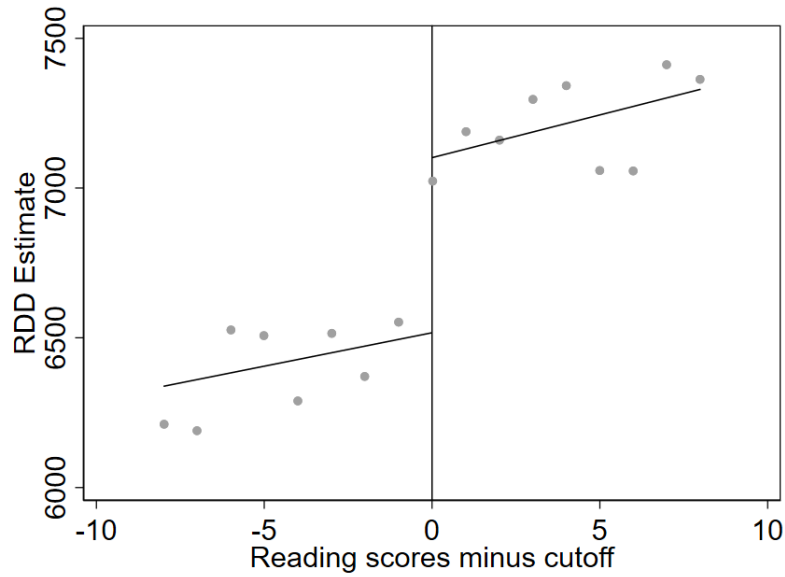
Notes: This table displays the fuzzy RDD estimate of the effect of grade retention on the age of first wages, first wages, and total work experience between 6 to 15 years after grade 3. The sample includes three cohorts of first-time third-grade students who took the third-attempt reading test from 2002–03 to 2004–05 school years. The bandwidth for each regression is chosen by optimizing the mean squared error, as introduced in Calonico et al. (2017). The standard errors are clustered at the school level to account for similarities between students within the same school and are reported in parentheses.

Figure 10: Effect of failing the third-attempt reading test on the age of the first earnings



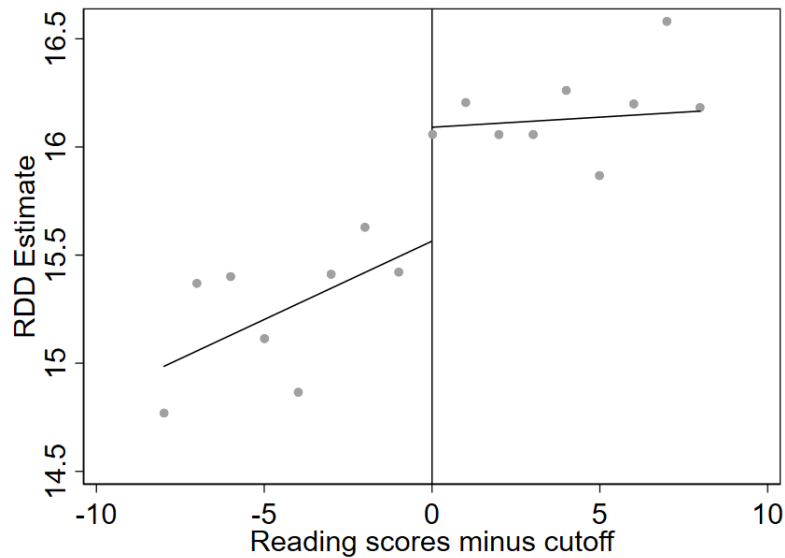
Notes: This figure displays the reduced-form estimate (α_1 from equation (1)) of failing the third-attempt reading test on the age of the first earnings. I define the first earnings as the initial wages that students earned after the legal dropout age of 18, provided that the annual wages surpass \$200.

Figure 11: Effect of failing the third-attempt reading test on the first earnings



Notes: This figure displays the reduced-form estimate (α_1 from equation (1)) of failing the third-attempt reading test on the first wages that students earned after the legal dropout age of 18 and the wages are greater than \$ 200.

Figure 12: Effect of failing the third-attempt test on work experience



Notes: This figure displays the reduced-form estimate (α_1 from equation (1)) of failing the third-attempt reading test on total work experience between 6 to 15 years after grade 3. The unit of work experience is a quarter.

Research has also shown that grade retention in 8th grade can lead to an increased likelihood of adult criminal convictions (Eren et al., 2022), which are strongly associated with low income (Looney and Turner, 2018; Craigie et al., 2020).⁶

6 Robustness Checks

In this section, I examine the potential limitations of the data and empirical strategy used to estimate the effect of grade retention on earnings. The TWC data only captures earnings information for individuals employed within Texas, raising the possibility that the effect of grade retention on earnings may be biased if it disproportionately leads to out-of-state employment. To examine this, I estimate the effect of grade retention on the likelihood of leaving Texas, as shown in Column (5) of Table 2. The results provide no evidence that grade retention increases the likelihood of leaving Texas public schools.⁷

Additionally, the out-of-state migration rate in Texas is low, with only 2% of residents reporting tax returns outside the state based on Internal Revenue Service Statistics of Income data for 2018-2019. Table 8 further supports the robustness of our results by showing no evidence of migration driving the effect of grade retention on earnings outcomes, even when excluding 10% and 20% of students from counties with high out-of-state migration rates.

⁶Specifically, approximately 49% of ex-prisoners earn less than \$500 in the first full year after release (Looney and Turner, 2018). Furthermore, people convicted of a crime experience reductions in their subsequent annual earnings, ranging from 22% to 50% of their prior annual earnings (Craigie et al., 2020).

⁷It is important to note that students who leave Texas public schools might attend private schools within Texas or attend schools in another state.

Table 8: Exclusion of Top 10% and 20% Students from High Out-of-State Migration Counties

	(1)	(2)
	Earnings 15 or 14 years after grade 3	
Outcomes	15 (age 24)	14 (age 23)
Panel A: Results in table 4	-4,384***	-3,667***
	(1,401)	(1,221)
Observations	22,070	22,070
Panel B: Exclude the top 10% students	-4,011***	-3,380***
	(1,520)	(1,271)
Observations	18,339	18,339
Panel C: Exclude the top 20% students	-3,763**	-3,570***
	(1,570)	(1,326)
Observations	16,471	16,471

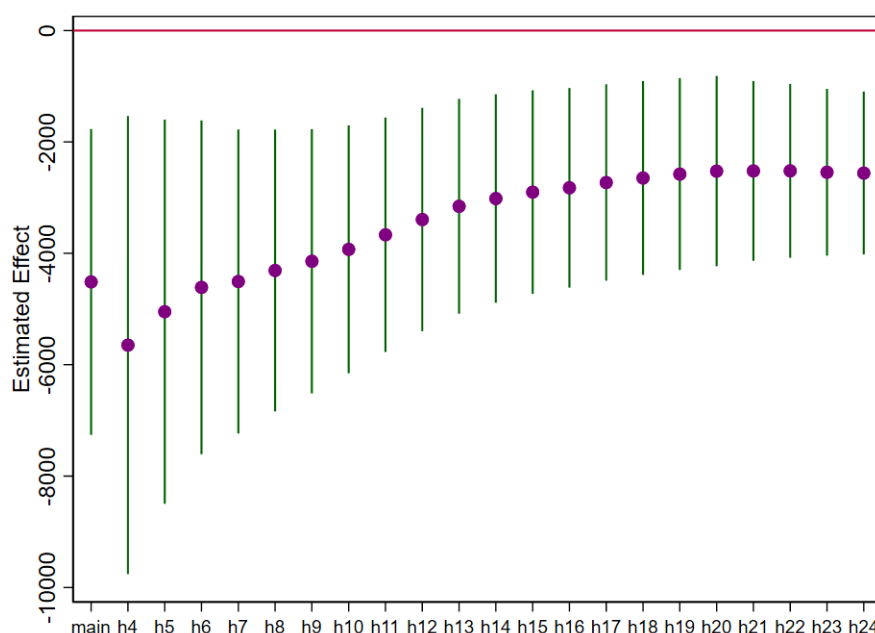
Notes: This table employs a fuzzy Regression Discontinuity Design to gauge the impact of third-grade retention on earnings. This estimation is performed by excluding the upper deciles (10% and 20%) of students from the counties exhibiting high out-of-state migration rates, as represented by the top 7% and 11%, respectively. The regression models control for students' characteristics, including gender, eligibility for free meals, and cohort fixed effects. The standard errors are clustered at the school level to account for similarities between students within the same school and are reported in parentheses.

Another limitation of the TWC data is the exclusion of government earnings. However, this is not a significant concern as government employment typically requires a high school diploma. The results in Table 6 demonstrate that retained students are less likely to graduate from high school relative to promoted students, reducing their likelihood of working in government. Therefore,

there is no evidence that grade retention disproportionately increases government employment for retained students.

This study employs a fuzzy regression discontinuity design to analyze the effect of grade retention on earnings, with the estimate limited to students near the promotion cutoff of the third-attempt reading test. To enhance the robustness of the results, I perform a sensitivity analysis varying the bandwidth around the cutoff and find that the negative impact of grade retention on earnings remains consistent, suggesting a robust effect. Figure 13, which presents the impact of grade retention on earnings 15 years after grade 3 using bandwidths from 4 to 24 points around the cutoff, supports the argument above.⁸

Figure 13: The effect of grade retention on earnings using different bandwidths



Notes: These figures show the fuzzy RDD estimates of grade retention's effect on earnings 15 years after grade 3 using bandwidths varying from 4 (h4) to 24 (h24). The bandwidth 24 includes all students who took the third attempt reading test. The main on the left of the x-axis represents the result in column (2) of Table 4 in which the bandwidth is chosen based on minimizing the mean square errors.

⁸Bandwidth 24 encompasses all students who took the third attempt reading test.

7 Conclusion

The impact of grade retention on low-achieving students has been a subject of debate. While some argue that it can improve cognitive performance, research suggests that it may cause behavioral issues, leading to tension around its overall effectiveness. This study uses a discontinuity generated by Texas' test-based retention policy to investigate the long-run effects of repeating the third grade on labor market outcomes. Results from a fuzzy RDD estimate show that holding low-achieving students back in the third-grade results in substantial reductions in earnings around age 23.

An examination of the mechanisms behind these findings reveals that while third-grade retention decreases the risk of repeating a grade, it concurrently increases behavioral problems, such as absenteeism and violent behaviors. The adverse effects on behavioral issues outweigh the initial gain on cognitive measures, leading to lower high school graduation rates. The negative effects of grade retention on behavioral issues and high school graduation jointly explain approximately 21% of the reductions in earnings 15 years after grade 3 caused by third-grade retention.

The results of this study challenge the notion that grade retention helps low-achieving students academically and instead suggest that it holds them back as they enter the labor market. The disparities in retention rates between Black, Hispanic, and White students, combined with the significant reductions in earnings, highlight the exacerbation of racial and ethnic inequalities in earnings due to grade retention.

Notably, this study focuses on the impact of holding back students versus promoting them rather than the broader effects of the retention policy. Texas's test-based retention policy also includes instructional support for students who failed the test, which may positively impact their educational attainment.

References

- Barr, A. and Gibbs, C. R. (2022). Breaking the cycle? intergenerational effects of an antipoverty program in early childhood. *Journal of Political Economy*, 130(12):3253–3285.
- Barreca, A. I., Lindo, J. M., and Waddell, G. R. (2016). Heaping-induced bias in regression-discontinuity designs. *Economic inquiry*, 54(1):268–293.
- Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, 17(2):372–404.
- Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., and Yagan, D. (2011). How does your kindergarten classroom affect your earnings? evidence from project star. *The Quarterly journal of economics*, 126(4):1593–1660.
- Craigie, T.-A., Grawert, A. C., and Kimble, C. (2020). *Conviction, imprisonment, and lost earnings: How involvement with the criminal justice system deepens inequality*. Brennan Center for Justice at New York University School of Law.
- Currie, J. and Almond, D. (2011). Chapter 15 - Human capital development before age five. In *Handbook of Labor Economics*, volume 4, pages 1315–1486. Elsevier.
- Deming, D. (2009). Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start. *American Economic Journal: Applied Economics*, 1(3):111–134.
- Elder, T. E. and Lubotsky, D. H. (2009). Kindergarten entrance age and children’s achievement impacts of state policies, family background, and peers. *Journal of human Resources*, 44(3):641–683.
- Eren, O., Depew, B., and Barnes, S. (2017). Test-based promotion policies, dropping out, and juvenile crime. *Journal of Public Economics*, 153:9–31.
- Eren, O., Lovenheim, M. F., and Mocan, H. N. (2022). The effect of grade retention on adult crime: Evidence from a test-based promotion policy. *Journal of Labor Economics*, 40(2):361–395.
- Figlio, D. and 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.
- García, J. L., Heckman, J. J., Leaf, D. E., and Prados, M. J. (2020). Quantifying the life-cycle

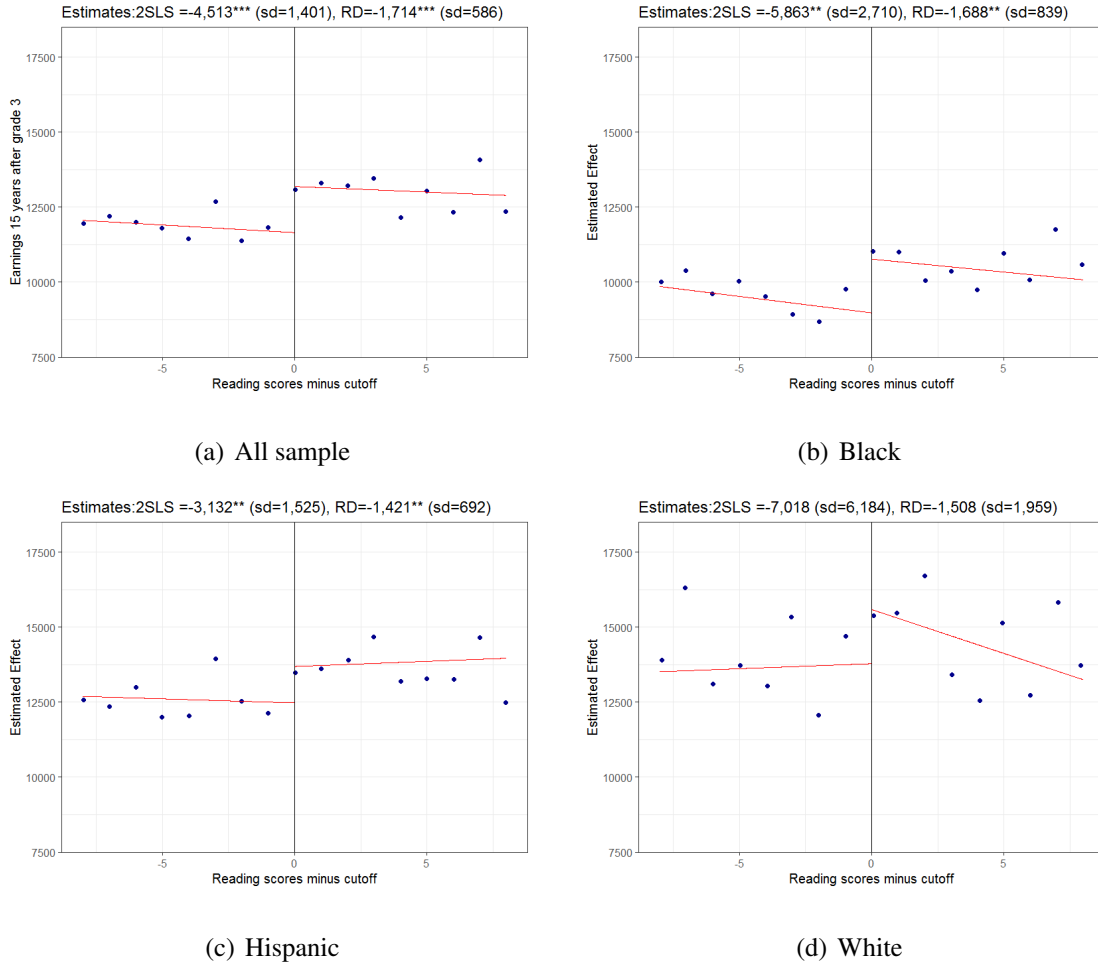
- benefits of an influential early-childhood program. *Journal of Political Economy*, 128(7):2502–2541.
- Greene, J. P. and Winters, M. A. (2007). Revisiting grade retention: An evaluation of florida’s test-based promotion policy. *Education Finance and Policy*, 2(4):319–340.
- Heckman, J. J. and Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the ged testing program. *American Economic Review*, 91(2):145–149.
- Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3):411–482. Publisher: The University of Chicago Press.
- Jacob, B. A. and Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *Review of economics and statistics*, 86(1):226–244.
- Jacob, B. A. and Lefgren, L. (2009). The effect of grade retention on high school completion. *American Economic Journal: Applied Economics*, 1(3):33–58.
- Kolesár, M. and Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, 108(8):2277–2304.
- Lindqvist, E. and Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the swedish enlistment. *American Economic Journal: Applied Economics*, 3(1):101–28.
- Looney, A. and Turner, N. (2018). Work and opportunity before and after incarceration. *Washington, DC: Brookings Institution*. Accessed October, 5:2018.
- Manacorda, M. (2012). The cost of grade retention. *Review of Economics and Statistics*, 94(2):596–606.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714.
- Özek, U. (2015). Hold back to move forward? early grade retention and student misbehavior. *Education Finance and Policy*, 10(3):350–377.
- Schwerdt, G., West, M. R., and Winters, M. A. (2017). The effects of test-based retention on

student outcomes over time: Regression discontinuity evidence from florida. *Journal of Public Economics*, 152:154–169.

A Appendix

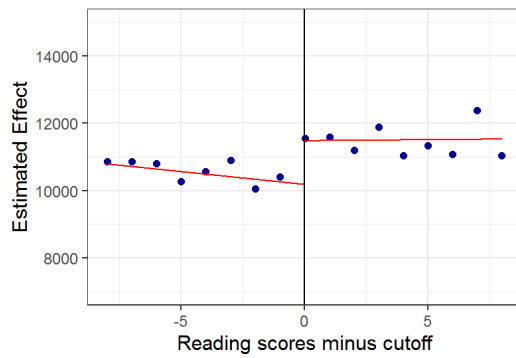
Online Appendix for "Early Grade Retention Harms Adult Earnings" by Jiee Zhong

Figure A.1: Failing the third-attempt reading test reduces earnings 15 years later

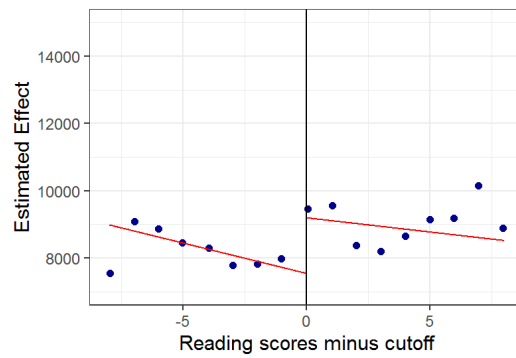


Notes: These figures show the reduced-form estimate (α_1 from equation (1)) of failing the third-attempt reading test on average earnings 15 years after grade 3 (around age 23). The blue dots are cell means, and the lines are fitted values from a first-order polynomial regression. The top of each sub-figure indicates the estimates from two-stage least squares (β_1 from equation (3)) and reduced-form (α_1 from equation (1)).

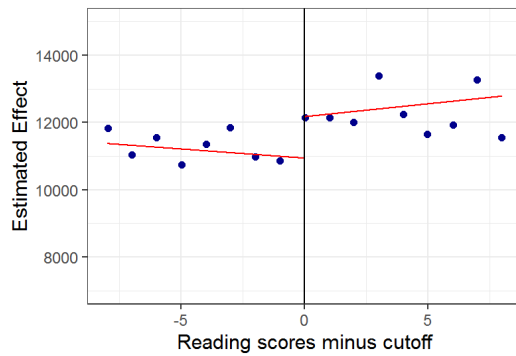
Figure A.2: Failing the third-attempt reading test reduces earnings 14 years later



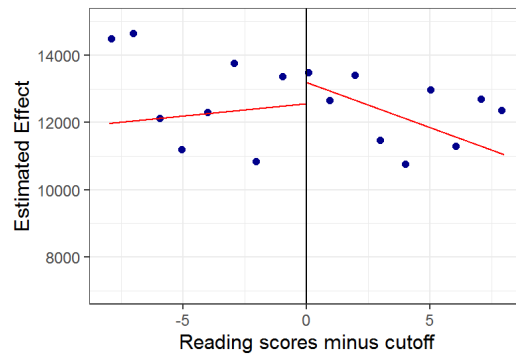
(a) All sample



(b) Black



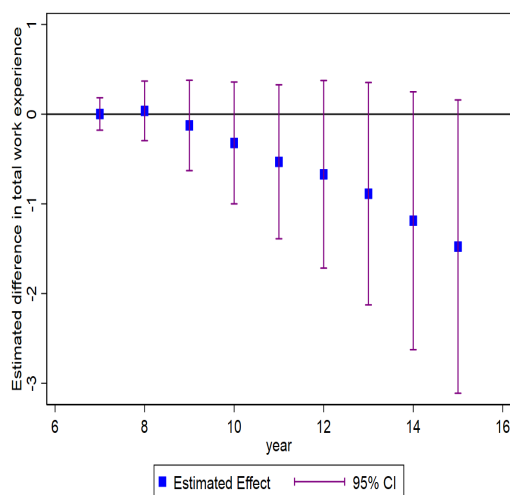
(c) Hispanic



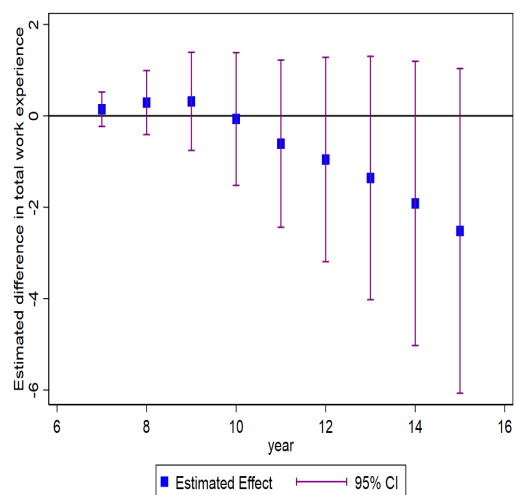
(d) White

Notes: These figures show the reduced-form estimate (α_1 from equation (1)) of failing the third-attempt reading test on average earnings 14 years after grade 3 (around age 22). The blue dots are cell means, and the lines are fitted values from a first-order polynomial regression.

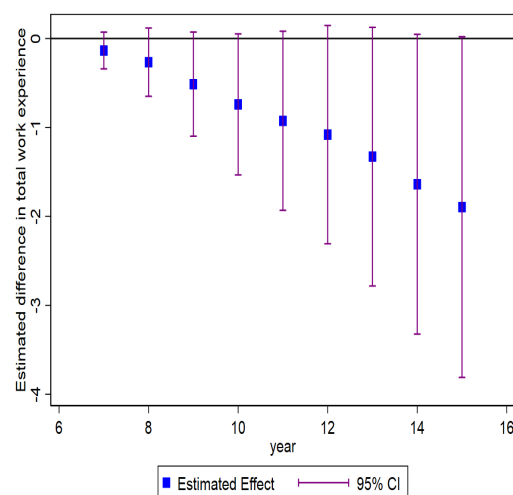
Figure A.3: Effect of retention on cumulative work experience by race and ethnicity



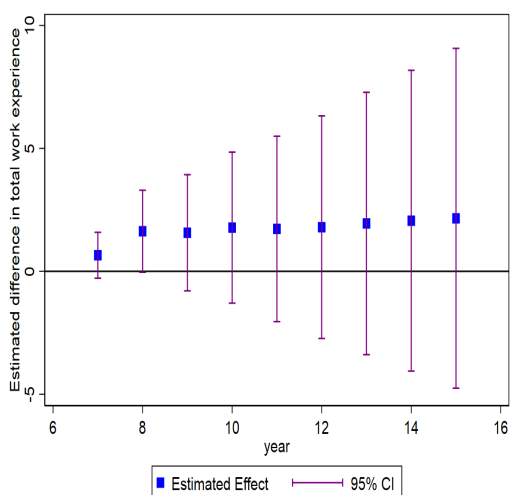
(a) Full sample



(b) Black



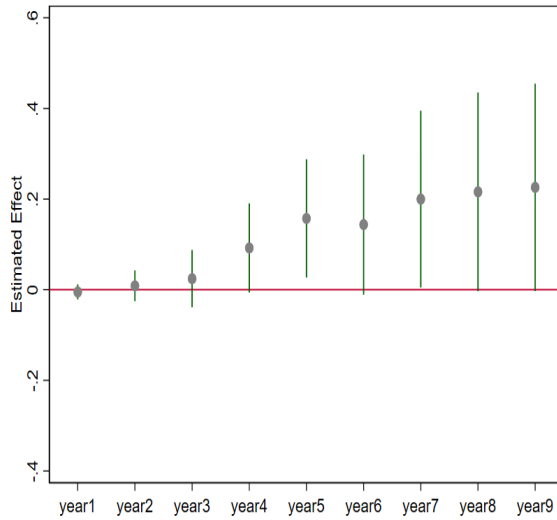
(c) Hispanic



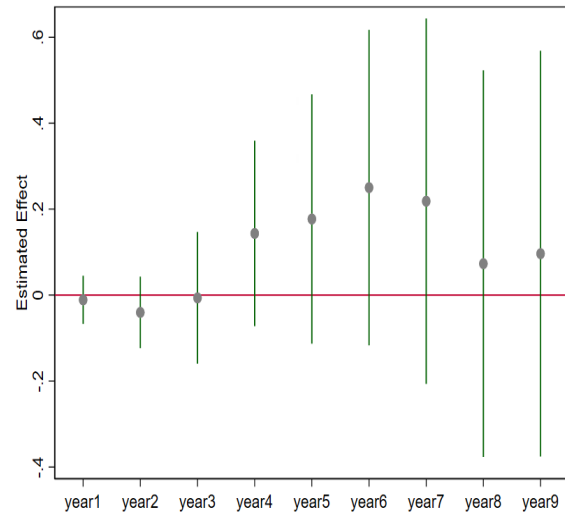
(d) White

Notes: This figure summarizes the results of nine fuzzy RDD regressions. The dependent variable is the total number of quarters a student has non-missing wages from 6 to X years after grade 3, where X is plotted on the horizontal axis. The figure plots the estimated coefficients and 95% confidence intervals.

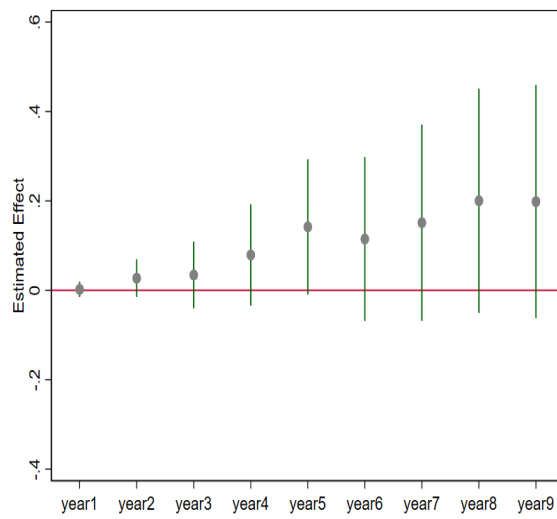
Figure A.4: Effect of grade retention on crime



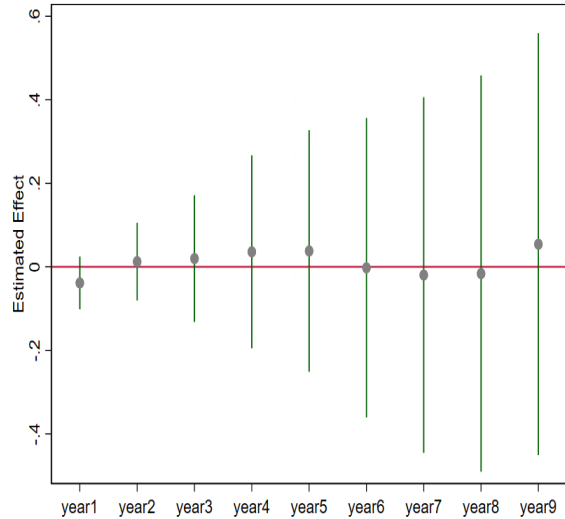
(a) Full sample



(b) Black



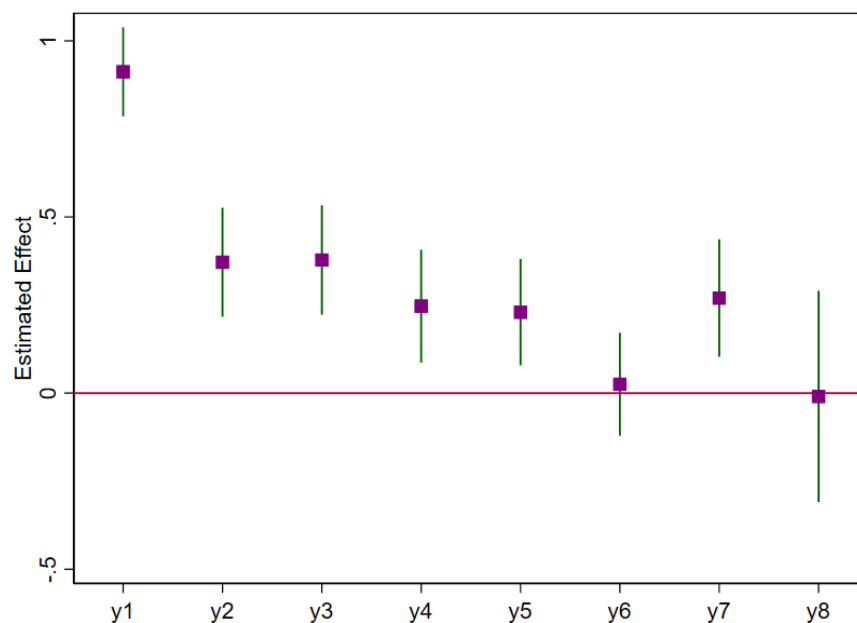
(c) Hispanic



(d) White

Notes: These figures show the fuzzy RDD estimate of the effect of grade retention on cumulative criminal activity in the years after grade 3. Criminal activity is a disciplinary problem categorized as a crime in the disciplinary data.

Figure A.5: The effect of grade retention on reading scores: same-age comparison



Notes: This figure displays the fuzzy RDD estimate of holding students back in third grade on the standardized reading scores in the subsequent years after the first time these students enter grade 3 (same-age comparison).

Table A1: Effect of retention on earnings outcomes by race and ethnicity

Outcomes	Black		Hispanic		White	
	(1)	(2)	(3)	(4)	(5)	(6)
Earnings 15 years later	-5846** (2607)	-5863** (2710)	-3189** (1610)	-3132** (1525)	-6322 (6445)	-7018 (6184)
Control group mean	[10,583]	[10,583]	[13,665]	[13,665]	[14, 642]	[14, 642]
Earnings 14 years later	-5,630** (2,243)	-5,728** (2,320)	-3,228** (1,443)	-3,230** (1,406)	-2,050 (5,201)	-2,828 (5,116)
Control group mean	[9,030]	[9,030]	[12,274]	[12,274]	[12,396]	[12,396]
Ave.earnings in 2018-2020	-5139** (2283)	-5213** (2374)	-2426* (1410)	-2424* (1354)	-2856 (5379)	-2653 (5249)
Control group mean	[10,307]	[10,307]	[13,374]	[13,374]	[13,769]	[13,769]
Observations	5052	5052	13315	13315	2727	2727
Demographic controls	No	Yes	No	Yes	No	Yes

Notes: This table displays the fuzzy RDD estimate of the effect of grade retention on earnings outcomes by race and ethnicity. The sample includes three cohorts of first-time third-grade students who took the third-attempt reading test from 2002–03 to 2004–05 school years. Odd-numbered columns report baseline estimates without controls. Even-numbered columns control for gender, eligibility for free meals, and cohort fixed effects. Values with square brackets represent the mean of outcomes among students whose scores are 0–8 points above the cutoff (or the control group). The bandwidth for each regression is chosen by optimizing the mean squared error, as introduced in Calonico et al. (2017). The standard errors are clustered at the school level to account for similarities between students within the same school and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A2: Effect of retention on cumulative absences after grade 3

	(1) 1 year	(2) 2 years	(3) 3 years	(4) 4 years	(5) 5 years	(6) 6 years	(7) 7 years	(8) 8 years	(9) 9 years
All	2.6** (1.0)	3.9*** (1.5)	4.4** (2.0)	5.2** (2.6)	6.8** (3.3)	8.4** (4.2)	10.1** (5.1)	9.0 (5.7)	8.0 (6.2)
Control group mean	[11.2]	[16.8]	[23.3]	[31.0]	[40.1]	[50.3]	[61.7]	[73.0]	[83.2]
Observations	17404	17404	17404	17404	17404	17404	17404	17404	17404
Black	5.3** (2.4)	8.1** (3.4)	11.4** (4.6)	14.3** (6.1)	19.4** (7.9)	23.2** (9.8)	28.8** (11.9)	28.0** (13.3)	26.0* (14.3)
Control group mean	[12.0]	[17.9]	[24.7]	[32.7]	[41.7]	[52.0]	[63.6]	[75.1]	[85.1]
Observations	5052	5052	5052	5052	5052	5052	5052	5052	5052
Hispanic	2.0** (1.0)	3.0** (1.5)	2.5 (2.0)	2.3 (2.7)	3.0 (3.4)	2.8 (4.4)	2.7 (5.3)	0.6 (6.1)	-1.1 (6.6)
Control group mean	[10.5]	[15.9]	[22.2]	[29.9]	[39.3]	[49.9]	[61.8]	[73.7]	[84.5]
Observations	12022	12022	12022	12022	12022	12022	12022	12022	12022
White	-4.1 (3.4)	-4.6 (5.1)	-6.1 (6.9)	-7.2 (9.2)	-8.6 (11.9)	-2.6 (14.5)	0.7 (16.9)	4.0 (19.0)	5.5 (20.3)
Control group mean	[13.2]	[19.9]	[26.9]	[34.4]	[43.0]	[51.9]	[61.3]	[70.4]	[78.4]
Observations	2727	2727	2961	2727	2727	2727	2727	2727	2727

Notes: This table shows the fuzzy RDD estimate of grade retention's effect on cumulative absences from school years after grade 3. The sample includes three cohorts of first-time third-grade students who took the third-attempt reading test from 2002–03 to 2004–05 school years. Each regression controls for gender, free lunch eligibility, and cohort fixed effect. The bandwidth for each regression is chosen by optimizing the mean squared error, as introduced in Calonico et al. (2017). The standard errors are clustered at the school level to account for similarities between students within the same school and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A3: Effect of retention on cumulative violent behavior after grade 3

	(1) 1 year	(2) 2 years	(3) 3 years	(4) 4 years	(5) 5 years	(6) 6 years	(7) 7 years	(8) 8 years	(9) 9 years
All	0.0 (0.0)	0.1 (0.0)	0.1* (0.1)	0.3*** (0.1)	0.4*** (0.1)	0.5*** (0.1)	0.5*** (0.2)	0.5*** (0.2)	0.6*** (0.2)
Control group mean	[0.1]	[0.1]	[0.2]	[0.3]	[0.5]	[0.6]	[0.7]	[0.8]	[0.9]
Observations	22070	19909	19909	17404	19909	19909	19909	17404	17404
Black	0.1 (0.1)	0.2 (0.2)	0.4* (0.2)	0.7** (0.3)	0.9*** (0.4)	1.2*** (0.4)	1.1** (0.5)	1.1** (0.5)	1.1** (0.5)
Control group mean	[0.1]	[0.2]	[0.4]	[0.6]	[0.8]	[0.9]	[1.1]	[1.2]	[1.3]
Observations	5052	5052	5052	5052	5052	5052	5052	5052	5052
Hispanic	0.0 (0.0)	0.0 (0.0)	0.0 (0.1)	0.2** (0.1)	0.3** (0.1)	0.3* (0.1)	0.3* (0.2)	0.4* (0.2)	0.4* (0.2)
Control group mean	[0.0]	[0.1]	[0.1]	[0.3]	[0.4]	[0.5]	[0.7]	[0.7]	[0.8]
Observations	13315	12022	12022	10512	12022	12022	12022	12022	12022
White	0.1 (0.1)	0.1 (0.2)	0.0 (0.2)	-0.0 (0.3)	-0.0 (0.3)	0.0 (0.3)	0.0 (0.4)	0.0 (0.4)	0.1 (0.4)
Control group mean	[0.0]	[0.1]	[0.1]	[0.2]	[0.2]	[0.3]	[0.4]	[0.4]	[0.5]
Observations	2727	2466	2466	2466	2466	2466	2727	2727	2727

Notes: This table shows the fuzzy RDD estimate of grade retention's effect on cumulative violent behaviors years after grade 3. Violent behavior is defined as a disciplinary problem categorized as violence in the disciplinary data. The sample includes three cohorts of first-time third-grade students who took the third-attempt reading test from 2002–03 to 2004–05 school years. Each regression controls for gender, eligibility for free school lunch, and cohort fixed effect. The bandwidth for each regression is chosen by optimizing the mean squared error, as introduced in Calonico et al. (2017). The standard errors are clustered at the school level to account for similarities between students within the same school and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A4: Effect of retention on cumulative crime after grade 3

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	1 year	2 years	3 years	4 years	5 years	6 years	7 years	8 years	9 years
All	-0.0	0.0	0.0	0.1*	0.2**	0.1*	0.2**	0.2*	0.2*
	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
Control group mean	[0]	[0]	[0]	[0.1]	[0.1]	[0.2]	[0.3]	[0.3]	[0.4]
Observations	19909	19909	17404	17404	17404	19909	17404	17404	17404
Black	-0.0	-0.0	-0.0	0.1	0.2	0.3	0.2	0.1	0.1
	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)	(0.2)	(0.2)	(0.2)	(0.2)
Control group mean	[0]	[0]	[0]	[0.1]	[0.1]	[0.2]	[0.3]	[0.3]	[0.4]
Observations	5052	5052	5052	5052	5052	5052	5052	5052	5052
Hispanic	0.0	0.0	0.0	0.1	0.1*	0.1	0.2	0.2	0.2
	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
Control group mean	[0]	[0]	[0]	[0.1]	[0.2]	[0.2]	[0.3]	[0.4]	[0.4]
Observations	12022	12022	12022	12022	12022	12022	12022	12022	12022
White	-0.0	0.0	0.0	0.0	0.0	-0.0	-0.0	-0.0	0.1
	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)	(0.2)	(0.2)	(0.2)	(0.3)
Control group mean	[0]	[0]	[0]	[0]	[0.1]	[0.1]	[0.2]	[0.2]	[0.2]
Observations	2961	2961	2961	2961	3173	2961	2961	2961	2727

Notes: This table shows the fuzzy RDD estimate of grade retention's effect on cumulative criminal activity years after grade 3. Crime is defined as a disciplinary problem categorized as a crime in the disciplinary data. The sample includes three cohorts of first-time third-grade students who took the third-attempt reading test from 2002–03 to 2004–05 school years. Each regression controls for gender, eligibility for free school lunch, and cohort fixed effect. The bandwidth for each regression is chosen by optimizing the mean squared error, as introduced in Calonico et al. (2017). The standard errors are clustered at the school level to account for similarities between students within the same school and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A5: Effect of retention on reading scores after grade 3: same grade comparison

	(1) Grade 4	(2) Grade 5	(3) Grade 6	(4) Grade 7	(5) Grade 8	(6) Grade 9
Raw scores	0.571*** (0.074)	0.331*** (0.087)	0.387*** (0.088)	0.275*** (0.089)	0.241*** (0.087)	0.164 (0.137)
Observations	19437	17253	19999	18759	18794	13858
Scale scores	0.273*** (0.067)	0.115 (0.086)	0.243*** (0.075)	0.173** (0.077)	0.016 (0.093)	0.001 (0.114)
Observations	19078	17253	19999	18544	16432	13672

Notes: This table shows the fuzzy RDD estimate of grade retention's effect on reading scores after grade 3 when retained and promoted students reach the same grade. The sample includes three cohorts of first-time third-grade students who took the third-attempt reading test from 2002–03 to 2004–05 school years. Each regression controls for gender, eligibility for free school lunch, and cohort fixed effect. The bandwidth for each regression is chosen by optimizing the mean squared error, as introduced in Calonico et al. (2017). The standard errors are clustered at the school level to account for similarities between students within the same school and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A6: Effect of retention on reading scores after grade 3: same age comparison

	(1) 1 year	(2) 2 years	(3) 3 years	(4) 4 years	(5) 5 years	(6) 6 years	(7) 7 years	(8) 8 years
Raw scores	0.912*** (0.065)	0.371*** (0.079)	0.378*** (0.079)	0.247*** (0.082)	0.230*** (0.077)	0.025 (0.075)	0.270*** (0.085)	-0.009 (0.153)
Observations	18435	18484	19261	17638	18105	17844	4786	4149
Scale scores	0.521*** (0.061)	0.132* (0.077)	0.307*** (0.069)	0.181** (0.072)	-0.070 (0.087)	-0.046 (0.069)	0.181** (0.083)	-0.006 (0.130)
Observations	18184	18412	17384	17552	15780	17765	9665	4679

Notes: This table shows the fuzzy RDD estimate of grade retention's effect on reading scores in the years after grade 3. The sample includes three cohorts of first-time third-grade students who took the third-attempt reading test from 2002–03 to 2004–05 school years. Each regression controls for gender, free lunch eligibility, and cohort fixed effect. The bandwidth for each regression is chosen by optimizing the mean squared error, as introduced in Calonico et al. (2017). The standard errors are clustered at the school level to account for similarities between students within the same school and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A7: The back-of-envelope calculation

Test scores	High school graduation	Violent behaviors	Retention after grade 3
Panel A: effects of retention on intermediate outcomes			
0.571*** (0.074)	-0.081* (0.043)	0.6*** (0.2)	0.552*** (0.059)
Panel B: conditional correlation between intermediate outcomes and earnings			
748 *** (195)	7266*** (307)	-599*** (64)	-288 (211)
Panel C: the back-of-envelope calculation			
427	-588	-359	-

Notes: This table shows the process of the back-of-envelope calculation. Panel A displays the causal effects of grade retention on outcomes shown in each column. Panel B shows the conditional correlation between these intermediate outcomes and earnings 15 years after grade 3 for the control group (students who scored 0 to 8 points above the cutoff). The standard errors are clustered at the school level to account for similarities between students within the same school and are reported in parentheses. Specifically, I use OLS to estimate the effects of these intermediate outcomes (include all these intermediate outcomes in one regression) on earnings 15 years after grade 3 by controlling for gender, eligibility for free meals, limited English proficiency, race, and cohort fixed effect. The values in Panel C equal the values in panel A multiplied by those in the same columns displayed in Panel B. I only show the back-of-envelope calculation when both the coefficients in Panel A and B are statistically significant at the 10% level. The effect of grade retention on days absent within nine years is 8 and is not statistically significant. Moreover, the conditional correlation between days absent within nine years and earnings is negligible. So, I do not include days absent in this table. Panel C implies the contribution of grade retention's effects on these intermediate outcomes to the reduction in earnings.