The Effect of Grade Retention on Adult Crime: Evidence from a Test-Based Promotion Policy*

Ozkan Eren
University of California - Riverside

Michael F. Lovenheim
Cornell University and NBER

H. Naci Mocan
Louisiana State University and NBER

March 2021

Abstract

This paper presents the first analysis in the literature of the effect of test-based grade retention on adult criminal convictions. We exploit math and English test cutoffs for promotion to ninth grade in Louisiana using administrative data on all public K-12 students combined with administrative data on all criminal convictions in the state. Our preferred model uses the promotion discontinuity as an instrument for grade retention, and we find that being retained in eighth grade has large long-run effects on the likelihood of being convicted of a crime by age 25 and on the number of criminal convictions by age 25. Effects are largest for violent crimes: the likelihood of being convicted increases by 1.05 percentage points, or 58.44%, when students are retained in eighth grade. Our data allow an examination of mechanisms, and we show that the effects are likely driven by changes to educational investments that result in lower non-cognitive skill acquisition. Using the method proposed by Angrist and Rokkanen (2015), we also estimate effects of grade retention away from the promotion cutoff and show that our results are generalizable to a larger group of low-performing students. Our estimates indicate that eighth grade test-based promotion cutoffs lead to nontrivial private and social costs in terms of higher levels of long-run criminal convictions that are important to consider in the development and use of these policies.

KEYWORDS: grade retention, crime, grade promotion policies, regression discontinuity

*We would like to thank Kevin Lang and two anonymous referees as well as seminar participants at University of Illinois-Urbana Champaign, National University-Singapore, Singapore Management University, University of California-Irvine, University of South Carolina, CESifo Area Conference on the Economics of Education, and the Association for Education Finance and Policy annual meeting for helpful comments and suggestions. Data used in this study were provided by the Louisiana Department of Education, Louisiana Office of Juvenile Justice, and Louisiana Department of Public Safety and Corrections, Adult Services. Access was provided by Louisiana State University. All results, conclusions and errors are our own.

1 Introduction

Education policies alter the skills and knowledge of young adults, which influence individual earnings and socially relevant outcomes such as civic engagement, health behaviors, and criminal activity later in life.[1] These social returns to education are critical to understand, because much of the argument for the large role played by the government in funding and delivering education services is predicated on the existence of positive externalities. While a sizable literature examines the private returns to education investments and policies in the form of long-run labor market outcomes, there is a paucity of evidence on the magnitude of social returns (Grossman 2006; Oreopoulos and Salvanes 2011).

This paper helps fill the gap in our understanding of the social returns to education policy by estimating the effect of a test-based grade retention policy for promotion to 9^{th} grade on adult criminal convictions in Louisiana. Crime is one of the most important social outcomes to examine because it generates enormous costs to society.[2] Investigation of the determinants of criminal activity also is important because delinquency and incarceration have long-run negative effects on outcomes, such as employment opportunities, family formation, and health (Johnson and Raphael 2009; Charles and Luoh 2010; Agan and Starr 2017).

Education can lead to reductions in adult criminal behavior by increasing the opportunity cost of engaging in crime (Lochner 2004), by altering time discounting (Becker and Mulligan 1997), and by building non-cognitive skills that contribute to better self-control and decision-making as adults (Fudenberg and Levine 2006). Despite the strong theoretical arguments linking education to crime, empirical work is sparse because of the difficulty in overcoming selection into criminal activity. Existing work with credible identification strategies has investigated the impact of the number of years of education on criminal propensity (e.g., Lochner and Moretti 2004; Machin et al. 2011). We instead exploit a grade retention policy in a regression discontinuity framework to examine the extent to which having been held back in 8^{th} grade impacts criminal delinquency up to age 25.

Louisiana adopted the test-based accountability program we analyze in 1998, which requires

[1]See Card (2001) and Oreopoulos (2007) for estimates of the effect of education on wages. For estimates of the impact of education on civic participation, production of health, and crime, see Dee (2004), Lochner and Moretti (2004), Lleras-Muney (2005) and Chou et al. (2010).

[2]Ferraz and Soares (2018) calculate that the expenditures on the criminal justice system and the cost of victimization add up to $450 billion in the U.S., and Anderson (1999) estimates a total cost of criminal activity of $1.7 trillion in 2018 dollars.

8^{th} grade students to achieve proficiency on the state math and English exams in order to be promoted to 9^{th} grade. Students who score below the cutoff on either test are offered remedial summer classes at no cost, and they retake the failed exam(s) in the summer regardless of whether they enroll in summer school. If a student fails to reach the passing threshold on one of the summer exams, she is retained in 8^{th} grade. We employ administrative records on all K-12 students from Louisiana merged with administrative criminal justice records. Our analysis is restricted to those who fail one of the initial exams.

Louisiana was an early of adopter of test-based accountability, which has grown markedly in the past several decades with the rise of accountability policies more broadly. Typically, accountability policies are aimed at teachers (e.g., teacher merit pay) or schools (e.g., No Child Left Behind). Test-based promotion policies are aimed directly at students, and currently 16 states as well as many large school districts incorporate end-of-year exams into the decision to promote students to the next grade (Zinth 2005; Marsh et al. 2009; Workman 2014).[3] By ending "social promotion," the goal of these policies is to provide incentives for students to increase learning. A growing literature examines the effect of these policies, and of grade retention more generally, on measures of academic performance and juvenile crime (e.g., Jacob and Lefgren 2004 and 2009; Manacorda 2012; Eren et al. 2017; Schwerdt et al 2017; Mariano et al. 2018). The effect of these policies on longer-run adult outcomes has remained unexamined prior to this analysis.

Being retained in 8^{th} grade increases the likelihood of being convicted of any adult crime by a non-statistically-significant 1.25 percentage points (10.85%). The effect is driven by violent crime: 8^{th} grade retention increases the likelihood of violent crime conviction by 1.05 percentage points (58.44%). This estimate is significant at the 13% level when we account for multiple hypothesis testing. We find no effect on the likelihood of committing a property or drug related crime, although the point estimates for drug crimes are positive and sizable in magnitude. People can be convicted of multiple crimes at once, and being convicted of more crimes at the same time is evidence of the severity of the offense. Retention increases the number of offenses

[3] The promotion and retention policies vary along several key dimensions including, but not limited to, grade level, subjects tested and waiver criteria. For example, Arizona and Florida target their promotion policies at grade 3, while Georgia and Texas focus on grades 5 and 8. The findings on the efficacy of test-based promotion policies on long-run outcomes is far from conclusive. This mixed evidence is not solely a function of variation in design features of promotion policies, because policies implemented at similar grades on identical subjects yield different long-run impacts (e.g., Jacob and Lefgren (2009) and Eren et al. (2017)).

for which individuals are convicted at the time they are first convicted by 0.025 (18.38%), which is composed primarily of a 0.017 (85.00%) violent crime increase. This estimate is significant at the 5% level after adjusting for multiple hypothesis testing.

Examining mechanisms, we find that grade retention reduces education quality and deteriorates non-cognitive skills, which manifests itself in terms of lower school attendance, more behavioral problems, and less educational attainment. The result is a higher probability of being convicted of a crime as an adult, and this effect is not simply a reflection of a higher likelihood of being in the criminal justice system as a juvenile.

Using the technique pioneered by Angrist and Rokkanen (2015), we estimate effects of retention on inframarginal students who pass the exam and who fail the exam. We find positive and statistically significant effects away from the cutoff for violent, property and drug-related crime. These results suggest that our regression discontinuity estimates understate the effects of grade retention on crime for a broader population of low-scoring students and indicate that the passing threshold cannot be simply moved up or down to eliminate the crime effect.

This paper adds to a body of research that examines the effect of grade retention on student outcomes. Several studies have investigated test-based student accountability policies in a regression discontinuity framework as we do here (Greene and Winters 2007; Jacob and Lefgren 2009; Schwert et al. 2017). They find that retaining students in higher grades has negative effects on educational attainment, but the same negative effects are not evident for lower-grade retention.[4] We are the first to examine longer-run outcomes more broadly and adult convictions in particular.

Eren et al. (2017) make use of the same Louisiana policy to estimate the effect of summer school on high school dropout. Their supplementary analysis shows that retention in 8^{th} grade has no impact on juvenile crime convictions.[5] This finding, which is also confirmed by our paper, indicates that the effect of grade retention on adult crime is not the result of the continuation of crime commission that started as a juvenile. This underscores the importance of directly examining longer-run outcomes. Recent evidence suggests that short-run effects of educational interventions differ substantively from longer-run effects, though we are the first to show that

[4]Ou and Reynolds (2010) and Fruehwirth et al. (2016) provide further evidence that the effect of grade retention varies as a function of when students are retained.

[5]Other work has found evidence of a link between high school attainment and juvenile crime (e.g., Anderson 2014).

this is the case with crime outcomes.[6]

Our analysis also contributes to a small literature on the effect of education on crime.[7] These papers use changes in compulsory schooling laws (Hjalmarsson et al. 2015; Oreopoulos and Salvanes 2011; Machin et al. 2011; Lochner and Moretti 2004) or winning a school choice lottery (Deming 2011; Cullen et al. 2006) for identification and generally find that more schooling leads to less crime. Critically, the marginal student who fails/passes a promotion exam is likely to differ from the marginal student in other studies. Additionally, as we show in this paper, grade retention can impact crime through mechanisms other than altering the number of years of schooling.

Because violent crimes are associated with large externalities, the sizable increase in violent crime convictions due to grade retention we document may reverse any social benefits to grade retention policies. We calculate that increase in violent crimes, driven by grade retention, implies a social cost between \$2.6 to \$18.4 million over the three cohorts of students we examine. Test-based promotion specifically, and accountability systems generally, can incentivize harder work and lead to more learning for all students. Our RD estimates identify the net effect of failing the promotion cutoff, not the effect of the accountability program itself. The benefits from the accountability program may be larger than the costs associated with grade retention that we estimate here. Whether test-based promotion policies are cost-beneficial is beyond the scope of our analysis, but the aggregate benefits would have to be large to overcome the costs from more crime.

Another policy implication relates to where the cutoff for promotion is located. At the current location of the pass-fail cutoff, there is a long-run effect on criminal behavior, and our estimates based on the Angrist and Rokkanen (2015) extrapolation method suggest we would obtain similar effects with a different cutoff. Our results point to specific types of students who are being harmed by this 8^{th} grade retention policy. These students could be targeted with additional resources, which could perhaps mitigate the effects we identify.

[6]For example, researchers have shown that short-run effects do not predict long run effects when examining head start (Ludwig and Miller 2007), class size (Chetty et al. 2011), school choice (Deming et al. 2013; Beuermann and Jackson 2019), teachers unions (Lovenheim and Willén 2019), and Medicaid (Cohodes et al. 2016).

[7]See Lochner (2011) for a detailed review of the literature.

2 Policy Background

The Louisiana School and District Accountability system was adopted in June 1998. School accountability is a core component of the policy, with the state setting long-run goals and requiring schools to demonstrate progress towards them. The policy also includes a student-focused accountability provision to limit "social promotion," whereby students were passed to the next grade regardless of their school performance. Under the new test-based promotion policy, students in eighth grade are required to score at predefined levels on the Louisiana Educational Assessment Program (LEAP) tests for both English Language Arts (ELA) and math to advance to the next grade.[8]

LEAP tests are criterion-referenced tests designed to directly align with state content standards. LEAP scores can be expressed as either a continuous scale score ranging from 100 to 500 points or as a discrete achievement level ranging from unsatisfactory to advanced. Students must score at least Approaching Basic in both subjects to advance to the next grade. This is equivalent to 269 and 296 scale points in ELA and math tests, respectively.[9] In addition to LEAP tests, students in earlier middle school grades were given Iowa Tests of Basic Skills (ITBS). The ITBS is a low-stakes norm-referenced test for which scores are compared to a national norm group. All tests are first administered in mid-March.

Students who fail to achieve the promotional standards in March are required to retake each failed subject exam in July. The school districts must offer, at no cost, a minimum 50 hours per subject of summer remediation in ELA and math to students who fail to meet the passing standards in March. Participation in summer school is optional. School districts are given the flexibility to determine curricula used in summer remediation classes, but the programs also are monitored by the state. Evidence from annual reports and monitoring visits suggests the summer school programs are staffed by qualified teachers who use high-quality pedagogical methods. However, there are large classrooms in some programs (around 20 students per classroom), which has raised concerns about their effectiveness (Pastorek 2010).

Those who pass the July exams move on to the next grade. Students who fail the July test are

[8] Similar test-based promotion policies are implemented in fourth grade as well.

[9] Raw scores are transformed to scaled scores in a three stage process. They are first mapped onto the Item Response Theory scale. They are then converted to a reporting scale and finally, they are equated to reflect the differences in item difficulty.

required to repeat the grade unless they receive a waiver. A school system may grant waivers for students following extenuating circumstances, such as physical illness or court-ordered custody issues. Similarly, students may receive waivers because of local education agency errors or other unique situations not covered under extenuating circumstances. The state also may grant a waiver to LEP students.[10] Conversely, students can fail a grade if state-mandated attendance requirements are not met, irrespective of LEAP scores. These institutional features lead us to estimate fuzzy regression discontinuity models, since failing or passing one of the July LEAP exams does not necessarily determine grade retention.

3 Data

We use data from a merge of administrative education and criminal records that come from two different sources. The first data source is administrative records from the Louisiana Department of Education (LDOE). These data include student demographic information (gender, race and free/reduced lunch status) and scores from the LEAP and ITBS tests. Unique state identification numbers allow us to track all the students through their tenure in the public school system and allow us to identify schools in which each student was enrolled. The data also include whether a student has dropped out, the number of days absent, and the number of disciplinary incidents. Because we observe all public school enrollments in the state, attrition only arises if students leave the state, attend a private school, or are home schooled. It is plausible that students who fail the July exam respond by leaving the state or enrolling in a private school, but as we show below, failure is uncorrelated with the likelihood of attriting from the public education sample. Only leaving the state causes attrition in our crime estimates.[11]

The second data source is the Louisiana Department of Public Safety and Corrections, Adult Services. These data span 1996 to 2012 and are merged with the LDOE data using common state identification numbers from both datasets. The crime data include demographic information, the exact type of crime committed, and sentence type (incarceration or probation)

[10] There were not well defined waiver policies over our sample period. Thus, while we can observe non-compliance in the data, we do not know why a student is non-compliant.

[11] While we cannot directly test for such attrition in the crime data, that we do not observe any differential attrition out of the education sample due to failing the promotion exam suggests that it is not a concern in the crime data either. Tabulations from the American Community Survey show that about 5% of those born between 1980 and 1985 in Louisiana moved out of the state between ages 18-25. This rate is just above 2% for those with at most a high school degree.

for each conviction. Importantly, the dataset is comprised only of those who are convicted of a crime. Thus, these data include the most serious offenses that led to criminal convictions. While 91% of property crime and 87% of drug-related crime convictions are given probation, only 55 percent of violent crime convictions receive probation. The high rate of probation for drug-related offenses suggest that most of these offenses are related to drug use rather than drug dealing.

Our outcome of interest is adult conviction by age 25. We use this age cutoff because of the timing of our education data and to be consistent with the literature (Aizer and Doyle 2015): most crime is committed by those under 26 (Hirschi and Gottfredson 1983). In order to measure criminal activity without any censoring, we limit our focus to eighth graders from the 1998-1999 to the 2000-2001 academic years.[12] Our analysis sample is composed solely of students who fail to meet the March promotional cutoff and who took the July exams. That both retained and promoted students had equal access to summer school means our estimates are net of any summer school effect (Matsudaira 2008). Our final analysis sample consists of 22,929 unique student observations, representing about 15% of the universe of eighth graders in the cohorts we study.

Table 1 presents descriptive statistics for the July testing sample. We show tabulations for the full sample as well as by retention status. Because we focus on low-performing students, the sample disproportionately includes disadvantaged students with low test scores. For example, 45% of the sample is on free/reduced prices lunch and the mean student in the full sample scores 0.86 and 0.69 standard deviations below the state average in math and ELA, respectively. Seventy-six percent of the sample is Black. Retained students are more likely to be Black and have lower test scores than students who are not retained, as shown in column (5). The math exam is more of an obstacle for promotion than the ELA test: only 5% of the retained sample is retained solely because of ELA failure. About 5 percent of the retained students in eighth grade during our sample period passed the July exams but were retained for other reasons, such as not meeting the minimum school attendance requirements. This is why the 8^{th} Grade LEAP test outcomes only sum to 95% in column (5). Furthermore, about 35% of those promoted

[12] We do not have access to data prior to 1998 because the retention policy was implemented at the same time as the current testing regime.

have also failed one or both of the tests of the July exam but were given exemptions for various reasons and were allowed to advance (see Section 2).

The mean dropout rate for the full sample is quite high, at 45.8%.[13] This reflects the low academic achievement level of this group of students. Among those retained, the dropout rate is even higher, at 53.5%. The higher dropout rate among those who are retained is due to their low academic achievement level as well as any causal effect of retention. Our regression discontinuity approach is designed to overcome this selection problem. We also find that 4 percent of these students were convicted of a crime as a juvenile. Not surprisingly, juvenile conviction rates are also higher for those who were retained in eighth grade.

The final set of descriptive statistics in Table 1 relate to adult crime. We show that 11.5 percent of the students are convicted as adults by age 25. The age at first-time adult conviction is almost 21, which underscores the importance of examining outcomes into the mid-20s. About 2 percent are convicted of a violent crime, 4 percent of a property crime, 5 percent of drug-related crime, and 1 percent of another type of crime.[14] Because an individual may have been convicted of more than one crime, the sum of all crime categories is slightly greater than the overall conviction rate. This is also reflected in the higher means for the number of convictions shown in the table. Second-degree battery, simple robbery and armed robbery are the most common violent crime convictions, making up more than 41 percent of violent offenses. Simple burglary, theft and illegal possession of stolen things make up 57 percent of all property crime convictions, and almost all drug related convictions fall under the category of possession and distribution of drugs.

4 Empirical Methodology

4.1 Reduced Form and IV Models

In order to obtain credible estimates of the effect of grade retention on adult crime convictions, we rely on the exogenous variation generated by the accountability policy in Louisiana in a regression discontinuity framework. Assuming that there is randomness around the July test

[13]Drop out statistics do not include students who transferred to private/home school.

[14]This last category (other crimes) is a very heterogeneous group and includes offenses ranging from hit and run driving to aggravated incest, from operating a vehicle intoxicated to perjury.

score cutoff, all observed and unobserved characteristics of students should move smoothly through the cutoff. Evidence consistent with this assumption is presented in Section 4.2. We estimate reduced form models that show the effect of barely failing one of the July exams and two stage least squares models in which we use the discontinuity created by the accountability policy as an instrument for grade retention. Because we use the sample of students who failed the March exam, our estimates should not be confounded by the presence of summer school: those on both sides of the July cutoff had equal access to summer school, and it is unlikely that summer school takeup was differential across the cutoff.[15] The reduced-form model is as follows:

$$Convicted_i = \gamma_0 + \gamma_1 FP_i + f(Index_i; \delta) + X_i'\theta + \mu_i, \tag{1}$$

where $Convicted_i$ is either an indicator variable that takes the value of one if student i is convicted of a crime by the age of 25 or is the count of first-time convictions by age 25 (including zeros). The first-time conviction count is the number of crimes for which an individual was convicted when he was found guilty by a court the first time as an adult. This can be greater than one because people sometimes are convicted of multiple offenses at the same time. Our two outcome measures thus capture somewhat different margins of criminal engagement. The variable FP_i is an indicator that takes the value of one if student i scored below Approaching Basic in either of the July ELA or Math exams. $Index_i$ is the running variable and denotes the minimum of the subject-specific distances from the respective cutoffs:

$$Index_i = \min[S_i^{Math} - Cutoff^{Math}, S_i^{ELA} - Cutoff^{ELA}], \tag{2}$$

where S_i and $Cutoff$ are the LEAP score and the relevant cutoff in math and ELA. The functional form between $Index_i$ and adult criminal convictions is described by the polynomial function $f(\cdot)$. Our baseline models employ a linear spline in $Index_i$ (Gelman and Imbens 2017), and we show that the results are robust to using quadratic splines and to the use of local linear estimation. The vector X_i is a set of covariates that include indicators for gender, race, free/reduced lunch and immigrant status as well as year fixed effects. In some specifications,

[15] We do not have data on which students took up summer school to directly test this assertion. Information from the annual summer remediation reports (Pastorek 2010) and discussions with the LDOE administrators indicate that the participation rate was more than 90 percent among eligible students.

we also control for 8^{th} grade school fixed effects, for composite test scores in ELA and math from the 7^{th} grade ITBS exams,[16] and for the 8^{th} grade March exam. These test scores are standardized against the statewide mean and standard deviation, separately by test year and subject. The key identifying assumption underlying the identification of γ_1 is that unobserved characteristics move smoothly through the cutoff, conditional on the controls for $Index_i$. Under this assumption, for students near the cutoff, the estimate of γ_1 can be interpreted as the intent to treat (ITT) effect of grade retention.

We also estimate a fuzzy regression discontinuity model using the threshold indicator as an instrument for actual grade retention. This model identifies the local average treatment effect (LATE) of grade retention among those who were retained because they failed a July exam and complied with the decision. The two stage least squares model is as follows:

$$Retention_i = \alpha_0 + \alpha_1 FP_i + f(Index_i; \pi) + X_i'\phi + \eta_i \quad (3)$$

$$Convicted_i = \beta_0 + \beta_1 \widehat{Retention}_i + f(Index_i; \lambda) + X_i'\psi + \varepsilon_i, \quad (4)$$

where $Retention_i$ is an indicator variable that takes the value of one if student i is retained in eighth grade and all other variables are as previously defined. Standard errors in the IV and reduced form models are clustered by 8^{th} grade school.[17] The baseline bandwidth we use is 30 index points on either side of the cutoff, which corresponds to 64 percent of our initial sample. We show that our results are robust to using different bandwidths.

The variable of interest in equation (4) is β_1. This parameter is likely to differ from retention effects at earlier ages (Fruehwirth et al. 2016) and from effect of retention that occur for non-test-based reasons. Nonetheless, the LATE estimates pertain to an extremely relevant group of students, as they are the ones impacted by the test-based promotion policy. Our estimates therefore are more likely to generalize to other states that have similar policies for 8^{th} grade promotion.

The assumptions underlying the identification of β_1 are very similar to those for γ_1 in the reduced form model. The additional assumption invoked here is that the entire effect of failing

[16]These exam scores are only available beginning with the 1999-2000 academic year.

[17]Our estimated standard errors are similar if we instead cluster at the level of the running variable (see Online Appendix Tables A-4 and A-5). We favor clustering at the school level because of recent work by Kolesar and Rothe (2018) that shows standard error estimation has poor properties when clustering on the running variable in regression discontinuity settings.

a July exam on adult crime comes through the effect of failing on retention. If there are independent effects of failing, then β_1 will over-state the impact of grade retention. Such an independent effect could be driven by changes in a student's self-esteem or beliefs about the value of investing in education from failing the exam. We argue it is unlikely that such independent effects exist. As we show below, there is only weak evidence of an effect on juvenile crime, which should be highly impacted by these forces. The non-cognitive responses to failing also take several years to show up, which suggests that any self-esteem effects are not immediate and are likely driven by retention rather than failing the July exam per se. Below, we show both ITT and IV estimates for completeness, since ultimately we cannot test this assumption. The IV estimates are our preferred results, and we focus on them in the results section.

To shed light on the generalizability of the estimates to students not local to the cutoff, we employ the method proposed by Angrist and Rokkanen (2015). This method is based on a conditional independence assumption, which states that conditional on a set of observed characteristics, there is no relationship between the running variable and the outcomes of interest on either side of the cutoff. Under this assumption, one can use these observed characteristics to estimate counterfactual outcomes among both the treated and untreated samples. This in turn permits estimation of treatment effects away from the cutoff. We show that the conditional independence assumption holds in our data, which allows us to estimate the causal effect of retention on adult criminal convictions for inframarginal treated and untreated students.

4.2 Validity Checks

The core assumption supporting our RD approach is that students are unable to sort around the cutoff. Absent direct cheating behavior, it is very unlikely such sorting occurs. In this section, we show evidence that student characteristics move smoothly through the cutoff and that there is no evidence of systematic sorting. Table 2 presents reduced form estimates that exclude observed characteristics as controls and use them as dependent variables. The first column shows that failing a July exam does not affect the likelihood of being in the data. Thus, students are not leaving the state because they fail the test.[18] The next five columns show how

[18]The LDOE codes reasons students leave the school prior to graduation that are based on U.S. Department of Education guidelines. Reasons for no longer attending include death, attending a private school, being homeschooled, dropping out, or leaving the state. Table 2 focuses on the set of students coded as leaving the state.

a series of immutable or pre-exam characteristics vary as a function of the promotion cutoff. All of the estimates are small in absolute value, and none is statistically significant at even the 10% level. Thus, there is no evidence of student sorting differentially across the promotion threshold.

Figure 1 displays the distribution of test scores as a function of the index value. As argued by McCrary (2008), heaping right over the threshold is evidence of sorting since the running variable is expected to be distributed smoothly through the cutoff. The test score distribution exhibits heaping at several relative scores, which is due to the underlying test score distributions. There is no evidence in the figure of excess score heaping right over the cutoff. The test proposed by McCrary (2008) fails to reject the null hypothesis of a continuous distribution (p-value=0.89). The estimates in Table 2 and Figure 1, together with the lack of statistical evidence of heaping above the threshold, support our regression discontinuity approach by showing no evidence of student sorting relative to the promotion cutoff.

5 Results

5.1 Main Results

First-stage estimates that show how July test failure relates to grade retention are presented in Table 3. All estimates include linear splines in the running variable and use a bandwidth of 30 index points on either side of the cutoff. Column (1) includes only test year fixed effects and the running variable controls, while in column (2) we add student observed characteristics: race, gender, free/reduced price lunch status and immigrant status. In column (3), we include 8^{th} grade school fixed effects.

The estimates across columns are extremely similar and indicate that failing a July exam increases the likelihood of being retained by 68 percentage points. The estimates all are significant at the 1% level, and the first-stage F-statistics are large. Figure 2 presents these results graphically. There is a large, significant jump in the likelihood of being retained at the cutoff of 68 percentage points. The estimate is not 100 because some students who score above the cutoff are nevertheless retained because of disciplinary problems or lack of attendance, and some students who fail wind up being promoted after they appeal. In this fuzzy regressions discontinuity design, the July test score cutoff for promotion is a very strong instrument for grade retention.

Reduced form and IV estimates both are presented in Table 4. We examine whether individuals are convicted of any crime (Panel A), a violent crime (Panel B), a property crime (Panel C), or a drug-related crime (Panel D). It is possible to be convicted of multiple types of crime, so these categories are not mutually exclusive. Columns (1) and (3) present estimates using an indicator for being convicted of a crime by age 25 as the dependent variable. These estimates have been multiplied by 100 for ease of interpretation. In columns (2) and (4), we present estimates using the number of convictions (including zeros) at first conviction as an adult up to age 25. The first two columns of Table 4 show reduced form results, while the last two columns present IV estimates.[19]

The impact of failing a July exam on any conviction in Panel A is positive but not significant at conventional levels. In column (1), the point estimate suggests that failing a July exam increases the likelihood of being convicted of any crime by 0.85 of a percentage point. This is a 7.36% increase relative to the baseline mean shown in Table 1. The IV estimates also are positive, but similar to the reduced form results, they are not significantly different from zero. They suggest that retention increases the likelihood of an adult criminal conviction by 1.25 percentage points (10.85% relative to the mean).

The clearest evidence of an effect of grade retention on crime is obtained from examining violent crimes. Failing a July exam increases the likelihood of an adult violent crime conviction by 0.71 of a percentage point, which is 39.61% relative to the mean. The IV estimate shows that grade retention leads to a 1.05 percentage point (58.44%) higher probability of an adult violent crime conviction. Almost all of these convictions are for assault or robbery; there are few murder convictions in this sample. The violent crime estimates are individually significant at the 5% level. When we adjust the p-values for multiple hypothesis testing for the four crime outcomes using the Romano-Wolf stepdown procedure (Romano and Wolf 2005; 2016), the estimate is significant at the 13% level. These adjusted p-values are shown in brackets in Table 4.

There is no evidence of an effect on the likelihood of being convicted of a property crime:

[19]Online Appendix Table A-1 shows that the estimates are similar when we exclude controls for student characteristics.

estimates are negative, close to zero, and are not statistically significant at even the 10% level. However, we show in Section 5.3 that there is an effect of retention on property crime away from the cutoff.[20] Estimates for drug-related crimes are positive and non-trivial in size, though they are not statistically significant. The point estimates suggest that failing a July exam increases the likelihood of a drug conviction by 0.35 of a percentage point (6.76%) and grade retention increases the probability of a drug conviction by 0.51 of a percentage point (10.02%). Figure 3 presents these results graphically. As in Table 4, there is an increase in violent crime at the cutoff,[21] but the other crime categories show less evidence of an effect.

An indicator for whether any conviction occurs may not capture variation coming from the number of convictions upon first conviction. People are sometimes convicted of multiple offenses at once even within a given crime category. For example, someone can be convicted of both battery and armed robbery. One may reasonably interpret such convictions as representing more serious crimes than do convictions for one offense. In columns (2) and (4) of Table 4, we present reduced form and IV estimates using the number of convictions (including zeros) when we observe an individual being convicted of a crime for the first time as an adult up to age 25.[22] This intensive margin arguably better measures the severity of the crime than does a conviction indicator. The results are quantitatively and qualitatively similar to those in columns (1) and (3) of Table 4. Column (2) shows that failing a July exam increases the number of crimes committed by 0.017 (12.50% relative to the mean), although the estimate is not statistically significant. As displayed in column (4), being retained increases the number of adult convictions by 0.025, or 18.38%, which also is not significant at even the 10% level.

Similar to the extensive margin results, the intensive margin estimates indicate a large and statistically significant positive effect of July exam failure on violent crime convictions, no effect for property crimes, and a positive but not significant effect for drug convictions. The violent crime effect of failing the July exam is 0.012 additional crimes, which is 60.00% of the mean. Being retained increases the number of violent crimes of which one is convicted by

[20]We additionally show in Section 5.4 that there are positive effects of retention on robbery convictions. Robberies are technically violent crimes, but they share many characteristics with property crime (such as burglary).

[21]The discontinuity is difficult to see in Figure 3 because all of the panels are on the same scale. Online Appendix Figure A-1 shows the same figure with different scales across panels; the discontinuity in violent crime is evident.

[22]We cannot examine recidivism because we only can observe subsequent offenses up to age 25. A subset of convicted individuals are still in prison at age 25, and those who are incarcerated are most likely to re-offend. We thus are unable to measure the full distribution of recidivism.

0.017, or 85.00%. This estimate is significant at the 5% level even after adjusting for multiple hypothesis testing, with an adjusted p-value of 0.044.[23] We also find suggestive evidence of an increase in the number of drug convictions: retention increases the number of drug convictions at first conviction by 0.012 (20.00%), but the estimate is not significant at conventional levels. Graphical representations of the regression discontinuity are shown in Figure 4 for the number of conviction outcomes. They match the estimates in Table 4 closely.[24]

The results thus far indicate that grade retention increases the likelihood of being convicted of a violent crime and increases the number of violent crime convictions upon first conviction. We also find suggestive evidence of a positive effect on drug convictions. There are large private costs and externalities associated with these outcomes. We examine convictions, not arrests, which means that the increased criminal activity is both likely to be severe and lead to a permanent criminal record.[25]

To put the increase in violent crime into perspective, we combine the estimates in Donohue (2009) of the weighted average cost of an assault and robbery with the discontinuity estimate from our preferred specification in column (4) of Table 4. Taking the number of convicted adults from violent crimes in the control group as our benchmark, the 58.44% increase in violent crime implies that 104 more adults are convicted as a result of retention over our sample period. This increase corresponds to a \$2.6 million social cost using Donohue's lower bound estimates of assault and robbery and a \$18.4 million cost using the upper bound estimates. Considering that about half of violent incidents are reported to the police and about half of reported incidents result in an arrest, the actual social cost of retention likely is even larger (FBI 2012).

5.2 Mechanisms

There are several mechanisms that could lead grade retention to increase adult crime. Holding students back a year may affect their accumulation of non-cognitive skills. Indeed, a core

[23] We consider the set of null hypotheses for the multiple hypothesis correction to be the four crime categories separately for each dependent variable. We argue this is appropriate because the two dependent variables provide different measurements of the same underlying outcome: the intensive margin measure is a more fine-tuned version of the extensive margin. When we perform corrections assuming all eight outcomes are in the same family of null hypotheses, the p-value for the effect of retention on violent adult criminal convictions is 0.160 and is 0.070 for the number of convictions. P-values using this correction for the other estimates are available from the authors upon request.

[24] Online Appendix Figure A-2 shows the same figure with different scales across panels; the discontinuity in violent crime is more evident.

[25] Online Appendix Table A-11 shows suggestive evidence that retention increases sentence length by 3 months (off of a mean of 79 months) among those convicted. The average age of first conviction also rises slightly, as shown in column (2), but the estimate is imprecise.

argument for "social promotion" is that it is damaging for students to be in grades with those who are younger (Black et al. 2011; Cook and Kang 2016), and qualitative evidence indicates that grade retention is a determinant of student stress (Anderson, Jimerson, and Whipple 2005). If retention reduces human capital accumulation and/or reduces students' social fit in the school, it could lead to lower academic performance, more behavioral problems, and to eventual dropout from school. Retaining students also can alter the educational environment, in particular the set of peers to which a student is exposed.[26] If retention reduces the quality of schooling inputs, student achievement may decline. Additionally, it is possible that being retained sends a signal to the student that they are not well equipped for high school, which could cause them to drop out. Reductions in non-cognitive skills can increase the likelihood young adults engage in criminal behavior as well.

Tables 5-7 provide evidence on some of these potential mechanisms. In Table 5, we examine how grade retention affects the composition of students in the high school the student attends. Being retained in 8^{th} grade leads students to attend high schools in which the percentage of white students in 9^{th} grade is 2.7 percentage points lower (significant at the 12% level) and the mean 9^{th} grade test score is 6.2% of a standard deviation lower (significant at the 5% level). These findings indicate that retained students subsequently attend high schools that have lower-achieving peers and are less racially diverse.[27] The change in peer composition is facilitated by the availability of intra-district school choice in many districts throughout Louisiana. The ability of students to switch schools within a district is common across the US, so this mechanism is likely to generalize to other settings.[28] In order to assess the extent to which school sorting drives our estimates, in Online Appendix Table A-6 we present estimates that include fixed effects for the high school each student attends. The addition of high school fixed effects does not substantively change the estimates, which is suggestive of a minor role for peer composition as a mechanism for our main results.

The development of non-cognitive skills is particularly important to examine, since these

[26]The retention policy itself is unlikely to generate peer effects outside of shifting students across schools. While lower-performing students are taken out of each grade, they are added to the grade below. Hence, overall peer composition in a cohort is not affected on average by holding low-performing students back each year.

[27]Cook (2018) shows evidence that attending more segregated schools in which the percentage of African American students is higher can reduce short- and long-run educational attainment.

[28]Louisiana currently has extensive school choice policies in the form of vouchers and charter schools. These programs were established after Hurricane Katrina and thus are too late to impact the cohorts we study.

skills are highly correlated with the propensity to engage in criminal activity (Carneiro et al. 2007; Jackson et al. 2020). In Table 6, we analyze behavioral outcomes that reflect such skills: absenteeism and disciplinary incidents (Lounsbury et al. 2004). First, we examine whether retention leads to a change in the likelihood of observing information on these outcomes. Failing the July exam leads to a small but statistically significant increase in the likelihood of a student having missing data on these behavioral outcomes (Panel A).[29] If anything, this biases our estimates towards zero because those with behavioral problems are more likely to have missing data due to attrition. In columns (2) to (4) of Table 6, we estimate effects on students after 1, 2, and 3 years of grade retention to more closely examine the time pattern of any retention impacts. Students who are retained are absent 0.76 more days per year but are less likely to have a behavior incident in the first year following failure. Neither estimate is statistically significant, however. By the third year after retention, there is a statistically significant effect on absences of 1.97 days per year (11.14% relative to the mean). Students also are 1.12 percentage points more likely to have a disciplinary incident, although this estimate is not statistically different from zero at conventional levels. The year 3 estimate is statistically different from the year 1 estimate, which suggests grade retention induces a worsening pattern of behavior over time.

Grade retention also can lead to less education attainment, as shown in Table 7. The first column indicates that students who barely missed the promotion cutoff are not more likely to leave the public school system before their status is determined as a graduating or dropping out student.[30] Column (2) of Table 7 shows that students who are retained in 8^{th} grade are 7.17 percentage points more likely to drop out. This is a 15.66% increase relative to the baseline dropout likelihood of 45.8%. Online Appendix Table A-11 additionally shows that retention reduces the number of years of completed education post-July exam by 0.264 years.[31] Hjalmarsson et al. (2015) present the only estimates of which we are aware on the effect of educational attainment on violent crime convictions among young adults. They find no effect on

[29] Failing the July exam generates no change in the probability of a student having missing data on these behavioral outcomes after the first and second years of retention. This suggests that the attrition effect in Table 6 reflects higher dropout rates as students age rather than students leaving the state.

[30] The LDOE requires the use of three school years (the previous, current and the following) to identify the dropout status of a student. Therefore, a dropout flag is not complete until after the dropout correction period of the following year. Certain exit codes available in the administrative records exclude individuals from being coded as dropouts (e.g., death, out of state or private/home school) (Pastorek 2011). An individual is identified as a dropout only if he/she is flagged as one in the administrative records.

[31] Online Appendix Table A-11 also shows that retention has a positive but not statistically significant effect on being classified as special education.

the likelihood of committing any violent crime but a negative effect on the number of violent crime convictions for men (but not women). Using their results, we calculate that reduced educational attainment associated with grade retention predicts a 0.003 increase in the number of violent crime convictions.[32] Thus, reduced educational attainment explains only 18% of the total increase in the number of violent crimes and none of the increase in the likelihood of any violent crime conviction stemming from 8^{th} grade retention.

We supplement this evidence by estimating a regression of violent crime on a dropout indicator using the full set of controls (i.e., student demographics, school and year effects) for observations in the 30 index point bandwidth around the cutoff. The estimated correlation between violent crime and dropping out is 0.00017. Multiplying by the dropout estimate of 7.171 yields a predicted violent crime conviction effect of 0.001: increased prevalence of high school dropout explains only 6% of the total increase in violent crime stemming from 8th grade retention. These results are consistent with the calculations from Hjalmarsson et al. (2015) and indicate that high school dropout can explain only a small portion of the crime effect we estimate.

The last two columns of Table 7 present estimates of grade retention on juvenile crime and juvenile felony convictions.[33] Recall that our adult criminal conviction data exclude youth convictions, so there is no mechanical relationship between these two outcomes. The estimates in Table 7 are positive but not significant at even the 10% level. Furthermore, they are much smaller than the IV estimates in Table 4. Understanding the implication of these estimates for our results requires estimating how juvenile crime conviction affects adult criminal convictions. Aizer and Doyle (2015) use random assignment of judges to identify the effect of youth *incarceration* on adult incarceration.[34] They find that youth incarceration increases the likelihood of adult incarceration by 23.4 percentage points and of adult incarceration for violent crime by 14.9 percentage points. Using the latter estimate, and making the extreme assumption that all

[32]Hjalmarsson et al. (2015) examine the effect of educational attainment on criminal convictions in Sweden. While a different setting from the US, they emphasize that their findings are similar to US estimates from Lochner and Moretti (2004) for outcomes that overlap across the two papers. We calculate the implied effect of the educational attainment reduction from grade retention by weighting the gender-specific estimates for 19-29 year olds by the gender proportion in our sample and then multiplying by the number of years of reduced educational attainment in Table A-11: (0.52*0.004+0.48*-0.029)*-0.264=0.003.

[33]See Eren et al. (2017) for a more detailed analysis of the effect of the Louisiana accountability system on juvenile crime. They show that the total effect of summer school and potential grade retention is to lower juvenile crime, but the effect of retention per se is positive and not significantly different from zero.

[34]Sorensen, Bushway, and Gifford (2019) show that principal-driven student dismissals in middle school lead to more criminal convictions in early adulthood, but they do not examine how juvenile convictions affect subsequent adult convictions.

convicted youths are incarcerated, the estimates in Table 7 suggest that the juvenile conviction effect would increase adult violent crime convictions by between 0.050 and 0.079 percentage points. This is an upper bound of the juvenile crime effect, as Aizer and Doyle (2015) demonstrate the effect on adult criminal activity of a conviction is lower when it is not accompanied by incarceration.

Because individuals in our sample are exposed to conviction risk as an adult for a longer period of time than they are exposed to juvenile conviction risk, our results could simply reflect a longer risk of exposure. We provide two pieces of evidence against such an explanation. First, Online Appendix Figure A-3 presents conviction estimates by age that show the cumulative risk from ages 17-25. If our results were simply picking up a longer period of exposure, we would observe a linear increase in the effect by age. Instead, there is little effect among those under 20 and a discrete increase in the effect among those age 21-25 for violent crime. These results underscore the importance of examining crime outcomes in the early 20s. Second, Eren and Mocan (forthcoming) show that juvenile incarceration in Louisiana has no impact on future violent crime. Hence, our results are not predictable from an extrapolation of any juvenile crime effect and present independent information on the long-run effects of grade retention.

Taken together, the results from Tables 5-7 indicate that grade retention leads to lower peer quality in high school, increases behavioral problems and absenteeism several years after being retained, and increases the likelihood of dropout. There is little effect on juvenile crime, and any juvenile crime effect is too small to explain our adult conviction estimates. These results suggest that the increase in adult crime is not being driven by earlier engagement in the criminal justice system. We argue that this evidence is most consistent with students reducing investments in human capital over time and accumulating less non-cognitive skills, which makes it more likely they will engage in violent criminal behavior. Indeed, our results are consistent with the decline in violent crime from a summer jobs program (Heller 2014), which also was driven by changes in non-cognitive skill.[35] Those results relate to juvenile crime; our estimates indicate a similar pattern of effects for adult crime.

[35] Heller et al. (2017) show that several RCTs in Chicago designed to alter the decision-making of youth from disadvantaged backgrounds reduced crime overall and violent crime in particular. They find that the interventions caused students to "slow down" and make more measured decisions.

5.3 Heterogeneous Treatment Effects

The regression discontinuity estimator identifies the effect of retention at the promotion cutoff. From a policy perspective, it is important to understand whether our effects extend to other parts of the test score distribution. If effects are local to the cutoff, then it would be possible to mitigate or eliminate this adverse effect by moving the passing threshold. Conversely, if our results generalize to a larger group of lower-performing students, it means any reasonable test-based retention policy might increase criminal convictions later in life. If students below the threshold who are retained also engage in more crime later in life, the size of the externality implied by our estimates is much larger than if the effect is localized to the test score cutoff. We use the method proposed by Angrist and Rokkanen (2015) to estimate treatment effects away from the discontinuity.

The Angrist and Rokkanen (2015) approach relies on the ability to predict outcomes conditional on the running variable using observed characteristics. Under the "conditional independence" assumption that conditional on the observed characteristics the running variable is no longer related to outcomes, one can use estimated linear relationships between the observables and outcomes to predict counterfactual outcomes away from the discontinuity that can be used to estimate treatment effects in a linear reweighting framework (Kline 2011). Our prediction variables consist of indicators for gender, race, free/reduced lunch status and immigrant status, composite (ELA and math) seventh grade standardized test scores, and school fixed effects. Following Angrist and Rokkanen (2015), we test for the conditional independence assumption by estimating the relationship between outcomes and the running variable on each side of the cutoff, conditional on the set of observed characteristics. These tests are shown in Appendix Table A-12: the estimates are universally small, and none is statistically significantly different from zero at conventional levels. These results support the use of the linear reweighting method to estimate treatment effects away from the discontinuity.

Table 8 presents the results from implementing the Angrist and Rokkanen (2015) estimator. The first two columns show the effect of retention among students to the left of the test score cutoff who are retained as a result of the test, and columns (3) and (4) present estimates for those to the right of the cutoff who are not retained. Standard errors are calculated by taking the standard deviation of 500 block bootstrap replications at the school level.[36] The table displays consistent evidence of positive effects on the treated and untreated that are similar in magnitude to one another. For violent crime, retention increases the likelihood of an adult conviction by 1 percentage point and increases the number of convictions by over 0.01. That the effects are of similar size to each other and to the estimates in Table 4 suggests that the effect of retention on violent crime is more general among low-performing students. Any reasonable test-based cutoff is likely to generate an increase in violent crime later in life.

Panels B and C of Table 8 show effects on property and drug crimes that are larger than the RD estimates in Table 4. These results indicate an effect on property crime conviction of about 1.4 percentage points and an increase in the number of property crime convictions of 0.017. For drug convictions, retention increases the likelihood of conviction by about 1 percentage point and increases the number of convictions by 0.014. Only the property crime estimates are consistently statistically significant, however. These findings help resolve the somewhat puzzling result that retention does not affect property crime. The estimates in Table 8 shows that this is the case only at the cutoff: retention affects property crime convictions among a larger set of treated students away from the cutoff.

Online Appendix Figures A-4 through A-7 present these results graphically and demonstrate why the RD estimates are smaller than the effects away from the cutoff for property and drug crimes. The “extrapolated” counterfactual lines have zero slope, while the actual fitted lines slope upward or downward away from the cutoffs. As a result, the extrapolated effects grow away from the cutoffs. The local average treatment effects estimated with the RD model understate the impact of test-based promotion on drug and property crime: inframarginal retained students experience large increases in both types of crime convictions later in life. The same is not the case for violent crime, where neither the extrapolated nor the fitted lines have nonzero slopes. This causes the RD and extrapolated effects to be the same size.

The estimates in Table 8 have two important implications for the interpretation of our results. First, the RD estimates for property and drug crimes are understated. When extrapolated away from the cutoff, grade retention has large effects on all types of crime. Second, increasing the

[36] Online Appendix Table A-13 additionally shows the nonparametric 95% confidence interval calculated using the 2.5th and 97.5th estimates from the block bootstrap replications. The nonparametric confidence intervals align closely with the confidence intervals calculated from the standard error using a normal approximation.

passing threshold would not mitigate these effects. In fact, it likely would increase the size of the crime effect as the untreated students would exhibit similarly-sized or larger treatment effects if they were retained. The results from Table 8 indicate that the effect of grade retention on adult crime is large and ubiquitous among the low-performing students we examine.

Online Appendix Tables A-7 and A-8 explore heterogeneity by gender. The effects of grade retention on adult conviction are more pronounced for males and the point estimates are generally more precise. This result is sensible because the underlying likelihood that women are convicted of a crime is very low.[37] We next analyze if the violent crime results are driven by robberies. This investigation is motivated by two concerns. First, robbery is the most prevalent crime in the violent crime category. Second, although robbery is a violent crime because it involves the use of a weapon, robbery creates a financial return to the perpetrator, and therefore it is different from other violent crimes. Estimates of the effect of grade retention on robberies and other violent crimes are similar to one another.[38]

5.4 Robustness Checks

In this section, we explore the sensitivity of our results to alternative modeling assumptions. In Online Appendix Tables A-2 and A-3, we alter the bandwidth and the modeling of the running variable using criminal conviction and the number of convictions as dependent variables, respectively. The estimates for violent crime change little across specifications. The drug crime estimates are somewhat sensitive to specification, but they are not inconsistent with our baseline results.

We next examine how estimates are affected by the use of different samples, controls, and clustering assumptions. Column (1) in Online Appendix Tables A-4 and A-5 show results that use all convictions up to age 28 for our analysis sample.[39] The estimates are extremely similar to baseline. The remaining results shown in Appendix Tables A-4 and A-5 exclude the regions most affected by Hurricane Katrina and Hurricane Rita in column (2),[40] add 7^{th} and

[37] Online Appendix Tables A-7 and A-8 show estimates away from the cutoff by gender; non-parametric confidence intervals of these estimates are shown in Online Appendix Tables A-14 and A-15. As with the main results, estimated effects away from the cutoff are evident for men but not women.

[38] Specifically, we obtain an estimate of 0.504 (s.e.=0.302) for the likelihood of being convicted of a robbery due to grade retention.

[39] We only observe outcomes at age 28 for the 1984 birth cohort, at age 27 for the 1984-1985 birth cohorts, at age 26 for the 1984-1986 birth cohorts, and at age 25 for the 1984-1987 birth cohorts.

[40] Hurricane Katrina hit the Gulf Coast in August of 2005 and caused massive flooding and damage in areas of Louisiana around New Orleans. Hurricane Rita hit Louisiana less than a month after Katrina and caused additional flooding and damage in coastal

8^{th} grade March composite test score controls in columns (3) and (4), respectively, include 8^{th} grade school fixed effects in column (5), and cluster standard errors at the level of the running variable (i.e., the index) in column (6). These changes to the baseline model have minimal impacts on the results.

We also examine the effect of failing to achieve the March promotional standards on the likelihood of being convicted of an adult crime and the number of crimes of which individuals are convicted. The results, displayed in Online Appendix Table A-9, show that the impact of March discontinuity is very close to zero.[41]

To address concerns about compositional changes of students across the subject-specific distribution of the running variable, we use the math test score as the sole running variable. The estimated effects of grade retention on adult criminal convictions are very similar to baseline (Online Appendix Table A-10).

Finally, we conduct placebo tests in which we estimate our baseline model using false cutoffs at each test score over a range of 50 index points on either side of the true cutoff. We incrementally increase the cutoff from -50 to 50 by one point, and in Online Appendix Figures A-8 and A-9 we plot the distribution of point estimates. The vertical red line shows the effects from Table 4, and we report the percentage of placebo estimates that are larger than the baseline effects on the x-axis of each figure. For any violent crime conviction (Figure A-8) as well as for the number of violent crime and drug crime offenses (Figure A-9), the RD estimates are in the tails of the distribution of point estimates.

6 Conclusion

We present the first estimates in the literature of the effect of grade retention on adult criminal convictions. Our data come from a merge of administrative education records with administrative data on adult criminal convictions. To overcome the endogeneity of student retention, we exploit test score cutoffs imposed by the State of Louisiana for promotion to 9^{th} grade whereby a student must repeat 8^{th} grade if they fail the math or ELA state exam in July.

areas and in parishes most impacted by Katrina.

[41] We can use these results to calculate the effect of summer school on adult crime (Jacob and Lefgren 2004): $\gamma^S=\gamma^M$-γ_1*p(retained), where γ^M is the impact of failing the March exam, γ_1 is the effect of failing the July exam in equation (1), and p(retained) is the probability of retention. Summer school reduces the likelihood of violent crime conviction by -0.114 (=0.018-0.713*0.185) of a percentage point.

We first show that failing a July exam increases the likelihood of being retained by about 68 percentage points. Reduced form and IV estimates both show a large effect of test failure and subsequent retention on being convicted of a crime by age 25. Being retained increases the likelihood of being convicted of any crime by 1.25 percentage points, or 10.85% relative to the baseline mean, and increases the number of convictions by 0.025 (18.38%). The effects are largest for violent crime: retention increases the likelihood of a violent crime conviction by 1.05 percentage points (58.44%) and increases the number of violent crime convictions by 0.017 (85.00%). We find no evidence of an effect of retention on property crime and suggestive evidence of an increase in drug-related crime at the cutoff.

We next examine mechanisms that drive the effects of retention on adult criminal convictions. We find that retention leads students to subsequently attend high schools with lower-performing and more disadvantaged peers and substantially increases dropout. Our estimates also indicate that retention causes higher levels of absence from school three years later, and there is suggestive evidence of increased disciplinary incidents after 3 years. We do not find a statistically significant impact on juvenile crime, and the estimate is too small to explain much of the adult crime effect. The evidence is most consistent with retention lowering the rate of human capital accumulation, most importantly non-cognitive skills, which results in higher levels of criminal engagement in early adulthood. Using the method proposed by Angrist and Rokkanen (2015), we estimate effects away from the test passing threshold and find sizable increases in all types of crime due to grade retention among a broader set of low-performing students.

Our results indicate that retaining students in 8^{th} grade due to failing a promotion exam induces higher criminal activity as an adult among students on the margin of passing as well as among low-scoring inframarginal students. This evidence is consistent with a sizable effect of education on crime, and it suggests that the promotion policy, in its current form, harms at least a subset of students and creates important negative externalities in the form of more violent felonies being committed in the future. Importantly, our results speak to the effect of 8^{th} grade retention, and effects may differ for retention in earlier grades. Understanding whether this relationship exists for retention among younger students is an important question for subsequent research.

While our estimates may not generalize to younger students, they generalize to a larger set of inframarginal 8^{th} grade students. The effects we report would be detected at any potential promotion cutoff and is not driven by the particular threshold being set for promotion. The policy implications of these findings are important, as they suggest test-based retention in 8^{th} grade among low-performing students leads to more criminal engagement later in life. This generates substantial private and social costs: considering only violent crimes, back-of-the-envelope calculations imply a social cost between \$2.6 and \$18.4 million over the three cohorts of students we examine. Test-based promotion policies are designed to ensure that students meet basic knowledge levels to advance in school, which may have benefits to marginal and inframarginal students. Our results speak to the effect of retaining students relative to promoting them within this system rather than the effect of the accountability system itself. The accountability system induces a threat of retention, combined with summer school supports, which could generate positive impacts on both retained and promoted students. How to structure promotion and accountability policies in a way that retains the benefits while minimizing the costs to students and society is an important question for future work.

References

[1] Agan, Amanda and Sonja Starr. 2017. "Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment." *Quarterly Journal of Economics* 133(1): 191-235.

[2] Aizer, Anna and Joseph J. Doyle Jr. 2015. "Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges." *The Quarterly Journal of Economics* 130(2): 759-803.

[3] Anderson, David A. 1999. "The Aggregate Burden of Crime." *The Journal of Law and Economics* 42(2): 611-642.

[4] Anderson, Gabriella E., Shane R. Jimerson, and Angela D. Whipple. 2005. "Student Ratings of Stressful Experiences at Home and School." *Journal of Applied School Psychology* 21(1): 1-20.

[5] Anderson, D. Mark. 2014. "In School and Out of Trouble? The Minimum Dropout Age and Juvenile Crime." *Review of Economics and Statistics* 96(2): 318-331.

[6] Angrist, Joshua D. and Miikka Rokkanen. 2015. "Wanna Get Away? Regression Discontinuity Estimation of Exam School Effects Away From the Cutoff." *Journal of the American Statistical Association* 110(512): 1331-1344.

[7] Becker, Gary S. and Casey B. Mulligan. 1997. "The Endogenous Determination of Time Preference." *Quarterly Journal of Economics* 112(3): 729-758.

[8] Beuermann, Diether and C. Kirabo Jackson. 2019. "The Short and Long-Run Effects of Attending The Schools that Parents Prefer." Mimeo.

[9] Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2011. "Too Young to Leave the Nest? The Effects of School Starting Age." *Review of Economics and Statistics* 93(2): 455-467.

[10] Card, David. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica* 69(5): 1127-1160.

[11] Carneiro, Pedro, Claire Crawford, and Alissa Goodman. 2007. "The Impact of Early Cognitive and Non-Cognitive Skills on Later Outcomes." CEE Discussion Papers 0092, Centre for the Economics of Education, LSE.

[12] Charles, Kerwin Kofi and Ming Ching Luoh. 2010. "Male Incarceration, the Marriage Market, and Female Outcomes." *Review of Economics and Statistics* 92(3): 614-627.

[13] Chetty, Raj, John Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star." *Quarterly Journal of Economics* 126(4): 1593-1660.

[14] Chou, Shin-Yi, Jin-Tan Liu, Michael Grossman, and Ted Joyce. 2010. "Parental Education and Child Health: Evidence from a Natural Experiment in Taiwan." *American Economic Journal: Applied Economics* 2(1): 33-61.

[15] Cohodes, Sarah, Daniel Grossman, Samuel Kleiner and Michael F. Lovenheim. 2016. "The Effect of Child Health Insurance Access on Schooling: Evidence from Public Insurance Expansions." *Journal of Human Resources* 51(3): 727-759.

[16] Cook, Jason. 2018. "Race-Blind Admissions, School Segregation, and Student Outcomes: Evidence from Race-Blind Magnet School Lotteries." CESifo Working Paper No. 7335.

[17] Cook, Philip J. and Songman Kang. 2016. "Birthdays, Schooling, and Crime: Regression-discontinuity Analysis of School Performance, Delinquency, Dropout, and Crime Initiation." *American Economic Journal: Applied Economics* 8(1): 33-57.

[18] Cullen, Julie Berry, Brian A. Jacob and Steven Levitt. 2006. "The Effect of School Choice on Participants: Evidence from Randomized Lotteries." *Econometrica* 74(5): 1191-1230.

[19] Dee, Thomas S. 2004. "Are There Civic Returns to Education?" *Journal of Public Economics* 88(9-10): 1697-1720.

[20] Deming, David J. 2011. "Better Schools, Less Crime?" *The Quarterly Journal of Economics* 126(4): 2063-2115.

[21] Deming, David J., Sarah Cohodes, Jennifer Jennings, and Christopher Jencks. 2013. "School Accountability, Postsecondary Attainment and Earnings." NBER Working Paper No. 19444.

[22] Donohue, John J. III. 2009. "Assessing the Relative Benefits of Incarceration: Overall Changes and the Benefits on the Margin." In S. Raphael and M. A. Stoll (Eds.), *Do Prisons Make Us Safer? The Benefits and Costs of the Prison Boom*, pp. 269.342. Russell Sage Foundation.

[23] Eren, Ozkan, Briggs Depew and Stephen Barnes. 2017. "Test-based Promotion Policies, Dropping Out, and Juvenile Crime." *Journal of Public Economics* 153: 9-31.

[24] Eren, Ozkan and Naci Mocan. Forthcoming. "Juvenile Punishment, High School Graduation and Adult Crime: Evidence from Idiosyncratic Judge Harshness." *Review of Economics and Statistics.*

[25] Federal Bureau of Investigation. 2012. "Crime in the United States 2012," Uniform Crime Report-U.S. Department of Justice: Washington, DC.

[26] Ferreira, Eduardo Ferraz Castelo Branco and Rodrigo Reis Soares. 2018. "Socially Optimal Crime and Punishment." Escola de Economia de Sao Paulo Dissertation Thesis.

[27] Fruehwirth, Jane Cooley, Salvadore Navarro, and Yuya Takahashi. 2016. "How the Timing of Grade Retention Affects Outcomes: Identification and Estimation of Time-Varying Treatment Effects." *Journal of Labor Economics* 34(4): 972-1021.

[28] Fudenberg, Drew and David K. Levine. 2006. "A Dual-self Model of Impulse Control." *American Economic Review* 96(5): 1449-1476.

[29] Gelman, Andrew and Guido Imbens. 2017. "Why High-order Polynomials Should not be Used in Regression Discontinuity Designs." *Journal of Business & Economic Statistics.*

[30] Greene, Jay P. and Marcus A. Winters. 2007. "Revisiting Grade Retention: An Evaluation of Florida's Test-Based Promotion Policy." *Education Finance and Policy* 2(4): 319-340.

[31] Grossman, Michael. 2006. "Education and Nonmarket Outcomes." In E. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Chapter 10, pp. 577-633. Elsevier.

[32] Heller, Sara B. 2014. "Summer Jobs Reduce Violence Among Disadvantaged Youth." *Science* 346(6214): 1219-1223.

[33] Heller, Sara B., Anuj K. Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan, and Harold A. Pollack 2017. "Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago." *Quarterly Journal of Economics* 132(1): 1-54.

[34] Hirschi, Travis and Michael Gottfredson. 1983. "Age and the Explanation of Crime." *American Journal of Sociology* 89(3): 552-584.

[35] Hjalmarsson, Randi, Helena Holmlund, and Matthew J. Lindquist. 2015. "The Effect of Education on Criminal Convictions and Incarceration: Causal Evidence from Microdata." *The Economic Journal* 125(587): 1290-1326.

[36] Imbens, Guido and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79(3): 933-959.

[37] Jackson, C. Kirabo. Forthcoming. "Can Introducing Single-Sex Education into Low-Performing Schools Improve Academics, Arrests, and Teen Motherhood?" *Journal of Human Resources.*

[38] Jackson, C. Kirabo, Shanette C. Porter, John Q. Easton, Alyssa Blanchard, and Sebastin Kiguel. 2020. "School Effects on Socio-emotional Development, School-based Arrests, and Educational Attainment." NBER Working Paper No. w26759.

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

[40] Johnson, Rucker C. and Steven Raphael. 2009. "The Effects of Male Incarceration Dynamics on Acquired Immune Deficiency Syndrome Infection Rates among African American Women and Men." *The Journal of Law and Economics* 52(2): 251-293.

[41] Kline, Patrick. 2011. "Oaxaca-Blinder as a Reweighting Estimator." *American Economic Review* 101(3): 532-537.

[42] Kolesr, Michal and Christoph Rothe. 2018. "Inference in Regression Discontinuity Designs with a Discrete Running Variable." *American Economic Review* 108(8): 2277-2304.

[43] Lleras-Muney, Adriana. 2005. "The Relationship Between Education and Adult Mortality in the United States." *Review of Economic Studies* 72(1): 189-221.

[44] Lochner, Lance. 2004. "Education, Work, and Crime: A Human Capital Approach." *International Economic Review* 45(3): 811-843.

[45] Lochner, Lance. 2011. "Nonproduction Benefits of Education: Crime, Health, and Good Citizenship." in Eric A. Hanushek, Stephen J. Machin and Ludger Woessman (Eds.) *Handbook of the Economics of Education, Volume 4*. Elsevier: Amsterdam

[46] Lochner, Lance and Enrico Moretti. 2004. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." *American Economic Review* 94(1): 155-189.

[47] Lounsbury, John W., Robert P. Steel, James M. Loveland, and Lucy W. Gibson. 2004. "An Investigation of Personality Traits in Relation to Adolescent School Absenteeism." *Journal of Youth and Adolescence* 33(5): 457-466.

[48] Lovenheim, Michael F. and Alexander Willén. 2019. "The Long-run Effect of Teacher Collective Bargaining." *American Economic Journal: Economic Policy* 11(3): 292-324.

[49] Machin, Stephen, Olivier Marie, and Suncica Vujic. 2011. "The Crime Reducing Effect of Education." *Economic Journal* 121(552): 463-484.

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

[51] Mariano, Louis T., Paco Martorell, and Tiffany Tsai. 2018. "The Effects of Grade Retention on High School Outcomes." Mimeo.

[52] Marsh, Julie A., Daniel Gershwin, Sheila N. Kirby, and Nailing Xia. 2009. "Retaining Students in Grade-Lessons Learned Regarding Policy Design and Implementation." RAND Education. Technical Report.

[53] Matsudaira, Jordan. 2008. "Mandatory Summer School and Student Achievement." *Journal of Econometrics* 142(2): 829-850.

[54] McLaughlin, Michael, Carrie Pettue-Davis, Derek Brown, Chris Veeh, and Tanya Renn. 2016. *The Economic Burden of Incarceration in the US.* Washington University: Seattle.

[55] Oreopoulos, Philip. 2007. "Do Dropouts Drop Out Too Soon? Wealth, Health and Happiness from Compulsory Schooling." *Journal of Public Economics* 91(11-12): 2213-2229.

[56] Oreopoulos, Philip and Kjell G. Salvanes. 2011. "Priceless: The Nonpecuniary Benefits of Schooling." *Journal of Economic Perspectives* 25(1): 159-184.

[57] Ou, Suh-Ruu and Arthur J. Reynolds. 2010. "Grade Retention, Postsecondary Education, and Public Aid Receipt." *Education Evaluation and Policy Analysis* 32(1): 118-139.

[58] Pastorek, P. 2010. "LEAP Remediation Programs." Louisiana Department of Education.

[59] Romano, Joseph P. and Michael Wolf. 2005. "Stepwise Multiple Testing as Formalized Data Snooping." *Econometrica* 73(4): 12371282.

[60] Romano, Joseph P. and Michael Wolf. 2016. "Efficient Computation of Adjusted P-values for Resampling-based Stepdown Multiple Testing." *Statistics & Probability Letters* 113: 3840.

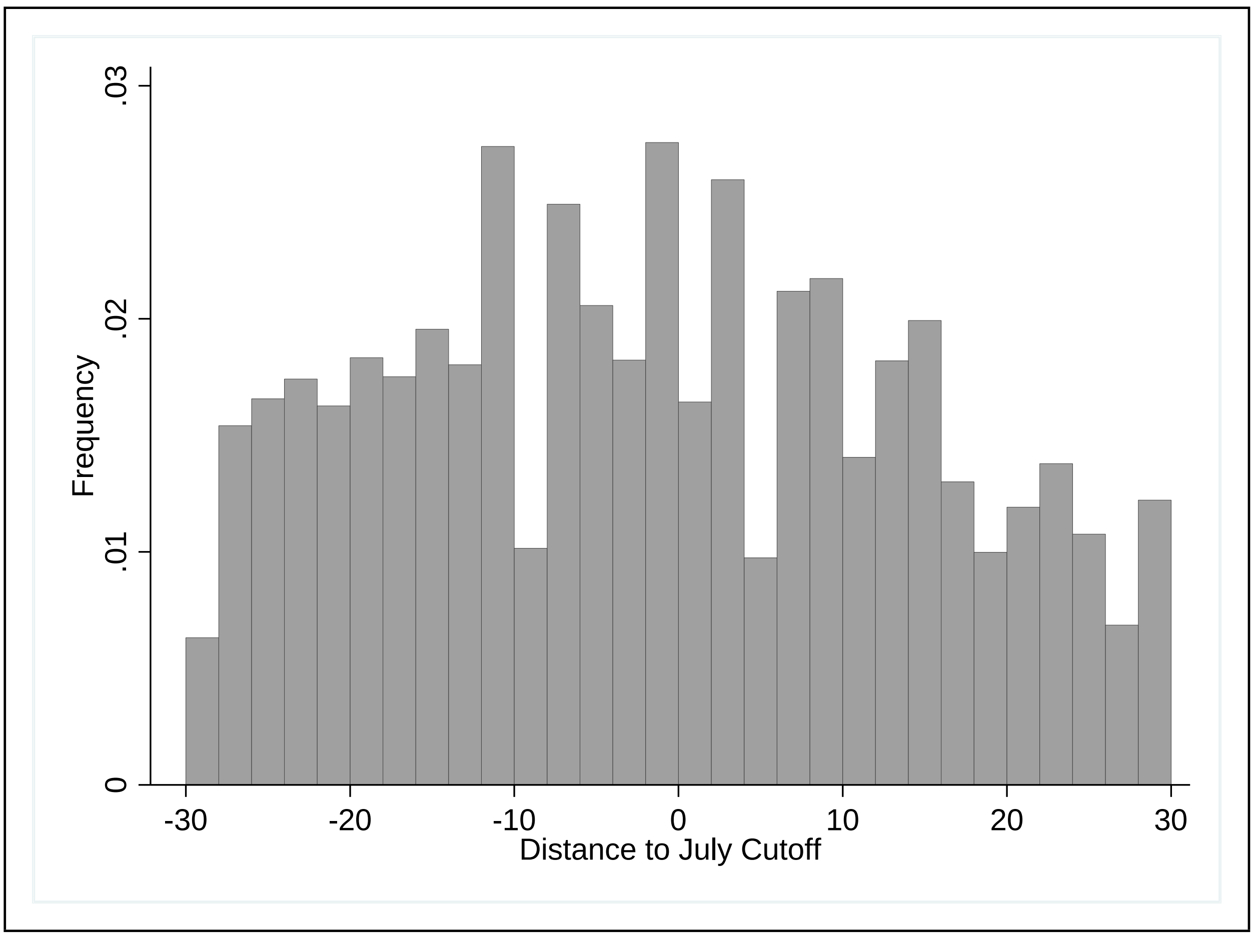
[61] Schwerdt, Guido, Martin R. West and Marcus A. Winters. 2017. “The Effects of Test-based Retention on Student Outcomes over Time: Regression Discontinuity Evidence from Florida.” *Journal of Public Economics* 152: 154-169.

[62] Sorensen, Lucy C., Shawn Bushway, and Elizabeth J. Gifford. 2019. “Getting Tough? The Effects of Discretionary Principal Discipline on Student Outcomes.” Mimeo.

[63] Workman, Emily. 2014. *Third-Grade Reading Policies.* Education Commission of the States.

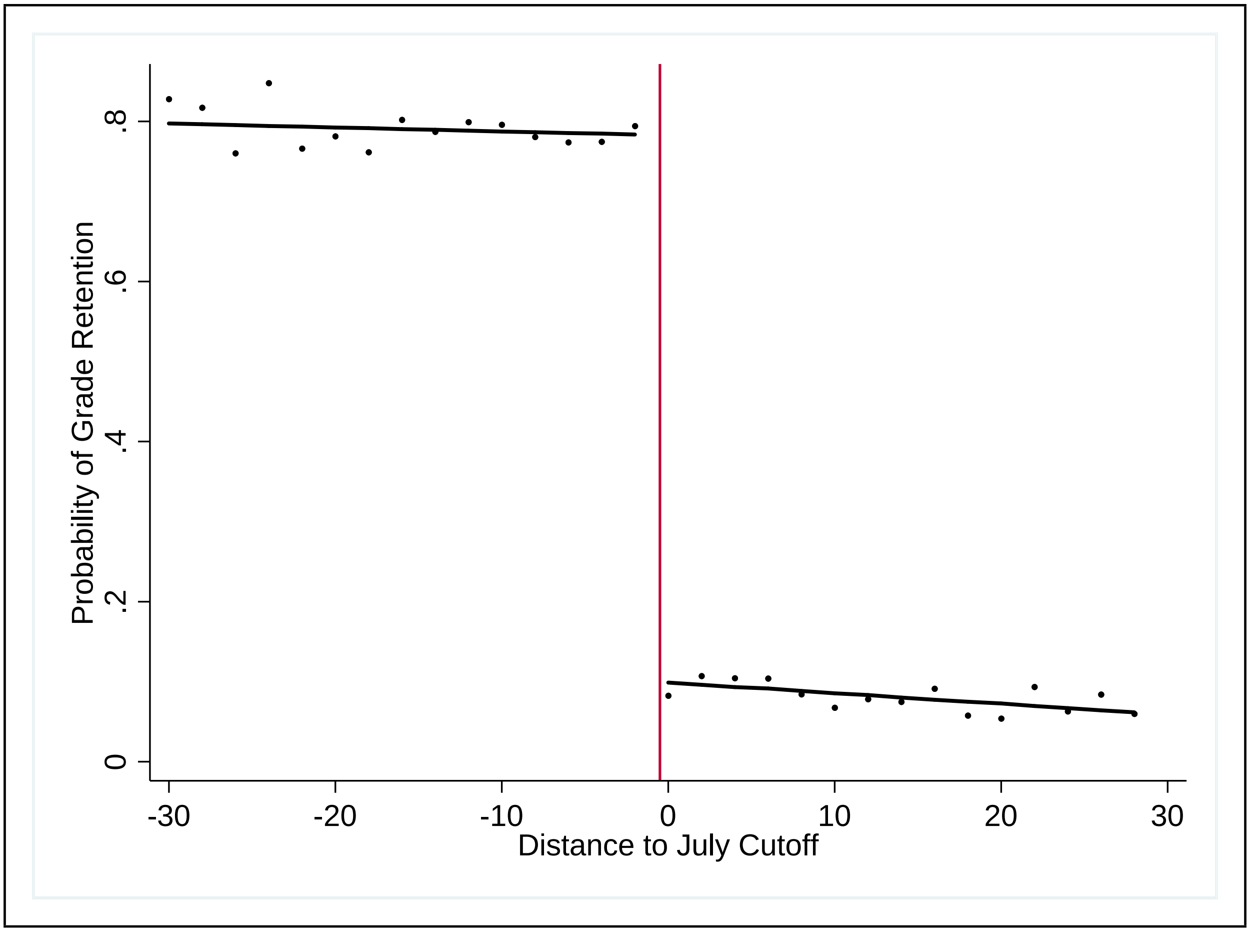
[64] Zinth, Kyle. 2005. *Student Promotion/Retention Policies.* Education Commission of the States.

Figure 1: Distributions of Students around the July Promotional Cutoff



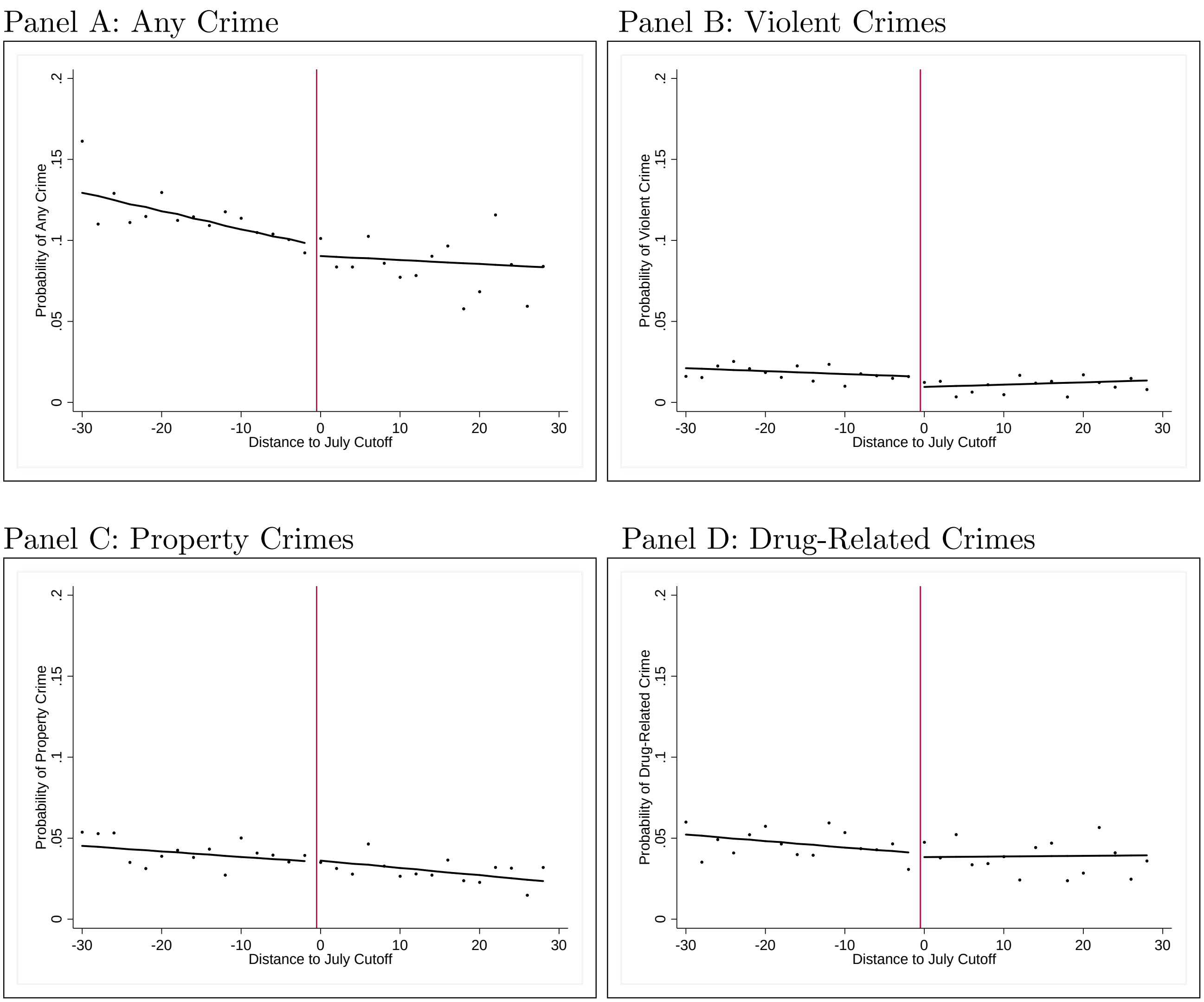
The figure shows the distribution of index scores relative to the July promotional cutoff (centered at zero), using the analysis sample with a bandwidth of 30 as described in the text.

Figure 2: Probability of Grade Retention and Distance to the July Promotional Cutoff



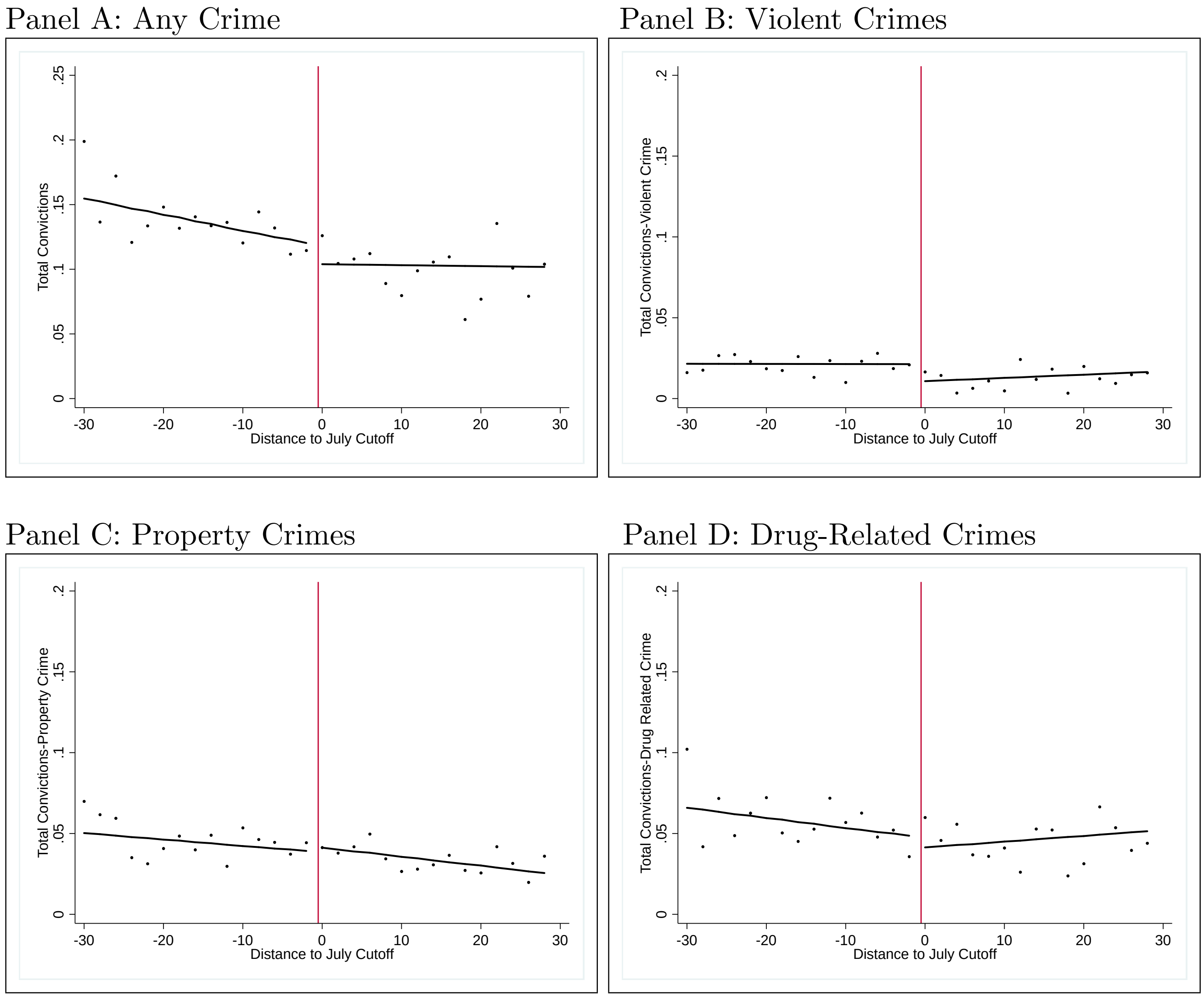
The vertical line denotes -0.5 points to the left of the July promotional cutoff (centered at 0). Each circle represents the unconditional mean of grade retention in two index point bins, based on the distance to July cutoff. The solid lines are fitted values of probability of grade retention from a linear spline over an index bandwidth of 30 points.

Figure 3: Any Criminal Convictions by Distance to the July Promotional Cutoff



The vertical line denotes -0.5 points to the left of the July promotional cutoff (centered at 0). Each circle represents the unconditional mean of adult crime conviction and types in two index point bins, based on the distance to July cutoff. The solid lines are fitted values of probability of criminal conviction from a linear spline over an index bandwidth of 30 points.

Figure 4: The Number of Adult Crime Convictions by Distance to the July Test Cutoff



The vertical line denotes -0.5 points to the left of the July promotional cutoff (centered at 0). Each circle represents the unconditional mean of the number of adult crime convictions, overall and by type, in two index point bins based on the distance to July cutoff. The solid lines are fitted values of the number of criminal convictions from a linear spline over an index bandwidth of 30 points.

Table 1: Summary Statistics

	Full Sample		Promoted		Retained	
	Mean	SD	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Student Characteristics						
Black	0.763	0.425	0.661	0.473	0.854	0.352
White	0.218	0.413	0.315	0.464	0.131	0.337
Female	0.520	0.499	0.504	0.500	0.535	0.499
Free/Reduced Lunch	0.454	0.498	0.473	0.499	0.438	0.496
Prior (7th Grade) ITBS Math z-score	-0.857	0.549	-0.727	0.569	-0.975	0.502
Prior (7th Grade) ITBS ELA z-score	-0.685	0.659	-0.586	0.695	-0.774	0.610
8th Grade LEAP Test Outcomes						
Failed Only ELA (July LEAP Exam)	0.039	0.194	0.027	0.162	0.050	0.21
Failed Only Math (July LEAP Exam)	0.452	0.498	0.189	0.391	0.689	0.463
Failed ELA and Math (July LEAP Exam)	0.173	0.379	0.133	0.340	0.209	0.407
Other Outcomes						
Drop Out of School	0.458	0.498	0.372	0.483	0.535	0.499
Committed a Juvenile Crime	0.042	0.199	0.039	0.193	0.044	0.206
Panel B: Adult Characteristics/Outcomes						
Any Adult Conviction (Extensive Margin)	0.115	0.319	0.110	0.313	0.119	0.324
Number of Convictions (Intensive Margin)	0.136	0.416	0.142	0.427	0.130	0.403
Adult Crime Type (Extensive Margin):						
Violent	0.018	0.131	0.016	0.127	0.019	0.136
Property	0.039	0.193	0.038	0.192	0.039	0.195
Drug Related	0.051	0.219	0.048	0.213	0.053	0.225
Other	0.010	0.102	0.011	0.103	0.010	0.100
Adult Crime Type (Intensive Margin):						
Violent	0.020	0.166	0.018	0.149	0.022	0.180
Property	0.043	0.222	0.043	0.226	0.043	0.219
Drug Related	0.060	0.248	0.057	0.277	0.063	0.290
Other	0.013	0.124	0.012	0.120	0.013	0.128
Age of Adult Crime	20.83	2.81	20.77	2.90	20.88	2.72
Sample Size	22,929		10,868		12,061	

The tabulations reflect our research sample as described in the text, which consists of students enrolled in regular classes in grade 8 between the 1998-1999 and 2000-2001 academic years. The students in the sample took July ELA or math (or both) LEAP exams in eighth grade and did not move out of the state. Prior achievement scores are available beginning with the 1999-2000 academic year and are standardized with respect to the statewide mean and standard deviation by test year, separately for each subject. The "Extensive Margin" measures whether adults have been convicted of a given type of crime, while the "Intensive Margin" measures the number of crimes of which people are convicted at first conviction (including zeros). The full set of sample statistics is available upon request from the authors.

Table 2: Regression Discontinuity Validation Tests

	Dependent Variable:					
	Moved Out of State (Attrition) (1)	Female (2)	Black (3)	White (4)	Free/ Reduced Lunch (5)	7^{th} Grade Composite Test Score (6)
Failed July Promotion Cutoff	-0.003	0.011	-0.011	0.016	-0.017	0.006
	(0.007)	(0.014)	(0.014)	(0.013)	(0.014)	(0.013)
Sample Size	15,501	14,728	14,728	14,728	14,728	14,378

Authors' estimation of equation (1) as described in the text. The sample is restricted to students who scored within 30 points of one of the cutoffs on the July promotion exams. All specifications control for linear splines in index score as well as test year fixed effects. Column (1) includes students who moved out of state, while the other columns exclude these students. The 7^{th} grade composite score is the average of the standardized test scores in ELA and math from this grade. These scores are available beginning with the 1999-2000 academic year and are standardized against the statewide mean and standard deviation, separately by test year and subject. Standard errors shown in parentheses and are clustered at the school level: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 3: First Stage Estimates of the Effect of Failing to Meet a July Promotion Cutoff on Grade Retention

Dependent Variable: Retained in 8^{th} Grade	(1)	(2)	(3)
Failed July Promotion Cutoff	0.677***	0.678***	0.676***
	(0.018)	(0.018)	(0.019)
First-Stage F-Statistic	1,352.53	1,357.78	1,274.78
Sample Size	14,728	14,728	14,728
Controls:			
Test Year Fixed Effects	Yes	Yes	Yes
Covariates	No	Yes	Yes
School Fixed Effects	No	No	Yes

Authors' estimation of equation (3) as described in the text. The sample is restricted to students who scored within 30 points of one of the cutoffs on the July promotion exams. All specifications control for linear splines in index score. Covariates include indicators for gender, race, free/reduced lunch and immigrant status. Standard errors are shown in parentheses and are clustered at the school level: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 4: Reduced Form and IV Estimates of Test Failure and Grade Retention on Adult Crime Conviction

	Reduced Form		IV	
Margin:	Extensive (x100)	Intensive	Extensive (x100)	Intensive
	(1)	(2)	(3)	(4)
Panel A: Any Crime				
Failed July Promotion Cutoff	0.846	0.017		
	(0.864)	(0.012)		
	[0.624]	[0.320]		
Grade Retention			1.248	0.025
			(1.270)	(0.018)
			[0.626]	[0.318]
Panel B: Violent Crime				
Failed July Promotion Cutoff	0.713**	0.012**		
	(0.351)	(0.005)		
	[0.130]	[0.046]		
Grade Retention			1.052**	0.017**
			(0.516)	(0.007)
			[0.130]	[0.044]
Panel C: Property Crime				
Failed July Promotion Cutoff	-0.057	-0.002		
	(0.573)	(0.007)		
	[0.928]	[0.716]		
Grade Retention			-0.084	-0.003
			(0.845)	(0.010)
			[0.930]	[0.712]
Panel D: Convicted of a Drug-Related Crime				
Failed July Promotion Cutoff	0.346	0.008		
	(0.586)	(0.008)		
	[0.796]	[0.528]		
Grade Retention			0.511	0.012
			(0.862)	(0.011)
			[0.794]	[0.524]

Authors' estimation of equations (1) and (4) as described in the text. The sample is restricted to students who scored within 30 points of one of the cutoffs on the July promotion exams; N=14,728. All specifications control for linear splines in index score, test year fixed effects, and indicators for gender, race, free/reduced lunch and immigrant status. The dependent variable in columns (1) and (3) are indicators taking the value of one if the student is convicted as an adult by age 25. The dependent variable in columns (2) and (4) are the number of crimes for which the individual is convicted at first conviction by age 25 (including zeros). Outcome variables are multiplied by 100 in columns (1) and (3) to obtain percent values. Standard errors clustered at the school level are in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%. P-values adjusted for multiple hypothesis testing using the Romano-Wolf procedure assuming the family of null hypotheses consists of the four crime groups (any, violent, property, drug) separately for each dependent variable are shown in brackets.

Table 5: Mechanism - The Effect of Grade Retention on High School Peer Composition

	Percent of Special Ed Students in High School	Percent of White Students in High School	Average 9^{th} Grade Math Standardized Test Score
	(1)	(2)	(3)
Grade Retention	-0.001	-0.027	-0.062**
	(0.004)	(0.017)	(0.027)
Sample Size	11,115	11,115	11,115

Authors' estimation of equation (4) as described in the text. The sample is restricted to students who scored within 30 points of one of the cutoffs on the July promotion exams. All specifications control for linear splines in index score as well as test year fixed effects. Covariates include indicators for gender, race, free/reduced lunch and immigrant status. Standard errors clustered at the school level are in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 6: Mechanism - Regression Discontinuity Estimates of Grade Retention on Absenteeism and Disciplinary Incidents

	Missing Outcome Data	Elapsed Time Relative to July Exam		
		Year 1	Year 2	Year 3
	(1)	(2)	(3)	(4)
Panel A: Missing Data				
Failed July Promotion Cutoff	0.022***			
	(0.011)			
Panel B: Total Days Absent from School				
Grade Retention		0.762	1.006	1.967***
		(0.709)	(0.923)	(0.750)
Mean-Total Days Absent		[13.40]	[17.30]	[17.66]
Panel C: Any Disciplinary Incident (x100)				
Grade Retention		-2.338	0.915	1.124
		(2.408)	(2.356)	(2.114)
Mean-Disciplinary Incident-%		[35.97]	[37.13]	[32.67]
Sample Size	15,501	12,963	12,963	12,963

Authors' estimation of equations (1) and (4) as described in the text. The sample is restricted to students who scored within 30 points of one of the cutoffs on the July promotion exams. All specifications control for linear splines in index score as well as test year fixed effects. Missing Outcome Data indicator takes the value of one if total absenteeism/disciplinary involvement information is missing in any of the three years following the July exam. Disciplinary incidents include any actions involving in-school suspension, out-of-school suspension and expulsion. Covariates include indicators for gender, race, free/reduced lunch and immigrant status. The outcome in Panel C is multiplied by 100 to obtain percent values. Standard errors clustered at the school level are in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 7: Mechanism - Regression Discontinuity Estimates of Grade Retention on Dropping Out of School and Juvenile Crime

	HS Attrition (1)	Drop Out of School (2)	Juvenile Conviction (3)	Juvenile Felony (4)
Failed July Promotion Cutoff	0.000			
	(0.011)			
Grade Retention		7.171***	0.528	0.335
		(2.334)	(0.811)	(0.492)
Sample Size	15,501	12,963	14,728	14,728

Authors' estimation of equations (1) and (4) as described in the text. The sample is restricted to students who scored within 30 points of one of the cutoffs on the July promotion exams. All specifications control for linear splines in index score as well as test year fixed effects. Covariates include indicators for gender, race, free/reduced lunch and immigrant status. HS Attrition in Column (1) takes the value of one if student left the public school system before dropout status is determined. Standard errors clustered at the school level are in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 8: The Effect of Retention on Adult Criminal Convictions Away from the Cutoff

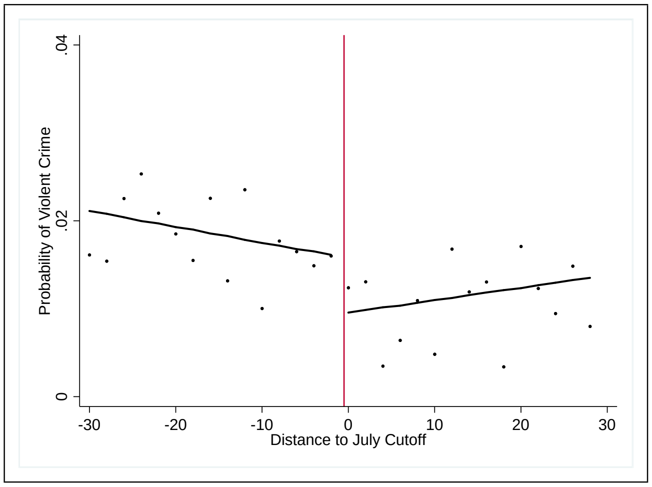
	Treatment Effect on the Treated (Extensive Margin) (1)	Treatment Effect on the Treated (Intensive Margin) (2)	Treatment Effect on the Untreated (Extensive Margin) (3)	Treatment Effect on the Untreated (Intensive Margin) (4)
Panel A: Violent Crime				
Grade Retention	1.070***	0.013***	0.976***	0.011***
	(0.283)	(0.003)	(0.300)	(0.004)
Panel B: Property Crime				
Grade Retention	1.462***	0.017***	1.424***	0.017***
	(0.454)	(0.005)	(0.520)	(0.006)
Panel C: Drug-Related Crime				
Grade Retention	0.962	0.015*	0.939*	0.014*
	(0.641)	(0.008)	(0.565)	(0.008)
Sample Size	14,378	14,378	14,378	14,378

Authors' estimation as described in the text. The sample is restricted to students who scored within 30 points of one of the cutoffs on the July promotion exams. Average treatment effect estimates are obtained using a linear reweighting estimator. Covariates include indicators for gender, race, free/reduced lunch, immigrant status, composite test scores and school and test year fixed effects. Outcome variables in Columns (1) and (3) are multiplied by 100 to obtain percent values, while total number of adult convictions up to age 25 are used in Columns (2) and (4). Standard errors computed by taking the standard deviation of 500 nonparametric block bootstrap replications at the school level are in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

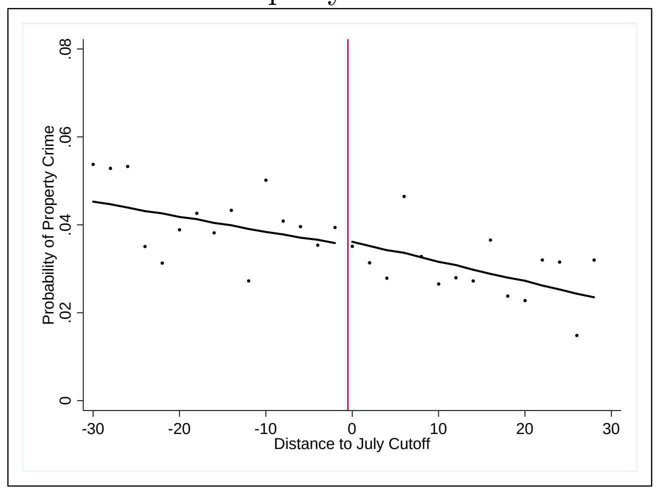
Online Appendix: Not for Publication

Figure A-1: Any Criminal Convictions by Distance to the July Promotional Cutoff - Different Scales

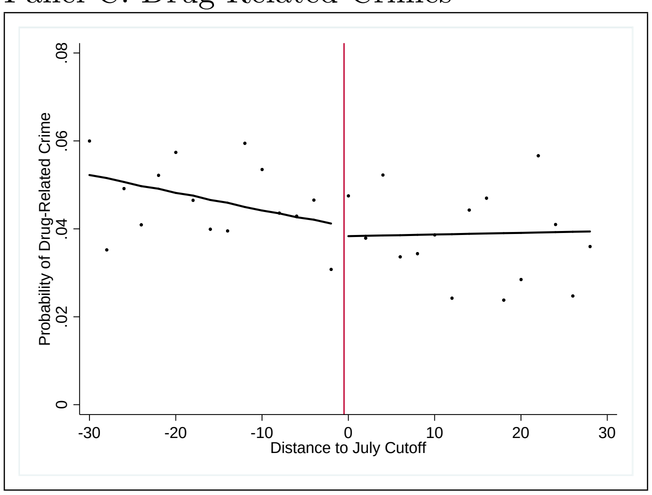
Panel A: Violent Crimes



Panel B: Property Crimes



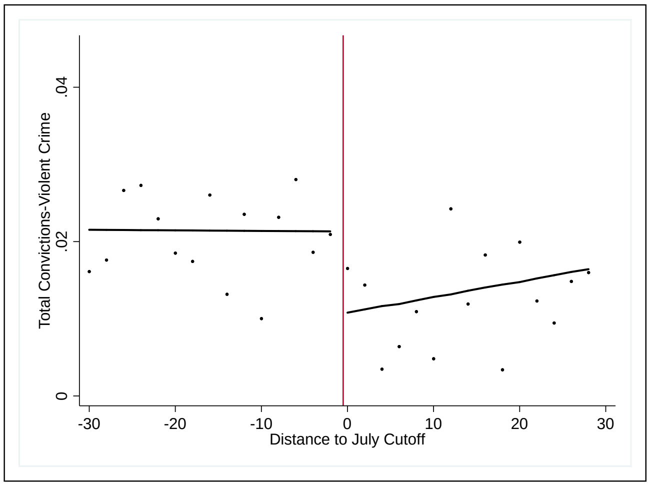
Panel C: Drug-Related Crimes



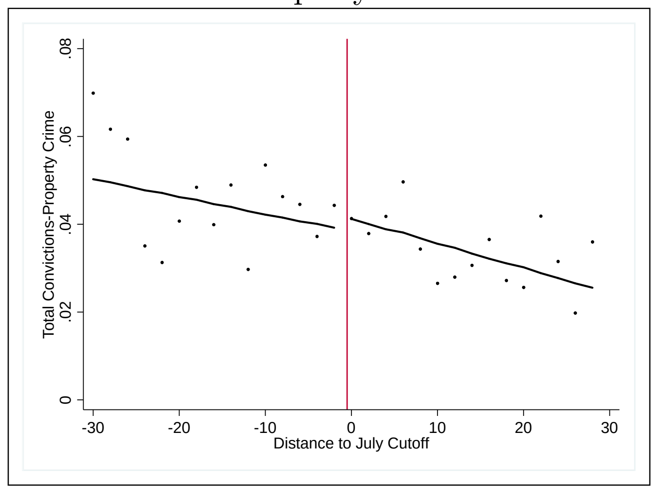
The vertical line denotes -0.5 points to the left of the July promotional cutoff (centered at 0). Each circle represents the unconditional mean of adult crime conviction and types in two index point bins, based on the distance to July cutoff. The solid lines are fitted values of probability of criminal conviction from a linear spline over an index bandwidth of 30 points.

Figure A-2: The Number of Adult Crime Convictions by Distance to the July Test Cutoff - Different Scales

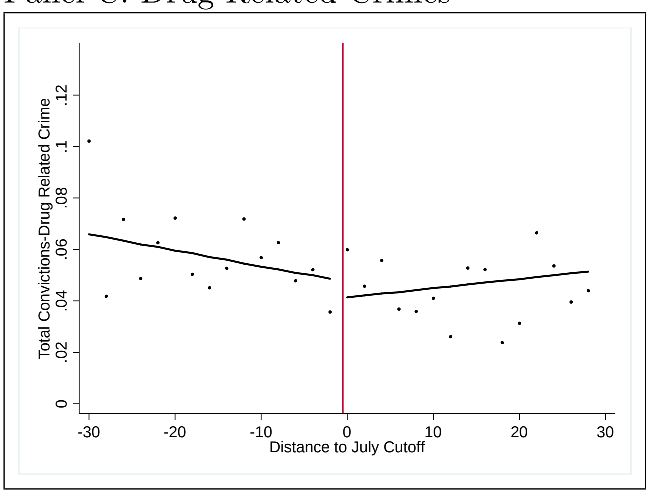
Panel A: Violent Crimes



Panel B: Property Crimes



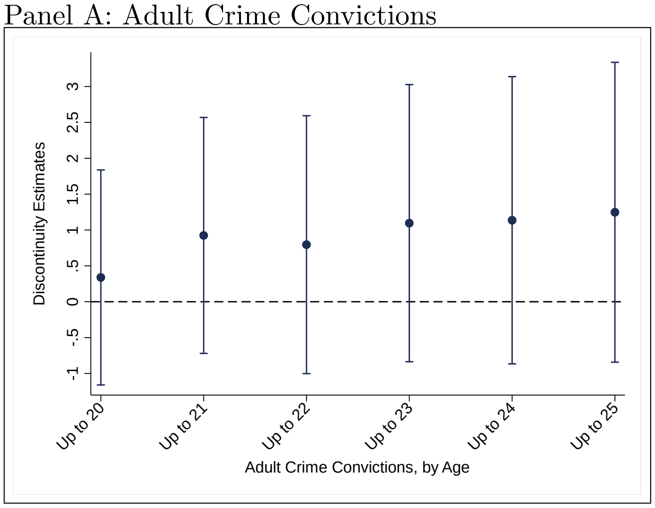
Panel C: Drug-Related Crimes



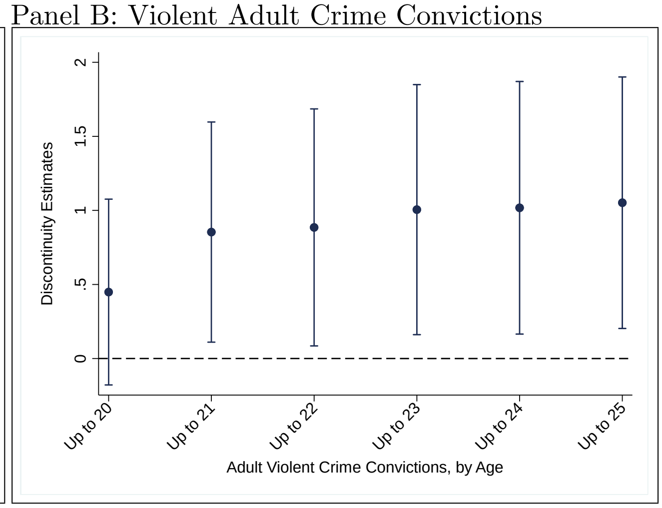
The vertical line denotes -0.5 points to the left of the July promotional cutoff (centered at 0). Each circle represents the unconditional mean of the number of adult crime convictions, overall and by type, in two index point bins based on the distance to July cutoff. The solid lines are fitted values of the number of criminal convictions from a linear spline over an index bandwidth of 30 points.

Figure A-3: The Effect of Grade Retention on Adult Criminal Convictions, by Age

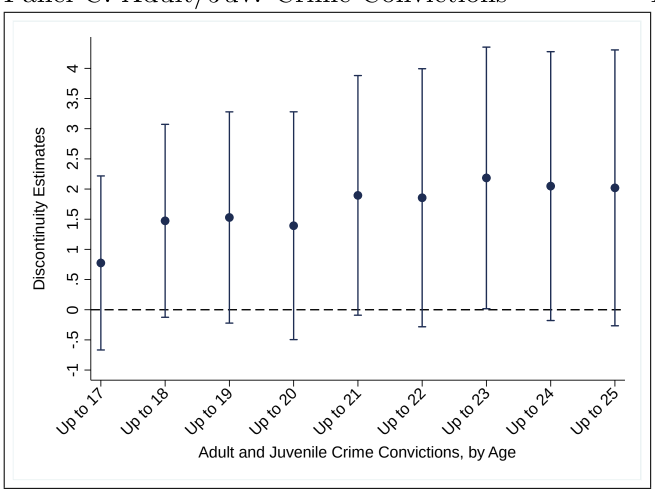
Panel A: Adult Crime Convictions



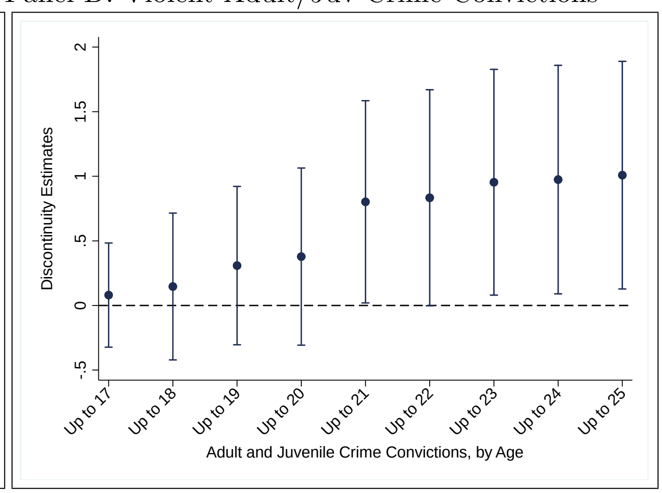
Panel B: Violent Adult Crime Convictions



Panel C: Adult/Juv. Crime Convictions



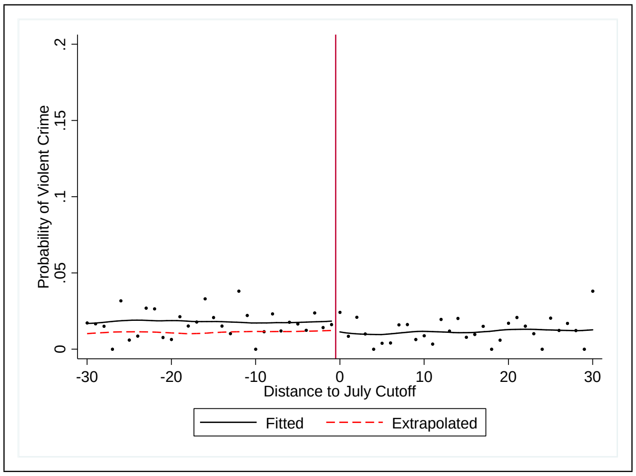
Panel D: Violent Adult/Juv Crime Convictions



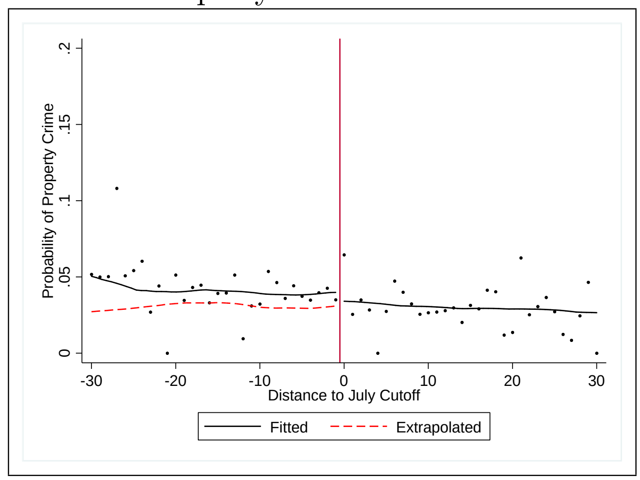
Authors' estimation of equation (4) as described in the text. The sample is restricted to students who scored within 30 points of one of the cutoffs on the July promotion exams. All specifications control for linear splines in index score as well as test year fixed effects. Covariates include indicators for gender, race, free/reduced lunch and immigrant status. Each point in each panel comes from a separate regression, using samples that increase in age moving rightward along the x-axis. The dependent variable is an indicator equal to 1 if the individual was first convicted of a crime by the given age. Each dot represents the regression discontinuity coefficient, with the bounds of the 95% confidence interval shown by the bars extending from each point. The confidence intervals are calculated from standard errors that are clustered at the school level.

Figure A-4: Linear Reweighting Estimates of the Counterfactual-Treatment Effect on the Treated (Extensive Margin)

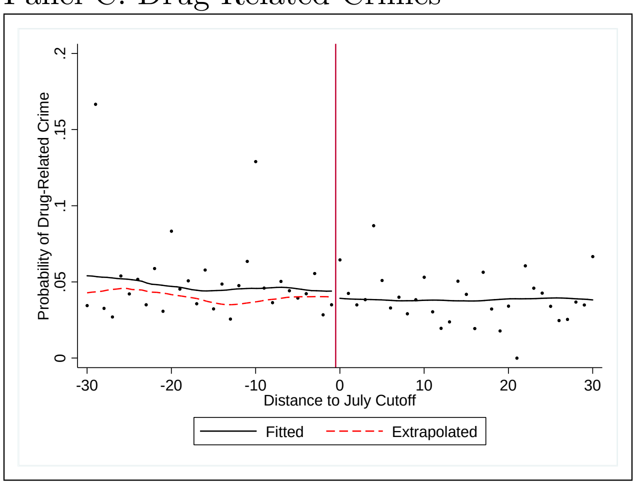
Panel A: Violent Crimes



Panel B: Property Crimes



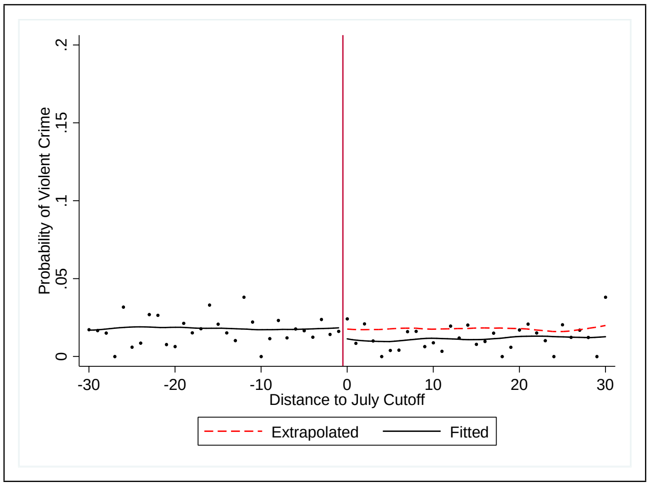
Panel C: Drug-Related Crimes



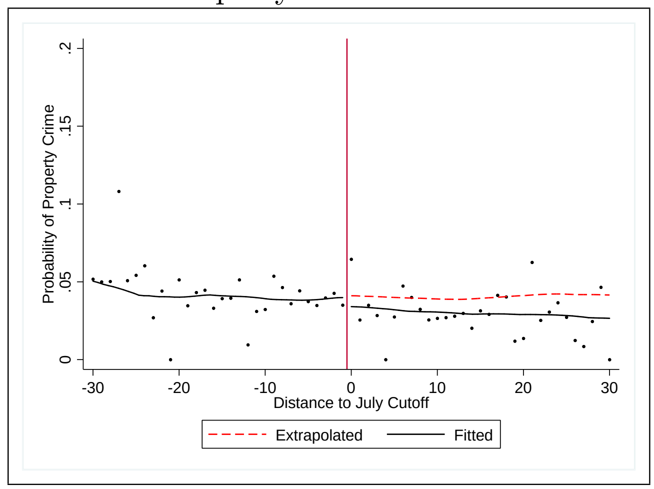
The Fitted line is the local linear estimate based on the actual data points shown. The Extrapolated line is the estimated counterfactual using the Angrist and Rokkanen (2015) method described in the text. The difference between the Fitted and the Extrapolated is the estimated effect away from the cutoff.

Figure A-5: Linear Reweighting Estimates of the Counterfactual-Treatment Effect on the Untreated (Extensive Margin)

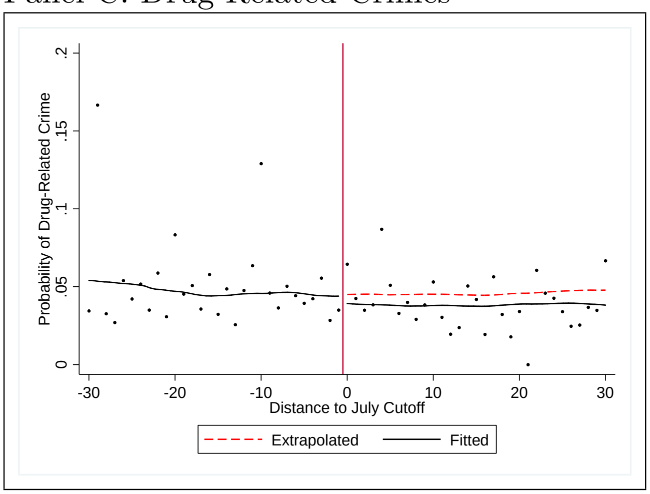
Panel A: Violent Crimes



Panel B: Property Crimes



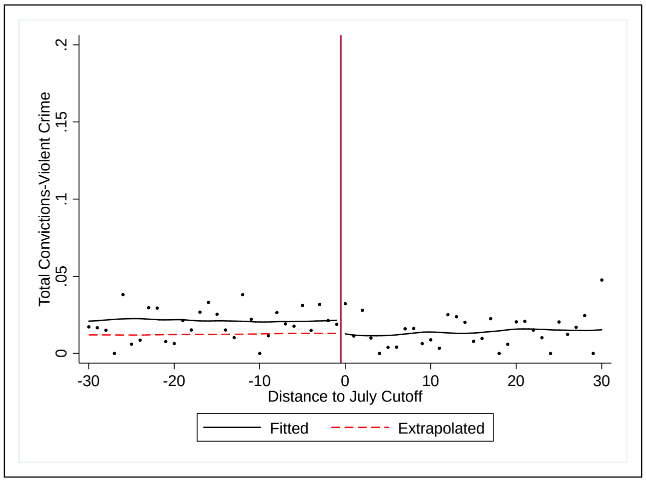
Panel C: Drug-Related Crimes



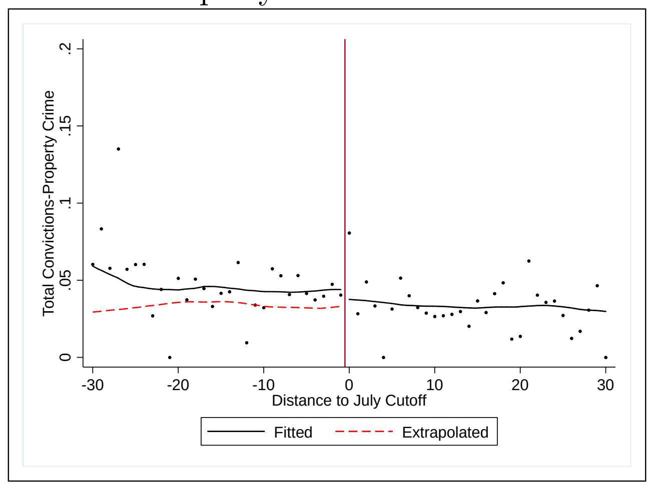
The Fitted line is the local linear estimate based on the actual data points shown. The Extrapolated line is the estimated counterfactual using the Angrist and Rokkanen (2015) method described in the text. The difference between the Fitted and the Extrapolated is the estimated effect away from the cutoff.

Figure A-6: Linear Reweighting Estimates of the Counterfactual-Treatment Effect on the Treated (Intensive Margin)

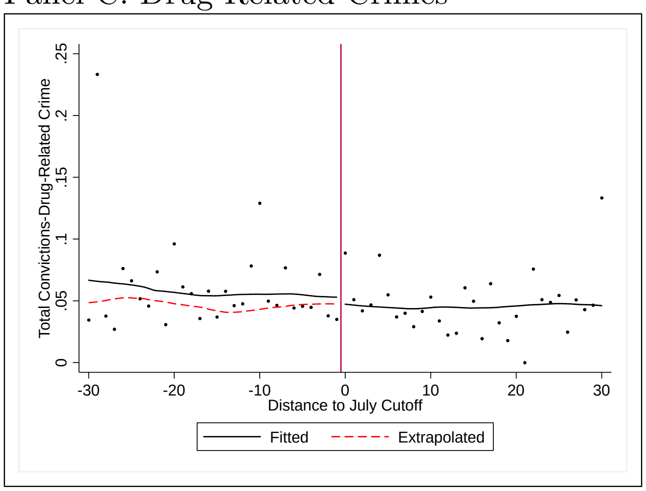
Panel A: Violent Crimes



Panel B: Property Crimes



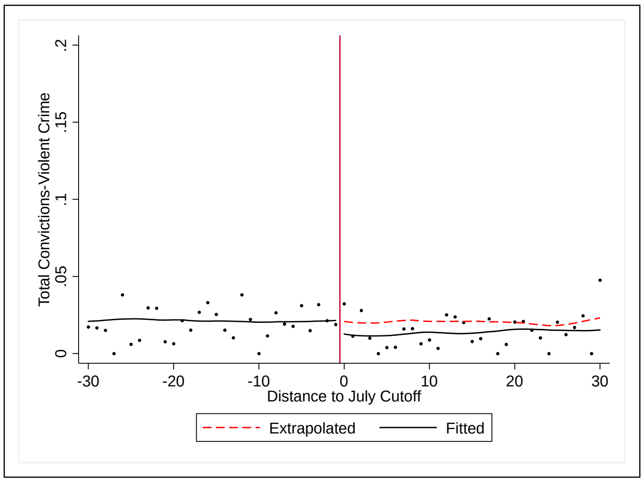
Panel C: Drug-Related Crimes



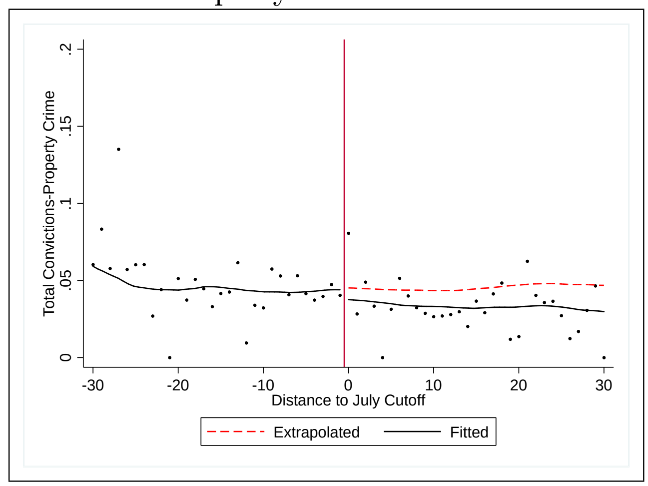
The Fitted line is the local linear estimate based on the actual data points shown. The Extrapolated line is the estimated counterfactual using the Angrist and Rokkanen (2015) method described in the text. The difference between the Fitted and the Extrapolated is the estimated effect away from the cutoff.

Figure A-7: Linear Reweighting Estimates of the Counterfactual-Treatment Effect on the Treated (Intensive Margin)

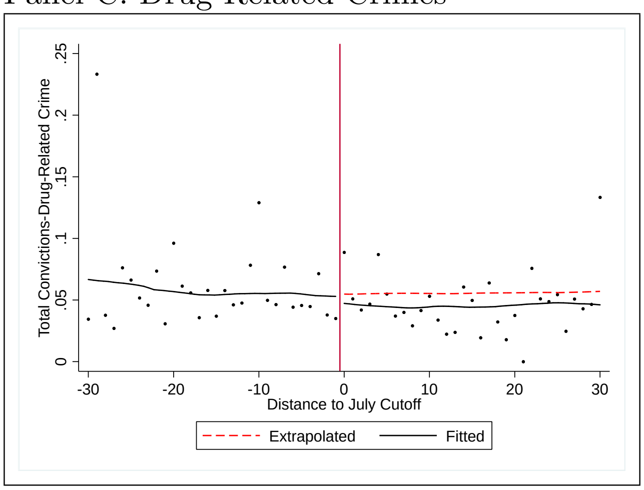
Panel A: Violent Crimes



Panel B: Property Crimes



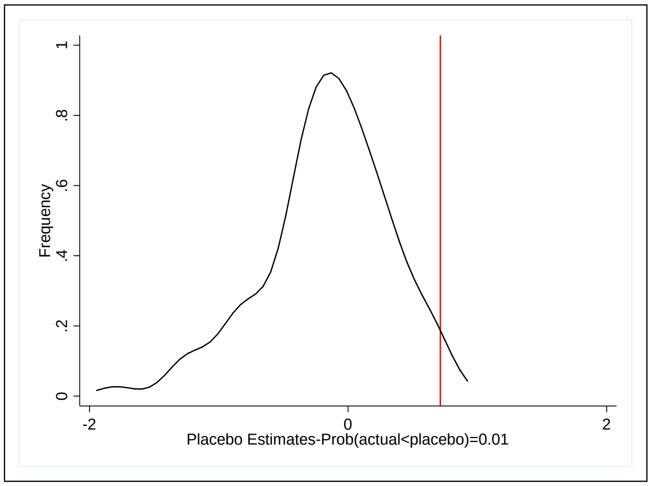
Panel C: Drug-Related Crimes



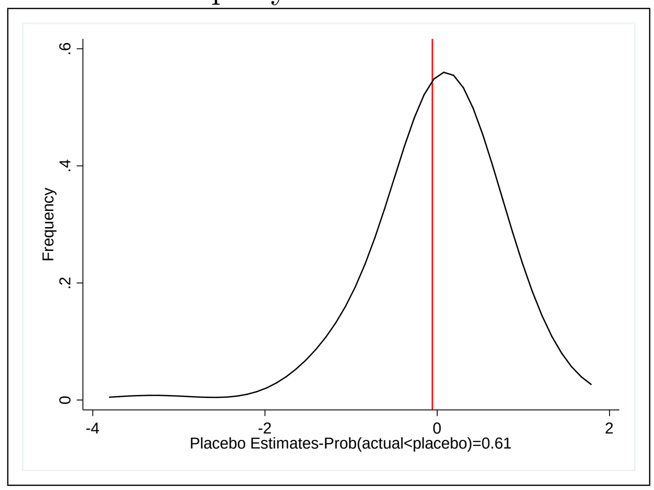
The Fitted line is the local linear estimate based on the actual data points shown. The Extrapolated line is the estimated counterfactual using the Angrist and Rokkanen (2015) method described in the text. The difference between the Fitted and the Extrapolated is the estimated effect away from the cutoff.

Figure A-8: Placebo Estimates of the Effect of Failing a July Promotional Cutoff on the Probability of Adult Criminal Conviction

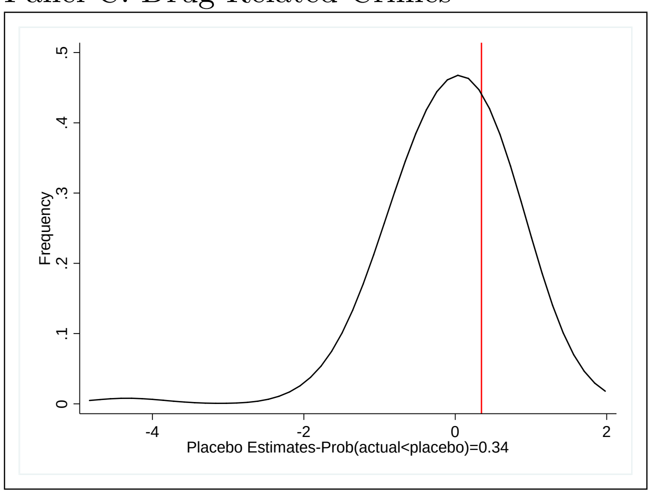
Panel A: Violent Crimes



Panel B: Property Crimes

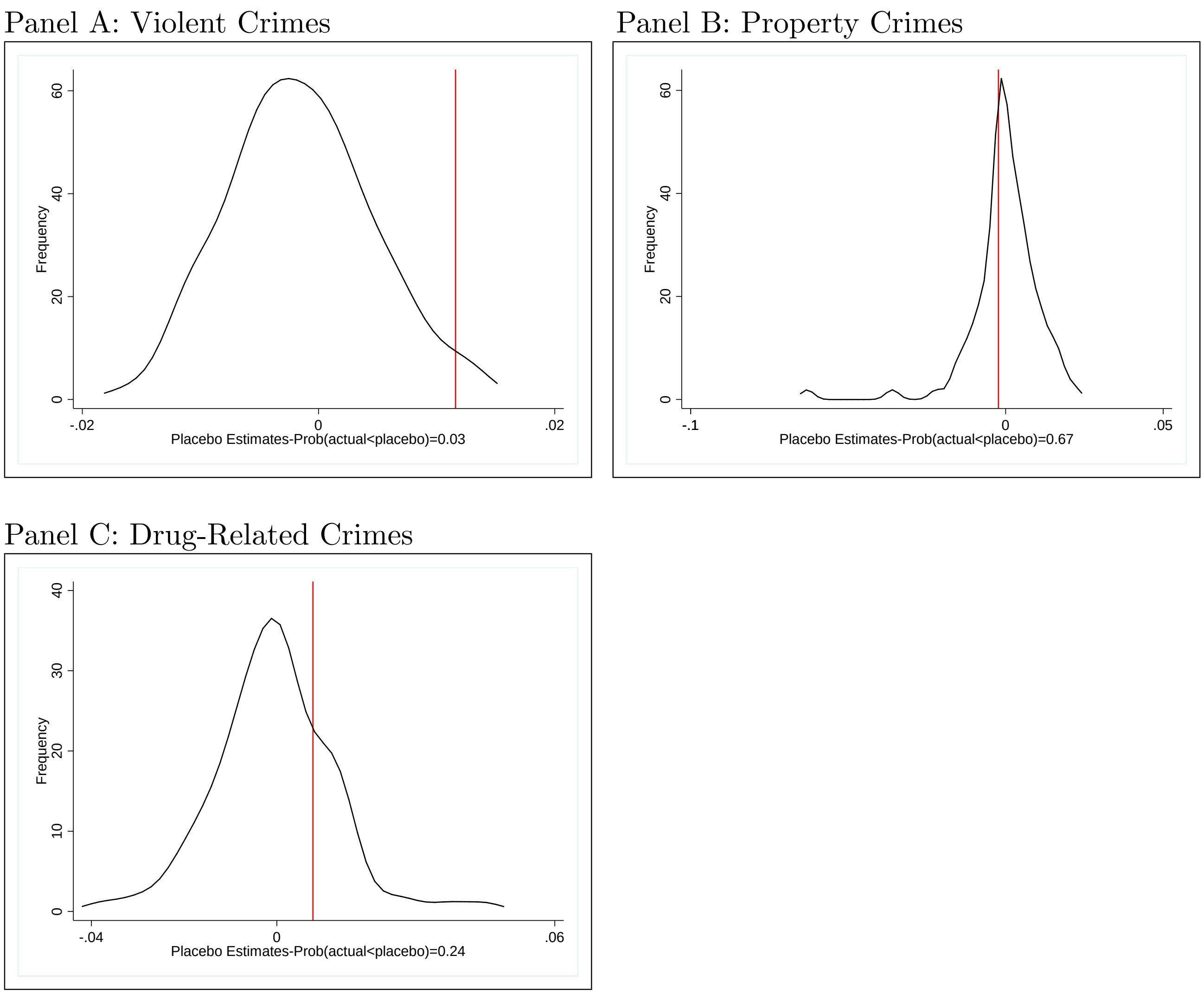


Panel C: Drug-Related Crimes



Each placebo estimate assigns a false promotion cutoff by incrementally adding one scale point to the subject-specific cutoffs (269 and 296 scale points in ELA and math LEAP tests, respectively) over a range of [-50, 50] scale points. A reduced form equation is then employed to estimate the effect of failing to meet the July cutoff on the probability of being convicted of different types of adult crime. All estimates are obtained from a linear spline using a bandwidth of 30 index points. The vertical lines denote the actual estimates. The fraction of placebo estimates larger than the actual estimate is also reported on the x-axis of each graph.

Figure A-9: Placebo Estimates of the Effect of Failing a July Promotional Cutoff on the Number of Adult Criminal Convictions



Each placebo estimate assigns a false promotion cutoff by incrementally adding one scale point to the subject-specific cutoffs (269 and 296 scale points in ELA and math LEAP tests, respectively) over a range of [-50, 50] scale points. A reduced form equation is then employed to estimate the effect of failing to meet the July cutoff on the number of adult criminal convictions by type of crime. All estimates are obtained from a linear spline using a bandwidth of 30 index points. The vertical lines denote the actual estimates. The fraction of placebo estimates larger than the actual estimate is also reported on the x-axis of each graph.

Table A-1: Reduced Form and IV Estimates of Test Failure and Grade Retention on Adult Crime Conviction - No Covariates

	Reduced Form		IV	
Margin:	Extensive (x100)	Intensive	Extensive (x100)	Intensive
	(1)	(2)	(3)	(4)
Panel A: Any Crime				
Failed July Promotion Cutoff	0.687	0.015		
	(0.875)	(0.012)		
Grade Retention			1.015	0.022
			(1.290)	(0.018)
Panel B: Violent Crime				
Failed July Promotion Cutoff	0.668*	0.011**		
	(0.352)	(0.005)		
Grade Retention			0.987*	0.016**
			(0.518)	(0.007)
Panel C: Property Crime				
Failed July Promotion Cutoff	-0.072	-0.002		
	(0.569)	(0.007)		
Grade Retention			-0.107	-0.004
			(0.841)	(0.010)
Panel D: Convicted of a Drug-Related Crime)				
Failed July Promotion Cutoff	0.254	0.007		
	(0.598)	(0.008)		
Grade Retention			0.375	0.010
			(0.883)	(0.011)

Authors' estimation of equations (1) and (4) as described in the text. The sample is restricted to students who scored within 30 points of one of the cutoffs on the July promotion exams; N=14,728. All specifications control for linear splines in index score and test year fixed effects. The dependent variable in columns (1) and (3) are indicators taking the value of one if the student is convicted as an adult by age 25. The dependent variable in columns (2) and (4) are the number of crimes for which the individual is convicted at first conviction by age 25 (including zeros). Outcome variables are multiplied by 100 in columns (1) and (3) to obtain percent values. Standard errors clustered at the school level are in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A-2: Robustness Checks of Bandwidth and Running Variable Controls - Regression Discontinuity Estimates of Grade Retention on the Likelihood of Being Convicted of a Crime

	Local Linear (Index=IK) (1)	Linear Spline Index=[-15,15] (2)	Linear Spline Index=[-45,45] (3)	Quadratic Spline Index=[-60,60] (4)	Quadratic Spline (Global) (5)
Panel A: Convicted of a Violent Crime (x100)					
Grade Retention	1.033*	0.807*	1.047**	1.033*	1.150***
	(0.582)	(0.430)	(0.438)	(0.607)	(0.419)
	[11,460]	[9,118]	[18,151]	[19,967]	[22,929]
Panel B: Convicted of a Property Crime (x100)					
Grade Retention	0.400	0.276	0.430	-0.151	0.330
	(1.025)	(1.105)	(0.690)	(0.987)	(0.695)
	[10,815]	[9,118]	[18,151]	[19,967]	[22,929]
Panel C: Convicted of a Drug-Related Crime (x100)					
Grade Retention	-0.194	-0.583	0.831	0.262	0.480
	(1.085)	(1.125)	(0.791)	(1.023)	(0.795)
	[9,741]	[9,118]	[18,151]	[19,967]	[22,929]

Authors' estimation as described in the text. All specifications in columns (2)-(5) control for linear or quadratic splines in index score using the listed bandwidth. Optimal bandwidths for local linear regression in the first column are obtained using the procedures in Imbens and Kalyanaraman (2012). All outcome variables are multiplied by 100 to obtain percent values. Covariates include indicators for gender, race, free/reduced lunch, immigration status, and test year fixed effects. Sample sizes underlying each specification are shown in brackets. Standard errors in parentheses are clustered at the school level: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A-3: Robustness Checks of Bandwidth and Running Variable Controls - Regression Discontinuity Estimates of Grade Retention on the Number of Adult Crime Convictions

	Local Linear (Index=IK) (1)	Linear Spline Index=[-15,15] (2)	Linear Spline Index=[-45,45] (3)	Quadratic Spline Index=[-60,60] (4)	Quadratic Spline (Global) (5)
Panel A: Violent Crime					
Grade Retention	0.015**	0.017*	0.013**	0.017**	0.015***
	(0.007)	(0.009)	(0.006)	(0.008)	(0.006)
	[14,493]	[9,118]	[18,151]	[19,967]	[22,929]
Panel B: Property Crime					
Grade Retention	0.001	-0.002	0.004	-0.008	0.000
	(0.013)	(0.014)	(0.008)	(0.012)	(0.009)
	[10,263]	[9,118]	[18,151]	[19,967]	[22,929]
Panel C: Drug-Related Crime					
Grade Retention	-0.009	-0.009	0.018*	0.007	0.013
	(0.014)	(0.014)	(0.010)	(0.013)	(0.010)
	[9,345]	[9,118]	[18,151]	[19,967]	[22,929]

Authors' estimation as described in the text. All specifications in columns (2)-(5) control for linear or quadratic splines in index score using the listed bandwidth. Optimal bandwidths for local linear regression in the first column are obtained using the procedures in Imbens and Kalyanaraman (2012). Covariates include indicators for gender, race, free/reduced lunch, immigration status, and test year fixed effects. Sample sizes underlying each specification are shown in brackets. Standard errors in parentheses are clustered at the school level: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A-4: **Robustness Checks of Sample Composition and Clustering - Regression Discontinuity Estimates of Grade Retention on the Likelihood of Being Convicted of a Crime**

	Conviction at Any Age Up to 28 (1)	Drop Hurricane Katrina/Rita Regions (2)	Add 7^{th} Grade Test Scores (3)	Add 8^{th} Grade March Test Scores (4)	8^{th} Grade FE (5)	Standard Errors Clustered at the Index Level (6)
Panel A: Convicted of a Violent Crime (x100)						
Grade Retention	1.178**	1.397**	1.080**	1.038**	0.959**	1.052***
	(0.528)	(0.612)	(0.526)	(0.515)	(0.379)	(0.404)
Panel B: Convicted of a Property Crime (x100)						
Grade Retention	-0.067	-0.499	0.058	-0.144	-0.037	-0.084
	(0.876)	(0.973)	(0.850)	(0.845)	(0.851)	(0.658)
Panel C: Convicted of a Drug-Related Crime (x100)						
Grade Retention	0.374	0.730	0.510	0.483	0.376	0.511
	(0.937)	(0.947)	(0.860)	(0.861)	(0.867)	(0.783)
Sample Size	14,728	10,199	14,378	14,728	14,728	14,728

Authors' estimation as described in the text. All specifications control for linear splines in index score with a bandwidth of 30. The dependent variable takes the value of one if a student is convicted as an adult by age 25 in columns (2)-(4). In column (1), the dependent variable takes the value of one if a student is convicted of a crime up to age 28. The analysis sample in column (2) excludes parishes that are known to be most affected from Hurricanes Katrina and Rita. Column (3) adds 7^{th} grade composite scores into the specifications, column (4) adds subject-specific 8^{th} grade test scores from the March exam, and column (5) adds in 8^{th} grade school fixed effects. Covariates include indicators for gender, race, free/reduced lunch, immigration status, and test year fixed effects. All outcome variables are multiplied by 100 to obtain percent values. Standard errors in parentheses are clustered at the school level in columns (1)-(5) and at the index (running variable) level in column (6): * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A-5: Robustness Checks of Sample Composition and Clustering - Regression Discontinuity Estimates of Grade Retention on the Number of Adult Crime Convictions

	Conviction at Any Age Up to 28 (1)	Drop Hurricane Katrina/Rita Regions (2)	Add 7^{th} Grade Test Scores (3)	Add 8^{th} Grade March Test Scores (4)	8^{th} Grade FE (5)	Standard Errors Clustered at the Index Level (6)
Panel A: Number of Violent Crime Convictions						
Grade Retention	0.018**	0.020**	0.018**	0.017**	0.017**	0.017***
	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)	(0.005)
Panel B: Number of Property Crime Convictions						
Grade Retention	-0.003	-0.009	-0.000	-0.004	-0.002	-0.003
	(0.010)	(0.011)	(0.010)	(0.010)	(0.010)	(0.008)
Panel C: Number of Drug-Related Crime Convictions						
Grade Retention	0.010	0.013	0.012	0.011	0.011	0.012
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.012)
Sample Size	14,728	10,199	14,378	14,728	14,728	14,728

Authors' estimation as described in the text. All specifications control for linear splines in index score with a bandwidth of 30. The dependent variable is the number of adult crime convictions by age 25 in columns (2)-(4). In column (1), the dependent variable is the number of adult convictions up to age 28. The analysis sample in column (2) excludes parishes that are known to be most affected from Hurricanes Katrina and Rita. Column (3) adds 7^{th} grade composite scores into the specifications, column (4) adds subject-specific 8^{th} grade test scores from the March exam, and column (5) adds 8^{th} school fixed effects. Covariates include indicators for gender, race, free/reduced lunch, immigration status, and test year fixed effects. Standard errors in parentheses are clustered at the school level in columns (1)-(5) and at the index (running variable) level in column (6): * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A-6: Regression Discontinuity Estimates of Grade Retention on Adult Crime Convictions, Including High School Fixed Effects

	Extensive Margin		Intensive Margin	
High School FE	No	Yes	No	Yes
	(1)	(2)	(3)	(4)
Panel A: Number of Violent Crime Convictions				
Grade Retention	1.026*	0.838	0.014*	0.012
	(0.575)	(0.593)	(0.007)	(0.008)
Panel B: Number of Property Crime Convictions				
Grade Retention	-0.399	-0.407	-0.004	-0.004
	(0.920)	(0.959)	(0.011)	(0.012)
Panel C: Number of Drug-Related Crime Convictions				
Grade Retention	1.763*	1.623*	0.028*	0.029*
	(0.946)	(0.920)	(0.012)	(0.012)
Sample Size	11,115	11,115	11,115	11,115

Authors' estimation as described in the text. All specifications control for linear splines in index score with a bandwidth of 30. The dependent variable is the number of adult crime convictions by age 25 in columns (1)-(2) and the number of convictions at first conviction by age 25 in columns (3)-(4). Estimates in even columns include high school fixed effects. Covariates include indicators for gender, race, free/reduced lunch, immigration status, and test year fixed effects. Standard errors in parentheses are clustered at the school level: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A-7: Regression Discontinuity Estimates of Grade Retention on Adult Crime Convictions Among Men

	Extensive Margin			Intensive Margin		
		ATT (away from (cutoff)	ATUT (away from cutoff)		ATT (away from cutoff)	ATUT (away from cutoff)
	RD			RD		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Violent Crime						
Grade Retention	2.072*	2.522***	1.753**	0.035**	0.033***	0.020**
	(1.106)	(0.684)	(0.759)	(0.015)	(0.008)	(0.009)
Dep. Var. Mean	0.030			0.036		
Panel B: Property Crime						
Grade Retention	-1.121	2.503***	1.532	-0.018	0.032***	0.022
	(1.167)	(0.903)	(1.112)	(0.021)	(0.010)	(0.014)
Dep. Var. Mean	0.060			0.067		
Panel C: Drug-Related Crime						
Grade Retention	0.046	1.521	2.028	0.029	0.031	0.035**
	(1.815)	(1.857)	(1.269)	(0.023)	(0.024)	(0.017)
Dep. Var. Mean	0.090			0.108		

Authors' estimation as described in the text. The sample is restricted to male students who scored within 30 points of one of the cutoffs on the July promotion exams; N=6,699. Results in columns (1) and (4) are IV estimates from equation (4) that use the July promotion cutoff as an instrument for retention. These specifications control for linear splines in index score with a bandwidth of 30. Results in other columns are extrapolated treatment effects away from the cutoff using the Angrist and Rokannen (2015) method. The ATT (average treatment effect on the treated) estimates show effects for those below the cutoff and the ATUT (average treatment effect among the untreated) estimates show effects for those above the cutoff. The dependent variable in columns (1)-(3) takes the value of one if a student is convicted as an adult by age 25. These estimates are multiplied by 100 to obtain percent values. The dependent variable in columns (4)-(6) is the number of adult crime convictions by age 25. Covariates include indicators for gender, race, free/reduced lunch, immigration status, and test year fixed effects. The estimates in columns (2), (3), (5), and (6) also include controls for composite test scores and school fixed effects. Standard errors are shown in parentheses. In columns (1) and (4) standard errors are clustered at the school level and in other columns they are computed using a nonparametric block bootstrap at the school level with 500 replications: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A-8: Regression Discontinuity Estimates of Grade Retention on Adult Crime Convictions Among Women

	Extensive Margin			Intensive Margin		
		ATT (away from (cutoff)	ATUT (away from cutoff)		ATT (away from cutoff)	ATUT (away from cutoff)
	RD			RD		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Violent Crime						
Grade Retention	0.271	0.027	0.268	0.003	-0.000	0.002
	(0.421)	(0.280)	(0.261)	(0.005)	(0.003)	(0.003)
Dep. Var. Mean	0.005			0.006		
Panel B: Property Crime						
Grade Retention	0.727	0.877	1.238**	0.008	0.008	0.012**
	(0.789)	(0.601)	(0.568)	(0.008)	(0.006)	(0.006)
Dep. Var. Mean	0.019			0.020		
Panel C: Drug-Related Crime						
Grade Retention	0.801	0.502	0.071	0.009	0.002	-0.003
	(0.725)	(0.405)	(0.394)	(0.010)	(0.006)	(0.005)
Dep. Var. Mean	0.013			0.016		

Authors' estimation as described in the text. The sample is restricted to female students who scored within 30 points of one of the cutoffs on the July promotion exams; N=8,029. Results in columns (1) and (4) are IV estimates from equation (4) that use the July promotion cutoff as an instrument for retention. These specifications control for linear splines in index score with a bandwidth of 30. Results in other columns are extrapolated treatment effects away from the cutoff using the Angrist and Rokannen (2015) method. The ATT (average treatment effect on the treated) estimates show effects for those below the cutoff and the ATUT (average treatment effect among the untreated) estimates show effects for those above the cutoff. The dependent variable in columns (1)-(3) takes the value of one if a student is convicted as an adult by age 25. These estimates are multiplied by 100 to obtain percent values. The dependent variable in columns (4)-(6) is the number of adult crime convictions by age 25. Covariates include indicators for gender, race, free/reduced lunch, immigration status, and test year fixed effects. The estimates in columns (2), (3), (5), and (6) also include controls for composite test scores and school fixed effects. Standard errors are shown in parentheses. In columns (1) and (4) standard errors are clustered at the school level and in other columns they are computed using a nonparametric block bootstrap at the school level with 500 replications: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A-9: Regression Discontinuity Estimates of The Effect of Failing the March Exam on Adult Criminal Convictions

	Dependent Variable:	
	I(Convicted of a Crime)x100	Number of Crimes Convicted
	(1)	(2)
Panel A: Violent Crime		
Failed March Promotion Cutoff	0.018	0.000
	(0.174)	(0.002)
Dep. Var. Mean	0.011	0.012
Panel B: Property Crime		
Failed March Promotion Cutoff	-0.041	0.002
	(0.264)	(0.003)
Dep. Var. Mean	0.026	0.029
Panel C: Drug-Related Crime (x100		
Failed March Promotion Cutoff	-0.036	-0.000
	(0.317)	(0.004)
Dep. Var. Mean	0.036	0.043
Sample Size	62,405	62,405

Authors' estimation of the effect of failing the March promotion cutoff as described in the text. All specifications control for linear splines in index score with a bandwidth of 30. The dependent variable takes the value of one if a student is convicted as an adult by age 25 in column (1) and is the number of crimes of which the individual is convicted at first conviction by age 25 in column (2). The estimates in column (1) are multiplied by 100 to obtain percent values. Covariates include indicators for gender, race, free/reduced lunch, immigration status, and test year fixed effects. Standard errors in parentheses are clustered at the school level: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A-10: The Effect of Grade Retention on Adult Criminal Convictions Using only the Math Score as the Running Variable

	Extensive Margin (x100) (1)	Intensive Margin (2)
Panel A: Violent Crime		
Grade Retention	1.121**	0.018**
	(0.426)	(0.006)
Panel B: Property Crime		
Grade Retention	-0.397	-0.007
	(0.888)	(0.101)
Panel C: Drug-Related Crime		
Grade Retention	0.063	0.004
	(0.879)	(0.013)
Sample Size	14,317	14,317

Authors' estimation as described in the text. All specifications control for linear splines in math score with a bandwidth of 30. The dependent variable takes the value of one if a student is convicted as an adult by age 25 in column (1) and is the number of crimes of which the individual is convicted at first conviction by age 25 in column (2). Covariates include indicators for gender, race, free/reduced lunch, immigration status, and test year fixed effects. Outcome variables are multiplied by 100 in column (1) to obtain percent values. Standard errors in parentheses are clustered at the school level: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A-11: Potential Channels - Regression Discontinuity Estimates of Grade Retention on Education and Crime Outcomes

	Classified as Special Ed in Any Year (Post-July Exam) (1)	Age at Conviction (2)	Sentence Length Imposed (# of Months) (3)	Highest Grade Attained (Post-July Exam) (4)
Grade Retention	0.233	0.262	3.042	-0.264***
	(1.651)	(0.267)	(8.492)	(0.055)
Mean of Outcome	20.3%	20.83	79.09	3.02
Sample Size	12,745	10,199	14,728	14,728

Authors' estimation of equation (4) described in the text. The sample is restricted to students who scored within 30 points of one of the cutoffs on the July promotion exams. All specifications control for linear splines in index score as well as test year fixed effects. Covariates include indicators for gender, race, free/reduced lunch, and immigration status. Standard errors clustered at the school level are in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A-12: Conditional Independence Tests

	Failed July Cutoff=1		Failed July Cutoff=0	
	Likelihood of Being Convicted	Number of Crime Convictions	Likelihood of Being Convicted	Number of Crime Convictions
	(1)	(2)	(3)	(4)
Panel A: Violent Crimes				
Running Variable (Index Score)	-0.016	0.0001	0.010	0.0001
	(0.020)	(0.0003)	(0.018)	(0.0002)
Panel B: Property Crimes				
Running Variable (Index Score)	-0.025	-0.0003	-0.027	-0.0004
	(0.028)	(0.0003)	(0.027)	(0.0003)
Panel C: Drug Related Crimes				
Running Variable (Index Score)	-0.021	-0.0004	-0.004	0.0003
	(0.028)	(0.0004)	(0.035)	(0.0005)
Sample Size	7,873	7,873	6,505	6,505

Authors' estimation as described in the text. All specifications control linearly for index score with a bandwidth of 30. Estimates use only observations to the left or right of the cutoff as indicated in column headings. Covariates include indicators for gender, race, free/reduced lunch, immigrant status, composite baseline test scores and school and test year fixed effects. Standard errors in parentheses are clustered at the school level: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A-13: The Effect of Retention on Adult Criminal Convictions Away from the Cutoff, Including Bootstrapped 95% Confidence Intervals

	Treatment Effect on the Treated (Extensive Margin) (1)	Treatment Effect on the Treated (Intensive Margin) (2)	Treatment Effect on the Untreated (Extensive Margin) (3)	Treatment Effect on the Untreated (Intensive Margin) (4)
Panel A: Violent Crime				
Grade Retention	1.070***	0.013***	0.976***	0.011***
	(0.283)	(0.003)	(0.300)	(0.004)
	[0.526,1.629]	[0.006,0.020]	[0.354,1.549]	[0.004,0.019]
Panel B: Property Crime				
Grade Retention	1.462***	0.017***	1.424***	0.017***
	(0.454)	(0.005)	(0.520)	(0.006)
	[0.504,2.290]	[0.007,0.026]	[0.500,2.492]	[0.006,0.029]
Panel C: Drug-Related Crime				
Grade Retention	0.962	0.015*	0.939*	0.014*
	(0.641)	(0.009)	(0.565)	(0.008)
	[-0.441,2.134]	[-0.004,0.030]	[-0.179,2.112]	[-0.003,0.030]
Sample Size	14,378	14,378	14,378	14,378

Authors' estimation as described in the text. The sample is restricted to students who scored within 30 points of one of the cutoffs on the July promotion exams. Average treatment effect estimates are obtained using a linear reweighting estimator. Covariates include indicators for gender, race, free/reduced lunch, immigrant status, composite test scores and school and test year fixed effects. Outcome variables in Columns (1) and (3) are multiplied by 100 to obtain percent values, while total number of adult convictions up to age 25 are used in Columns (2) and (4). Standard errors computed by taking the standard deviation of 500 nonparametric block bootstrap replications at the school level are in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%. Non-parametric 95% confidence intervals calculated using the 2.5^{th} and 97.5^{th} estimates from the block bootstrap replications are in square brackets.

Table A-14: The Effect of Retention on Adult Criminal Convictions Away from the Cutoff Among Males

	Treatment Effect on the Treated (Extensive Margin) (1)	Treatment Effect on the Treated (Intensive Margin) (2)	Treatment Effect on the Untreated (Extensive Margin) (3)	Treatment Effect on the Untreated (Intensive Margin) (4)
Panel A: Violent Crime				
Grade Retention	2.552***	0.033***	1.753**	0.020**
	(0.684)	(0.008)	(0.759)	(0.009)
	[1.077,3.789]	[0.017,0.048]	[0.327,3.236]	[0.004,0.038]
Panel B: Property Crime				
Grade Retention	2.503***	0.032***	1.532	0.022
	(0.903)	(0.010)	(1.112)	(0.014)
	[0.646,4.256]	[0.011,0.053]	[-0.479,3.522]	[-0.002,0.050]
Panel C: Drug-Related Crime				
Grade Retention	1.521	0.031	2.028	0.035**
	(1.857)	(0.024)	(1.269)	(0.017)
	[-3.536,3.751]	[-0.035,0.056]	[-0.616,4.459]	[0.005,0.067]
Sample Size	6,699	6,699	6,699	6,699

Authors' estimation as described in the text. The sample is restricted to students who scored within 30 points of one of the cutoffs on the July promotion exams. Average treatment effect estimates are obtained using a linear reweighting estimator. Covariates include indicators for gender, race, free/reduced lunch, immigrant status, composite test scores and school and test year fixed effects. Outcome variables in Columns (1) and (3) are multiplied by 100 to obtain percent values, while total number of adult convictions up to age 25 are used in Columns (2) and (4). Standard errors computed by taking the standard deviation of 500 nonparametric block bootstrap replications at the school level are in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%. Non-parametric 95% confidence intervals calculated using the 2.5^{th} and 97.5^{th} estimates from the block bootstrap replications are in square brackets.

Table A-15: The Effect of Retention on Adult Criminal Convictions Away from the Cutoff Among Females

	Treatment Effect on the Treated (Extensive Margin) (1)	Treatment Effect on the Treated (Intensive Margin) (2)	Treatment Effect on the Untreated (Extensive Margin) (3)	Treatment Effect on the Untreated (Intensive Margin) (4)
Panel A: Violent Crime				
Grade Retention	0.027	-0.000	0.268	0.002
	(0.280)	(0.003)	(0.261)	(0.003)
	[-0.545,0.574]	[-0.008,0.006]	[-0.212,0.797]	[-0.004,0.007]
Panel B: Property Crime				
Grade Retention	0.877	0.008	1.238**	0.012**
	(0.601)	(0.006)	(0.568)	(0.006)
	[-0.425,2.010]	[-0.006,0.020]	[0.150,2.433]	[0.001,0.025]
Panel C: Drug-Related Crime				
Grade Retention	0.502	0.002	0.071	-0.003
	(0.405)	(0.006)	(0.394)	(0.005)
	[-0.315,1.236]	[-0.011,0.013]	[-0.655,0.896]	[-0.013,0.008]
Sample Size	8,029	8,029	8,029	8,029

Authors' estimation as described in the text. The sample is restricted to students who scored within 30 points of one of the cutoffs on the July promotion exams. Average treatment effect estimates are obtained using a linear reweighting estimator. Covariates include indicators for gender, race, free/reduced lunch, immigrant status, composite test scores and school and test year fixed effects. Outcome variables in Columns (1) and (3) are multiplied by 100 to obtain percent values, while total number of adult convictions up to age 25 are used in Columns (2) and (4). Standard errors computed by taking the standard deviation of 500 nonparametric block bootstrap replications at the school level are in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%. Non-parametric 95% confidence intervals calculated using the 2.5^{th} and 97.5^{th} estimates from the block bootstrap replications are in square brackets.