THE EFFECT OF GRADE RETENTION ON ADULT CRIME:
EVIDENCE FROM A TEST-BASED PROMOTION POLICY

Ozkan Eren
Michael F. Lovenheim
Naci H. Mocan

The Effect of Grade Retention on Adult Crime: Evidence from a Test-Based Promotion Policy
Ozkan Eren, Michael F. Lovenheim, and Naci H. Mocan
NBER Working Paper No. 25384
December 2018
JEL No. I21,I26,J24

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 models use 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 declines in high school peer quality, lowered non-cognitive skill acquisition, and a reduction in educational attainment. However, we find little effect on juvenile crime, which suggests the effects on adult criminal engagement are driven by worse job market prospects and non-cognitive skills that stem from lower educational investments by students. 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 test- based promotion cutoffs lead to large 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.

Ozkan Eren
Department of Economics
University of California, Riverside
Sproul Hall
Riverside, CA 92521
ozkan.eren@ucr.edu

Michael F. Lovenheim
Department of Policy Analysis and Management
ILR School, and Department of Economics
Cornell University
264 Ives Hall
Ithaca, NY 14853
and NBER
mfl55@cornell.edu

Naci H. Mocan
Department of Economics
Louisiana State University
2439 BEC
Baton Rouge, LA 70803-6306
and NBER
mocan@lsu.edu

1 Introduction

Education policies alter the skills and knowledge of young adults, which influence not only individual earnings but also 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 such positive externalities. While a sizable literature exists that 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 existence and 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 9th grade in Louisiana on adult criminal convictions. 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 the ensuing incarceration have long-run negative effects on a number of outcomes ranging from employment opportunities to 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) as well as 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: those who obtain less schooling tend to have unobserved characteristics that also make them more likely to commit a crime. Existing work with credible identification strategies has investigated the impact of education, as measured by the number of years of schooling, on criminal propensity (e.g., Lochner and Moretti 2004; Machin 2011). In

---

[1] See Card (2001) and Oreopoulos (2006) 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 the estimates of Anderson (1999) imply that in 2018 the total cost of criminal activity, including lost productivity and diminished quality of life, is $1.7 trillion.

this paper, we instead estimate the effect on crime of a specific education policy that governs promotion to 9th grade. Specifically, we exploit a grade retention policy in Louisiana in a regression discontinuity framework to examine the extent to which having been held back in 8th grade impacts criminal delinquency up to age 25. Our analysis thus makes contributions both to the literature on the impact of human capital formation on crime and to the literature that investigates the consequences of test-based promotion policies.

The test-based accountability program we analyze was adopted in 1998 and requires 8th grade students to achieve a certain level of proficiency on both the state math and English exams in order to be promoted to 9th 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 8th grade.

Louisiana was an early of adopter of this test-based accountability program, but these programs have grown markedly in popularity 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 a type of accountability policy 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 the students to the next grade level. By ending "social promotion," the goal of these policies is to provide incentives for students to pass exams that policymakers hope will ultimately 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., Greene and Winters 2007; Jacob and Lefgren 2009; Manacorda 2012; Eren et al. 2017; Schwerdt et al. 2017). The effect of these policies on longer-run adult outcomes has remained unexamined prior to this analysis.

We employ two detailed administrative datasets from Louisiana to estimate the effect of being retained in 8th grade due to failing the summer promotion exam. Our data combine administrative records on all K-12 students in the state from the Louisiana Department of Education (LDOE) with administrative crime records on individuals up to age 25 from the Louisiana Department of Public Safety and Corrections, Adult Services. These data are unique

in allowing us to link together the universe of K-12 students in Louisiana with the universe of adult criminal convictions. Because of our focus on adult crime among those 25 years and younger, we focus on 8th graders from the 1998-1999 through 2000-2001 school years (the 1984-1987 birth cohorts). Our analysis is further restricted to those who fail one of the initial exams (in March of the relevant academic year), so the effects we estimate are relative to a low-performing set of students on the margin of failing to be promoted to 9th grade.

In our preferred IV model that uses the promotion cutoff as an instrument for grade retention, we find that being retained in 8th grade increases the likelihood of being convicted of any crime by a non-statistically-significant 1.25 percentage points. This is a 10.85% increase relative to the mean. The effect is driven by violent crime: 8th grade retention increases the likelihood of violent crime conviction as a young adult by a statistically significant 1.05 percentage points (58.44%). 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. Examining the number of convictions instead of just the probability of conviction shows that retention increases the number of criminal convictions by 0.025 (18.38%), which is composed primarily of a 0.017 (85.00%) violent crime increase that is significant at the 5 percent level.

The rich data on educational outcomes allows us to examine several potential mechanisms. We show that grade retention leads students to be exposed to lower-quality education environments, worsens non-cognitive outcomes, and lowers educational attainment. Specifically, failing the 8th grade promotion exam exposes students to education environments with a higher minority share and with lower-achieving peers. This occurs because students who fail the promotion cutoff subsequently attend high schools with lower-performing and more disadvantaged students. We also find that retention in 8th grade increases the number of days absent three years later by 1.97 (11.14% relative to the mean), increases the number of disciplinary incidents by at least 1.1 percentage points, and causes students to be 7.17 percentage points more likely to drop out of high school. We find some evidence that convictions for any juvenile crime increase by 0.53 of a percentage point and juvenile felony convictions increase by 0.34 of a percentage point, but neither estimate is statistically significant at conventional levels. Taken together, the evidence on mechanisms suggests that failing to be promoted to 9th grade on time

reduces education quality and human capital investments, 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. Thus, retaining students in 8th grade due to not passing this high stakes exam leads to substantially worse outcomes among a subset of these students into adulthood.

Finally, we show that the effects are largest for 20+ year olds and that the impact of grade retention on crime generalizes to students away from the test passing threshold. 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. The results indicate sizable positive effects of retention on crime for both groups. Indeed, we find positive and statistically significant effects away from the cutoff for violent, property and drug-related crime. These findings suggest that our regression discontinuity estimates understate the effects of grade retention on crime for a broader population of low-scoring students. They also indicate that one cannot simply move the passing threshold to eliminate the crime effect; the effect of grade retention on crime is evident at any reasonable retention cutoff in the test score distribution.

This paper is the first to estimate the causal effect of grade retention, due to test-based accountability policies, on adult criminal convictions. We thus make contributions to several strands of the literature. First, this paper adds to a growing body of research that examines the effect of grade retention on student outcomes. Most studies that credibly overcome selection concerns investigate test-based student accountability policies in a regression discontinuity framework as we do here. Jacob and Lefgren (2009) provide the earliest such analysis focused on the test-based accountability regime in Chicago. They find that test-based grade retention in the 6th grade has no effect on high school dropout but that 8th grade test-based retention significantly reduces the likelihood of high school completion. Florida's test-based retention program focusing on 3rd grade students has also been analyzed by several researchers. Students who barely fail the promotion cutoff perform better on standardized exams at first, but this effect fades out by high school. They also have higher GPAs, but the likelihood of graduating is unaffected (Schwert et al. 2017; Greene and Winters 2007). Retaining students in early grades

may have fundamentally different long-run effects than retaining them in higher grades. This may explain the differences in findings across these studies.[3]

Eren et al. (2017) make use of the same Louisiana policy and employ similar data as we do in this paper to estimate the effect of summer school on high school dropout. Their supplementary analysis of grade retention shows that grade retention in the 8th grade has no impact on juvenile crime convictions.[4] This finding, which is also confirmed by our paper, indicates that the effect of grade retention on adult crime we identify is not the results of a spillover effect; i.e. it is not due to the continuation of crime commission that started as a juvenile (before the age of majority). This aspect of our paper underscores the importance of directly examining longer-run outcomes, such as adult criminal convictions, using a design that can plausibly account for selection bias. To our knowledge, we are the first to be able to estimate this parameter.

Second, our analysis contributes to a small literature on the effect of education on crime.[5] Lochner and Moretti (2004) use state changes in compulsory schooling laws to estimate the effect of educational attainment on both arrests and incarceration.[6] They find that those who complete high school because of compulsory schooling are much less likely to be involved with the criminal justice system. Similarly, Machin et al. (2011) use the increase in the minimum schooling age in England and Wales to identify an effect of high school educational attainment on crime. They find that increased high school attainment reduces both property and violent crime convictions later in life. While similar in many ways to the parameter we estimate, students who do not complete high school because of compulsory schooling laws likely differ from those failing the promotion exam in important dimensions related to cognitive and non-cognitive skills. Additionally, grade retention can impact crime through mechanisms other than altering the number of years of schooling, as we show. This underscores the importance of providing direct evidence on how test-based promotion policies impact adult crime.[7]

Further evidence on the effect of education on crime comes from the school choice literature: Deming (2011) shows that an increase in school quality due to winning a school choice lottery in

---

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

[4]Other work has found evidence of a link between high school attainment and juvenile crime: Anderson (2014) shows that increasing the high school dropout age leads to a reduction in juvenile arrest rates.

[5]See Lochner (2011) for a more detailed review of the literature on the link between schooling and crime.

[6]Oreopoulos and Salvanes (2011) reproduce these results using an expanded sample and a slightly altered specification.

[7]Stephens and Yang (2014) cast doubt on the validity of the difference-in-difference approach surrounding changes in compulsory schooling laws based on evidence that states that change these policies experience different secular trends from those that do not. However, they lack the data to explicitly examine the common trends assumption with respect to crime outcomes.

Charlotte-Mecklenberg leads to lower levels of crime 7 years after assignment. These estimates are identified off of clearly exogenous lottery-based variation, but the affected students tend to be higher-achieving than those who are on the margin of being retained in our setting.[8] Thus, the school choice results do not necessarily provide insights into potential effects of test-based grade retention on adult criminal activity.

Our main findings indicate that grade retention increases the likelihood of being convicted of a violent crime and the number of violent crimes of which individuals are convicted. Thus, test-based promotion causes harm to a subset of students, and the effect persists into adulthood. Because violent crimes have large externalities associated with them, this effect reduces or may reverse any social benefits to these policies. We calculate that increase in violent crimes, driven by grade retention, implies a social cost between $2.6 to $18.4 million. Whether test-based promotion has other offsetting benefits that may make it cost-beneficial is beyond the scope of our analysis, but the size of the crime effect is such that any 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 do not address whether there is an optimal cutoff that balances any positive effects of grade promotion with the effects on crime and high school dropout. This is an important public policy question highlighted by our findings. Finally, our results point to specific types of students who are being harmed by this policy. These students could be targeted with additional resources, which could perhaps mitigate the effects we identify.

The rest of this paper is organized as follows: Section 2 presents the policy background, Section 3 describes the data, and Section 4 shows the estimating equation and discusses identification. Results are shown in Section 5, and Section 6 concludes.

---

[8]Cullen et al. (2006) show similar estimates among high school students who win a school choice lottery in Chicago. They use self-reported crime data, however, and do not have information on criminal activity after high school.

2 Policy Background

The Louisiana School and District Accountability system was adopted by the state's Board of Elementary and Secondary Education (BESE) in June 1998. The adoption of this policy was four years before the federal No Child Left Behind Act, which underscores that Louisiana was an early mover in terms of adopting accountability policies. A core component of the policy focused on school accountability, with the state setting ten and twenty year goals for all public schools and requiring schools to demonstrate progress toward these goals. As part of the overall accountability system, the BESE also included a student-focused accountability provision that was designed to end "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.[9]

LEAP tests are criterion-referenced tests and are designed to directly align with the state content standards. A student's LEAP score 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 LEAP tests, respectively.[10] 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.[11] 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. The school districts are given the flexibility to determine curriculum used in summer remediation classes, but the summer programs are monitored by the state on a regular basis. Evidence from the annual summer school remediation reports and monitoring

---

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

[10]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.

[11]Beginning in Spring 2006, the state replaced ITBS with *integrated* Louisiana Educational Assessment Program (*i*LEAP), a low-stakes test consisting of both criterion-referenced and norm-referenced components.

visits suggest that teachers in summer schools are proficient in the content area in which they are teaching and that they are using a variety of creative teaching strategies including but not limited to small group instruction, use of hands-on materials, and problem solving teams. However, there are large classrooms in some summer programs (around 20 students per classroom), which has raised concerns about their effectiveness (Pastorek 2010).

Students are not required to attend summer school programs to be eligible for the July testing. Those who pass the July exams move on to the next grade. Students who again fail are required to repeat the grade unless they are permitted to move to the next grade through 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.[12] 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

One of the main innovations of this study is to merge together administrative education records with administrative criminal records that allow us to observe rich information about the entire population of Louisiana students and those convicted of a crime. The data for this study come from two different sources. The first is administrative records from the Louisiana Department of Education (LDOE). These administrative data include student demographic information such as gender, race and free/reduced lunch status, as well as scores from the LEAP and ITBS tests in all years in which these tests were taken. Unique state identification numbers allow us to track all the students through their tenure in the public school system and therefore, we are able to identify in which school each student was enrolled. The data also include whether a student has dropped out without receiving a diploma, the number of days absent in each school

---

[12]Unfortunately, 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.

year, and the number of disciplinary incidents. Note that 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 option, but as we show below failure is uncorrelated with the likelihood of attriting from the public education sample. Furthermore, only leaving the state causes attrition in our crime estimates, as private school and home schooled students are included in the conviction data.

The second source of data is the Louisiana Department of Public Safety and Corrections, Adult Services. These data span from 1996 through 2012 and are merged with the LDOE data using common state identification numbers from both datasets. The crime data include basic demographic information, the exact type of crime committed and sentence type (incarceration or probation) for each conviction. Importantly, only those convicted of a crime are in this dataset; those who are arrested but never charged or found not guilty are not included. 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. Thus, serving prison time is common for violent crime convictions. 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 at age 25 or earlier. We use this age cutoff to be consistent with the literature (Aizer and Doyle 2015), because most crime is committed by those 25 and younger (Hirschi and Gottfredson 1983), and because of the timing of our education data. 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, roughly corresponding to the cohorts born between 1984 and 1987. Our analysis sample also is composed solely of students who fail to meet the March promotional cutoff and took the July exams. This is important because summer school can have independent effects on student outcomes (Matsudaira 2008). That both retained and promoted students had equal access to summer school means our estimates are net of any summer school effect. Our final analysis sample consists of 22,929 unique student observations. The total number of students in this sample roughly corresponds

to 15% of the universe of eighth graders from the 1998-1999 to the 2000-2001 academic years.

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 are focused on low-performing students, the sample disproportionately consists of disadvantaged students and average test scores are lower than the state average. 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. Disproportionate representation of black students and low test scores are even more pronounced for the sample of retained students, as shown in Column (5). The math exam is more of an obstacle for promotion than the ELA test, as only 5% of the retained sample is retained solely because of ELA failure. Note that 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 8th Grade LEAP test outcomes only sum to 95% in Column (5). Furthermore, about 35% of those promoted also failed 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 the summer school group. Among those retained, the dropout rate is even higher, at 53.5%. The much higher dropout rate among those who are retained likely reflects the lower academic achievement level of this group 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-

_____

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

related crime, and 1 percent of another type of crime.[14] Note that 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. 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. Interestingly, while those who are promoted are less likely to be convicted of a crime by age 25, the differences across the groups is small as are the distributions of crime types.

4 Empirical Methodology

4.1 Reduced Form and IV Models

The goal of this analysis is to estimate the effect of grade retention on crime during young adulthood. This parameter is difficult to identify, because students who are retained have lower academic achievement, are more likely to be minorities, and tend to come from lower-income households.[15] These characteristics also are correlated with the likelihood of having a criminal conviction as an adult. Even a rich conditioning set is unlikely to fully control for selection into retention, since students who are retained probably have unobserved skills and traits (e.g., patience) that make them more likely to commit a crime. This makes the assumptions underlying selection-on-observables models very strong.

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 local randomness to test scores, all observed and unobserved characteristics of students should move smoothly through the July test score cutoff. We estimate reduced form models that show the effect of barely failing one of the July exams as well as two stage least squares models in which we use the discontinuity created by the accountability policy as an instrument for grade retention among students who

---

[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.

[15]In our setting, minority students are either black or Hispanic. However, because of the demographic makeup of Louisiana, virtually all minority students in our sample are black.

failed the March exam. Note that 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.[16] 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 (until age 25). This can be greater than one because people are often 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 LEAP exams. That is, FP_i takes the value of one if the minimum of the difference between subject-specific July LEAP scores and their respective relevant cutoffs is negative. $Index_i$ 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 $Index_i$, as suggested by Gelman and Imbens (2017). We show as well 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 observed covariates that include indicators for gender, race, free/reduced lunch and immigrant status as well as year fixed effects. In some specifications, we also control for 8th grade school fixed effects and for

---

[16]Unfortunately, we do not have data on which students took up summer school to directly test this assertion because the administrative data do not include any information on summer school attendance. However, aggregate 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 for both subjects. While summer school could increase the likelihood of passing the test, it is doubtful that summer school participation leads to heaping right over the passing threshold. As long as summer school participation moves smoothly through the July cutoff, summer school is not a confounding factor in this analysis.

composite test scores in ELA and math from the 7th grade ITBS exams. 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. The key identifying assumption underlying the identification of γ_1 in equation (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.

Since we observe the actual retention status of the students in the sample, we also estimate a fuzzy regression discontinuity (RD) model. Using the threshold indicator as an instrument for actual grade retention yields the local average treatment effect (LATE), which is the impact for those students whose retention decisions are causally affected by the state's promotion policy. Thus, this parameter identifies the effect of grade retention among those who were retained because they failed a July exam and complied with the decision. This LATE 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 we identify are for an extremely relevant group, as they are the ones impacted by the test-based promotion policy. Our estimates therefore are more likely to generalize to other states and school districts that have similar policies for 8th grade promotion. The two stage least squares models is as follows:

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

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

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 8th grade school.[17] The baseline bandwidth we use is 30 index points on either side of the cutoff. This bandwidth corresponds to 64 percent of our initial sample. We show that our results are robust to using different bandwidths as well.

The variable of interest in equation (4) is β_1, which shows the effect of being retained in

---

[17]Our estimated standard errors are similar if we instead cluster at the level of the running variable (see Online Appendix Tables A-3 and A-4). 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. Furthermore, students in the same school experience similar education inputs and tend to be similar on observed and unobserved dimensions, which makes the 8th grade school a sensible variable on which to cluster.

8th grade on adult criminal convictions. The assumptions underlying the identification of this causal parameter are very similar to those discussed above for γ_1 in the reduced form model. The additional assumption invoked here is that the *entire* effect of failing a July exam on adult crime comes through the effect of failing on retention. If there are independent effects of failing to meet the threshold, 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 in investing in education that respond to 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, however, and we focus on them in the results section.

The estimates obtained from the RD specification identify effects of retention only among those who are local to the pass-fail threshold. If there are heterogeneous treatment effects by student ability, such estimates may not be representative of the effect of retention for larger groups. 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 an RD approach in this context is that students are unable to sort around the cutoff. Given that the running variable consists of test scores, 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 due to test failure.[18] The next five columns show how a series of immutable or pre-exam characteristics vary as a function of the promotion cutoff: female, black, white, free/reduced priced lunch and 7th grade exam scores. All of the estimates are small in absolute value, and none is statistically significant at even the 10% level. Thus, there is no evidence in Table 2 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. Indeed, the mass right over the cutoff is lower than the mass right above, but this difference is small and clearly reflects the choppiness of the underlying test score distribution rather than a behavioral response to the cutoff. Taken together, the estimates in Table 2 and Figure 1 support our regression discontinuity approach by showing no evidence of student sorting relative to the promotion cutoffs.

---

[18]Note that transferring to a private school does not generate attrition in our criminal conviction data. Thus, the only relevant margin of attrition is out-of-state migration. 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.

5 Results

5.1 Main Results

First-stage estimates that show how July test failure relates to grade retention are shown 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 8th grade school fixed effects in the model.

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 implied first-stage F-statistics are quite large. Figure 2 shows 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 pass the test are retained for other reasons, and some students who fail wind up being promoted after they appeal. This is a fuzzy regression discontinuity setting, and the July test score cutoff for promotion is a very strong instrument for grade retention.

Reduced form and second stage estimates are presented in Table 4, using an indicator for being convicted of a crime by age 25 as the dependent variable. The first three columns present reduced form effects (equation 1) while the last three show IV estimates of the effect of grade retention (equation 4). Each set of three results columns is laid out identically to those in Table 3. We examine whether individuals are convicted of any crime and whether they are convicted of a violent crime, a property crime or a drug-related crime. It is possible to be convicted of multiple types of crime, so these categories are not mutually exclusive. Note that dependent variables have been multiplied by 100 for ease of interpretation. The estimates in columns (2) and (5) that include test year fixed effects and observed student characteristics are our preferred estimates, and we focus on them throughout. Our estimates are robust to excluding these covariates and to including school fixed effects; the results change little across columns.

The impact of failing a July exam on any conviction in Panel A is positive but not significant at conventional levels. In column (2), 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 resulting IV estimates also are positive, but similar to the reduced form results, they are not significantly different from zero. Still, 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 a violent crime conviction as an adult by 0.71 of a percentage point, which is 39.61% relative to the baseline likelihood of such a conviction. 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 very few murder convictions in this sample. There is no evidence of an effect on the likelihood of being convicted of a property crime: estimates are negative, close to zero, and are not statistically significant at even the 10% level. Estimates for drug-related crimes are positive and non-trivial in size, though they are imprecisely estimated and 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 (6.78%) 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. The figure aligns well with the results in Table 4: there is an increase in violent crime at the cutoff, but the other crime categories show less evidence of an effect.

Examining an indicator for whether any conviction occurs may miss variation coming from the number of convictions. 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, which are both violent crimes. One may reasonably interpret such convictions as representing more serious crimes than do convictions for one offense. In Table 5, we therefore 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.[19]

---

[19]We also could examine recidivism using the total number of times an individual was convicted of a crime up to age 25. However, examining re-offending with a sample that imposes an age cutoff of 25 years old will not allow us to measure the full effect of grade retention on re-offending. This follows from the fact that a subset of convicted individuals would still be in prison by the time they reach 25, and the likelihood of incarceration is highest among those who are most likely to re-offend. Examining recidivism in our sample thus may provide misleading estimates.

This intensive margin arguably better measures the severity of the crime than does a conviction indicator, however the results are quantitatively and qualitatively similar to those in Table 4. In column (2), we find that failing a July exam increases the number of crimes committed by 0.017, which is a 12.50% rise relative to the mean. The estimate is not statistically significant, though. Being retained increases the number of adult convictions by 0.025, or 18.38%, which also is not significant at even the 10% level.

In Table 5 we again find 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 0.017, or 85.00%. We also find suggestive evidence of a sizable 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 5 closely in showing an increase in the number of criminal convictions, violent crime convictions, and drug convictions due to failing a July exam but no effect on property crime convictions.

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. We also find suggestive evidence of a positive effect on drug convictions. It is important to emphasize the large private costs and externalities associated with these outcomes. The outcomes we examine are convictions, not arrests, which means that the increased criminal activity is both likely to be severe and leads to a permanent criminal record.[20]

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 (5) of Table 4. Note that retention increases the probability of being convicted of a violent crime by around 58.44% (1.052/1.800). Taking the number of convicted adults from violent crimes in the control group as our benchmark,

---

[20]Online Appendix Table A-6 shows suggestive evidence that retention increases sentence length by 3 months (off of a mean of 79 months) among those convicted, however the estimate is imprecise. This is consistent with the severity of crimes committed increasing. The average age of first conviction also rises slightly, as shown in column (2) of the table, but the estimate is again imprecise and thus only suggestive of an effect.

the percent increase 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 than what is reported here (FBI 2012).

5.2 Mechanisms

There are several potential mechanisms that could generate the impact of grade retention on adult crime. Holding students back a year may affect their accumulation of non-cognitive skills. Indeed, a core argument for "social promotion" is that it is damaging for students to be in grades with those who are younger. 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 for some students. Retaining students also can alter the educational environment, in particular the set of peers to which a student is exposed. 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. Lower levels of human capital and educational attainment can lead to higher rates of criminal activity by worsening job prospects. Reductions in non-cognitive skills can increase the likelihood young adults engage in criminal behavior as well. To the extent that these mechanisms lead students to commit criminal acts as juveniles, adult crime could increase because juvenile incarceration has a strong positive effect on adult criminal engagement (Aizer and Doyle 2015).

Tables 6-8 provide evidence on some of these potential mechanisms. First, in Table 6, we examine how grade retention affects the composition of students in the high school the student attends. While retention does lead to an increase in the likelihood that 9th grade data are missing,[21] we observe a reduction in measures of peer quality. Being retained in 8th grade leads

---

[21]This differential attrition occurs predominantly because we only observe 9th grade outcomes if students enroll in a public high school in Louisiana. Retention reduces the likelihood that students enroll in high school, which is why the attrition rate in Table 6 is so much larger than in the subsequent tables that do not condition on enrolling in a public high school. Hence, the attrition in Table 6 provides evidence of reduced educational attainment from retention, which complements the attainment evidence in Tables 8 and A-6.

students to attend high schools in which the percentage of white students in 9th grade is 2.7 percentage points lower (significant at the 12% level)[22] and the mean 9th grade test score is 6.2% of a standard deviation lower (significant at the 5% level). This finding suggests that retained students subsequently attend high schools that have peer attributes associated with worse academic outcomes. The change in peer composition we document likely is facilitated by the large amount of school choice that exists in Louisiana, and resulting adverse peer effects could be a driver of the increases in crime we document.

Changes to the peer composition of one's high school as well as direct effects of retention can alter educational outcomes. The development of non-cognitive skills is particularly important to examine, since these skills are highly correlated with the propensity to engage in criminal activity (Carneiro et al. 2007). We now turn to an analysis of behavioral outcomes in Table 7 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 for students. 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. If anything, this likely 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 7, 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 and are actually less likely to have a behavior incident in the first year following failure, although neither estimate is statistically significant. By the third year after retention, however, 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, though 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.

In addition to reductions in non-cognitive skill acquisition, grade retention can lead to less education attainment, as shown in Table 8. The first column of Table 8 indicates that students

---

[22]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.

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.[23] Column (2) of Table 8 shows that students who are retained in 8th grade due to failing a July exam 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%.[24] Despite the large impact of retention on high school dropout, we provide suggestive evidence that this effect can explain only a small proportion of the violent crime increase. In this exercise, we estimate 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. Thus, increased prevalence of high school dropout explains only 12% of the total increase in violent crime stemming from 8th grade retention.

The last two columns of Table 8 present estimates of grade retention on juvenile crime and juvenile felony convictions. These estimates are positive but not significant at even the 10% level. Furthermore, they are much smaller than the IV estimates in Table 4, suggesting that increases in juvenile crime can at most explain a small part of the overall adult crime effect.[25] Given the evidence that increased adult crime from 8th grade retention is not predominantly driven by the influence of high school dropout or juvenile convictions, we argue that the effect most likely comes from a degradation of non-cognitive skills.

Taken together, the results from Tables 6-8 indicate that grade retention leads to lower peer quality in high school, increases behavioral problems and absenteeism several years after being retained, and substantially increases the likelihood of dropout. There is little effect on juvenile crime, and back-of-the-envelope calculations indicate that an increase in the propensity to drop

---

[23]The LDOE analyzes enrollment data employing an algorithm coupled with U.S. Department of Education guidelines and flags an individual as being a drop out using these guidelines. Specifically, 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.

[24]Online Appendix Table A-6 shows that retention has a positive but not statistically significant effect on being classified as special education and reduces the number of years of completed education post-July exam by 0.264 years. The latter estimate is significant at the 1% level.

[25]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.

out, triggered by grade retention, does not explain the crime effect in adulthood. 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 accumulating less non-cognitive skills, which both lowers their job market prospects and makes it more likely they will engage in violent criminal behavior.[26] Indeed, our results are consistent with the decline in violent crime from a summer jobs program (Heller 2014), which also was likely driven by changes in non-cognitive skills.[27]

5.3   Heterogeneous Treatment Effects

While we do not find much evidence of an effect on juvenile crime, it still is possible that the impact of retention on adult crime has an age pattern. This is particularly relevant because the likelihood of committing a crime declines with age after the late teens (Hirschi and Gottfredson 1983). In Table 9, we examine age heterogeneity by redefining convictions as occurring only at ages 19+ (column (1)) and 20+ (column (2)). The effect of 8th grade retention on adult violent crime largely carries over to considering convictions that only occur among those who are between 20 and 25. This finding indicates that much of the adult crime effect occurs among young adults 20 and older. As with the baseline estimates, there is no significant effect for either age group for the other crime categories.

The regression discontinuity estimator identifies the effect of retention at the promotion cutoff. This local average treatment effect has clear policy relevance for those proximate to the threshold, but from a policy perspective it is important to understand whether the effects we find 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.[28] If students below the threshold who are retained also engage in more crime later in life,

---

[26]One might expect property crime to increase if changing labor market prospects are an important mechanism underlying our result. However, there is some evidence that youth employment programs lead to more property crime among those likely to be employed absent the intervention (Davis and Heller 2017). It hence is not obvious that reduced labor market opportunities will increase property crime.

[27]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 do not find evidence of an effect on traditional non-cognitive skills but rather that the interventions caused students to "slow down" and make more measured decisions.

[28]We consider here the set of test-based promotion policies targeted towards lower-performing students. No promotion policy will

the size of the externality implied by our estimates is much larger than if the effect is localized to the test score cutoff. In order to examine the extent to which our regression discontinuity estimates generalize to a broader set of lower-performing student, we use the method proposed by Angrist and Rokkanen (2015) to estimate treatment effects away from a 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. Angrist and Rokkanen (2015) suggest a simple test for the conditional independence assumption that entails 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-5: the estimates are universally small, and none is statistically significantly different from zero at conventional levels. Thus, the tests strongly support the use of the linear reweighting method to estimate treatment effects away from the test score discontinuity.

Table 10 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 in the table are block bootstrapped at the school level using 500 replications. The table shows consistent evidence of positive effects of crime 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 Tables 4 and 5 suggests that the effect of retention on violent crime is more general among low-performing students. Thus, any

---

set thresholds targeted at the upper end of the test score distribution, so the only feasible set of cutoffs available to policymakers is within the 30 point bandwidth we examine.

reasonable test-based cutoff is likely to generate an increase in violent crime later in life.

For property and drug crimes in Panels B and C of Table 10, we find effects that are larger than the RD estimates in Tables 4 and 5. 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. The effect of retention on property and drug crimes away from the cutoff is larger than our RD estimates shown in Tables 4 and 5.

Online Appendix Figures A-1 through A-4 show these results graphically and demonstrate why the RD estimates are much smaller than the estimated effects away from the cutoff for property and drug crimes. The "extrapolated" counterfactual lines tend to have no 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. Thus, 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 10 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, these results suggest that increasing the 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 10 indicate that the effect of grade retention on adult crime is large and ubiquitous among the low-performing students we examine.

5.4 Robustness Checks

The results presented in Section 5.1 point to a large increase in violent criminal convictions among adults when they are held back because of failing the 8th grade promotion exam. In this section, we explore the sensitivity of these results to alternative modeling assumptions. In Online Appendix Tables A-1 and A-2, we alter the bandwidth and the modeling of the running variable using criminal conviction and the number of convictions as dependent variables, respectively. Column (1) shows estimates using local linear regression with the Imbens and Kalyanaraman (2012) optimal smoother. In column (2) we alter the baseline bandwidth of 30 index points to 15, while column (3) increases it to 45. Column (4) presents estimates that use a quadratic spline in the running variable with a bandwidth of 60, and finally column (5) employs a global quadratic spline. The estimates for violent crime change very little across specifications in both tables. The drug crime estimates are somewhat sensitive to specification, with negative estimates in columns (1) and (2). However, these estimates are very imprecise, so 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. Online Appendix Tables A-3 and A-4 show results that use all convictions up to age 28 for our analysis sample. We only observe outcomes at age 28 for the 1984 birth cohort, at age 27 for the 1983-1984 birth cohorts, and at age 26 for the 1982-1984 birth cohorts. Thus, these estimates are based on an unbalanced panel of ages, but they allow us to examine effects past age 25. The estimates are extremely similar to baseline, which is expected given the lower likelihood of criminal engagement when individuals age into their mid 20s. The remaining results shown in Appendix Tables A-3 and A-4 exclude the regions most affected by Hurricane Katrina and Hurricane Rita in column (2),[29] add 7th grade composite test score controls in column (3), and cluster standard errors at the level of the running variable (i.e., the index) in column (4). These changes to the baseline model have minimal effects on the results: the point estimates and standard errors change little across columns. Thus, our main results and conclusions are not sensitive to these core modeling assumptions.

---

[29]Hurricane Katrina hit the Gulf Coast in August of 2005 and caused massive flooding and damage in areas of Louisiana around New Orleans, much of which was due to flooding from breached levees in several New Orleans parishes. Hurricane Rita hit Louisiana less than a month after Katrina and caused additional flooding and damage in coastal areas and in parishes most impacted by Katrina.

Finally, we conduct placebo tests in which we estimate our baseline model using false cut-offs 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 Figures 5 and 6 we plot the distribution of point estimates. The vertical red line shows the effects from Tables 4 and 5, 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 5) as well as for the number of violent crime and drug crime offenses (Figure 6), the RD estimates are in the tails of the distribution of point estimates. In Figure 5, the actual point estimate is larger than all but 1% of the placebo estimates for violent crime, and the point estimate for the number of violent offenses is larger than all but 3% of the placebo estimates in Figure 6. Thus, our violent crime results are driven by changes in criminal convictions at the retention cutoff and provide larger estimates than changes at the vast majority of placebo cutoffs.

6   Conclusion

We present the first estimates in the literature of the effect of grade retention on adult criminal convictions. The data we use come from a merge of administrative education records from the Louisiana Department of Education with administrative data on adult criminal convictions from the Louisiana Department of Public Safety and Corrections, Adult Services. To overcome the endogeneity of student retention, we exploit test score cutoffs imposed by the State of Louisiana for promotion to 9th grade whereby a student must repeat 8th grade if they fail the math or ELA state exam in July. We estimate both reduced form effects of failing the exam as well as the effect of 8th grade retention in an IV framework that imposes the additional assumption that the only reason test failure in 8th grade is related to crime is because of retention.

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.

The rich administrative data we use allow us to 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. 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. Retention has a large effect on dropout, increasing the likelihood by 7.17 percentage points (or 15.66% relative to the mean). Similar to Eren et al. (2017), we do not find a statistically significant impact on juvenile crime, though the estimate is positive. Thus, the main mechanism is unlikely to be increased juvenile criminal proclivity, which then leads to higher criminal activity as an adult. We argue that the evidence is most consistent with retention lowering the rate of human capital accumulation, and in particular non-cognitive skills, which results in higher levels of criminal engagement in early adulthood. Finally, 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 8th 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. That our estimates generalize to a larger set of inframarginal students highlights that the effects we report would be detected at any potential promotion cutoff set by the Louisiana Department of Education and is not driven by the particular threshold being set for promotion. The policy implications of these findings are very important, as they suggest any test-based retention in 8th grade among low-performing students leads to more criminal engagement later in life. Test-based promotion policies are designed to ensure that students meet basic knowledge levels to advance in school, which may

have benefits to both marginal and inframarginal students. Our results show that there are also private and social costs, which are substantial in magnitude. Considering only violent crimes, simple back-of-the-envelope calculations imply a social cost between $2.6 and $18.4 million. How to structure test-based promotion policies, or promotion policies more broadly, in a way that retains the benefits from imposing educational standards 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] 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.

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

[5] 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.

[6] 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.

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

[8] 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.

[9] Davis, Jonathan M.V. and Sara B. Heller. 2017. "Rethinking the Benefits of Youth Employment Programs: The Heterogeneous Effects of Summer Jobs." NBER Working Paper No. 23443.

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

[11] 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.

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

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

[14] 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.

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

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

[17] 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.

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

[19] 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.

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

[21] Imbens, Guido and Karthik Kalyanaraman. 2015. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *The Review of Economic Studies* 79(3): 933-959.

[22] 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.

[23] 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.

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

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

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

[27] 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

[28] 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.

[29] 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.

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

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

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

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

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

[35] 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.

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

[37] 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.

[38] Stephens, Melvin Jr. and Dou-Yan Yang. 2014. "Compulsory Education and the Benefits of Schooling." *American Economic Review* 104(6):1777-1792.

[39] Wakefield, Sara and Christopher Wildeman. 2011. "Mass Imprisonment and Racial Disparities in Childhood Behavioral Problems." *Criminology and Public Policy* 10(3): 793-817.

[40] Wildeman, Christopher and Bruce Western. 2010. "Incarceration in Fragile Families." *The Future of Children* 20(2): 157-177.

Figure 1: Distributions of Students around the July Promotional Cutoff
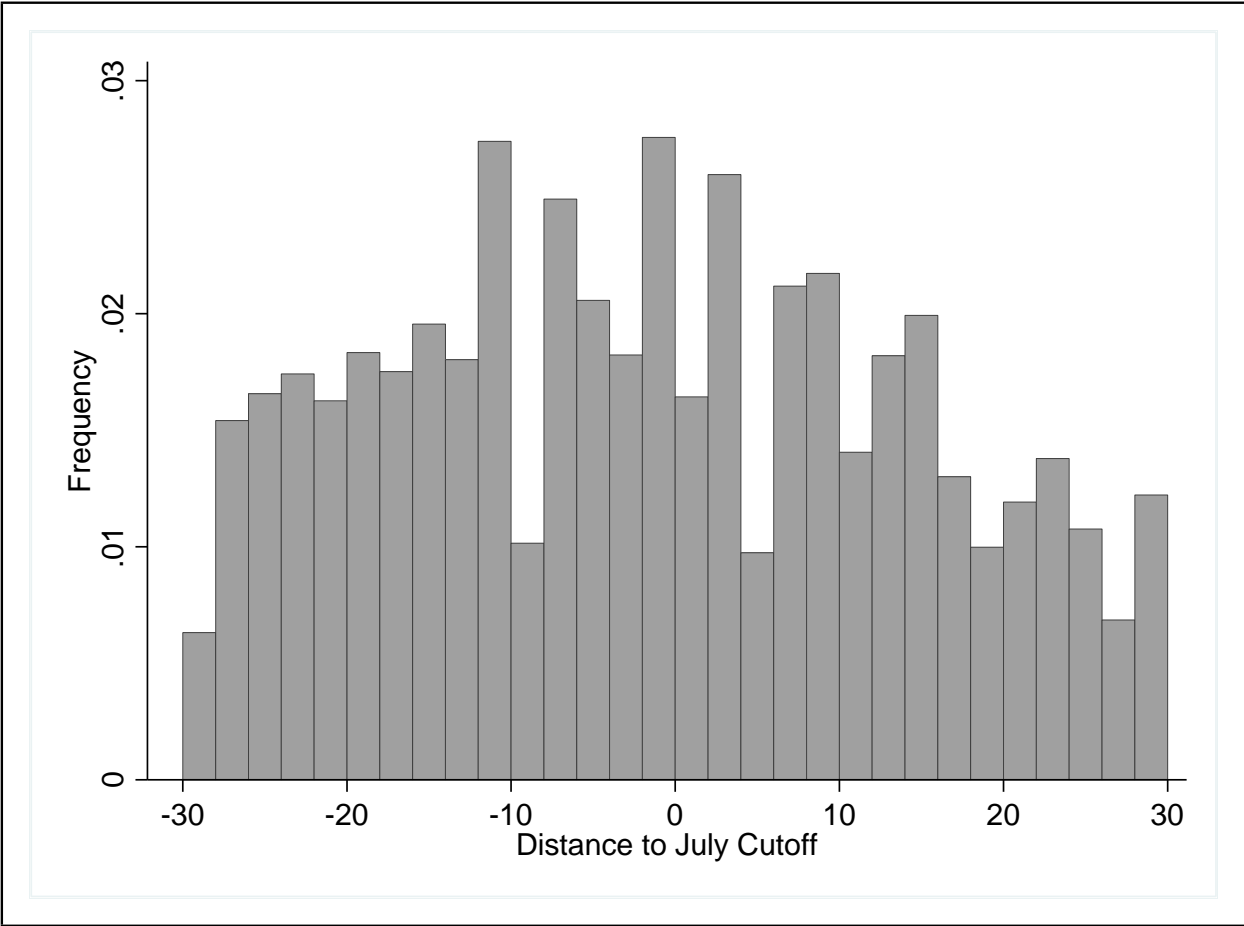
The figure shows the distribution of index scores relative to the July promotional cutoff (centered at zero), using the analysis sample with a bandwidth of 30 as described in the text.
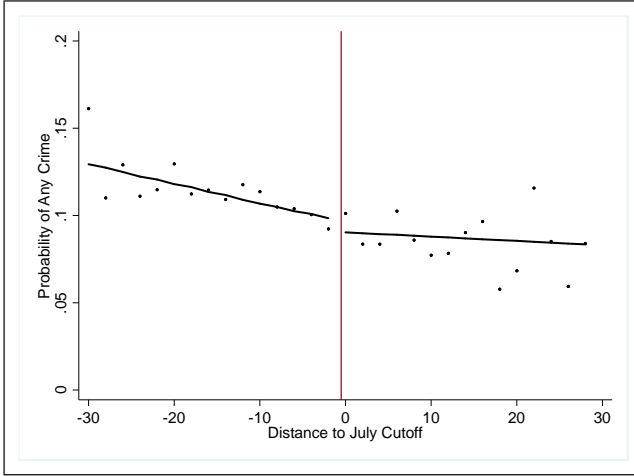
Figure 2: Probability of Grade Retention and Distance to the July Promotional Cutoff
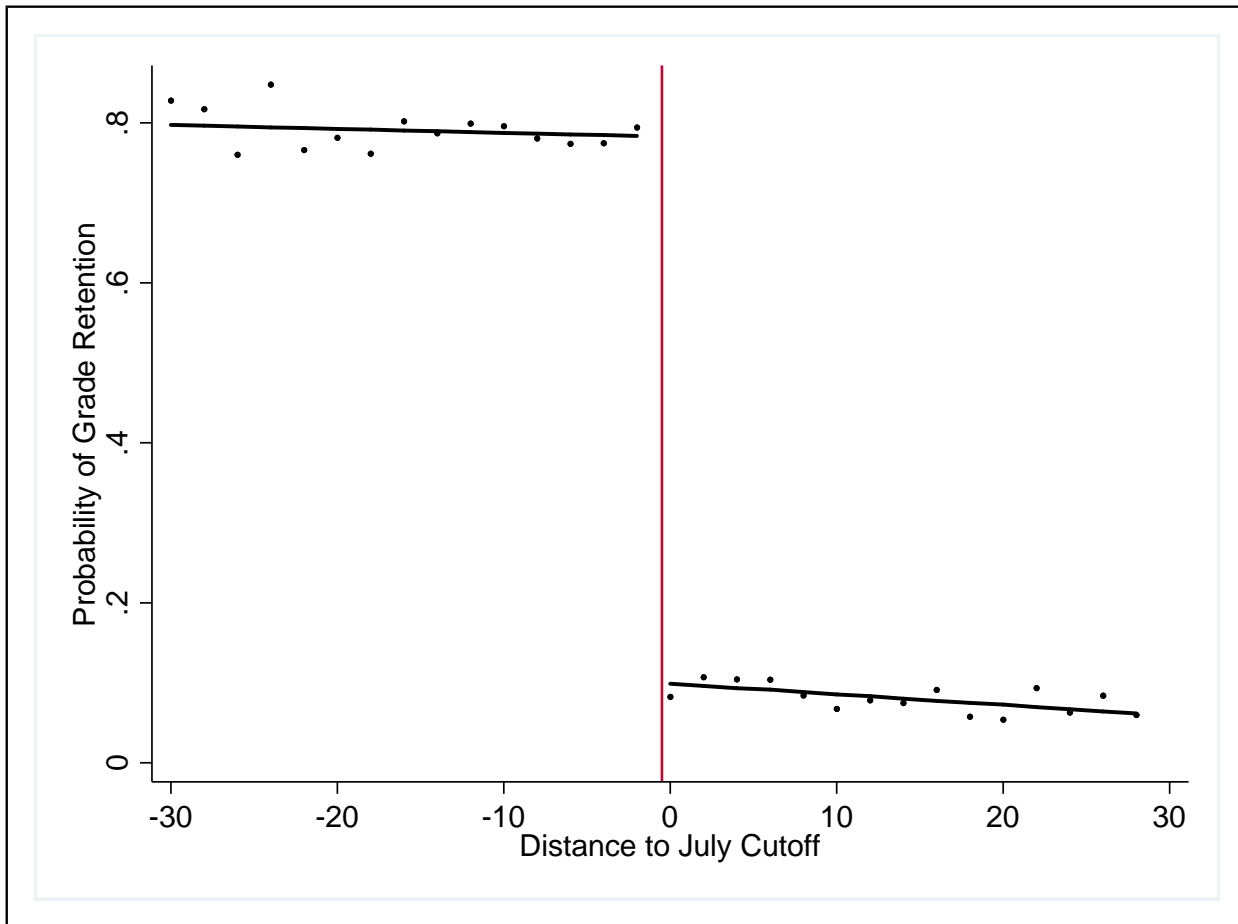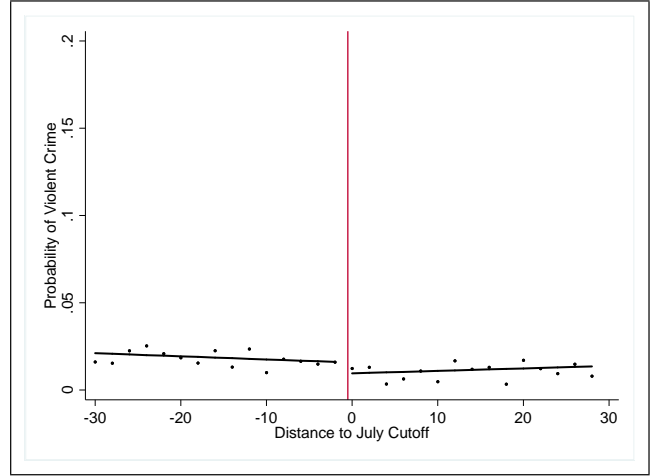

The vertical line denotes -0.5 points to the left of the July promotional cutoff (centered at 0). Each circle represents the unconditional mean of grade retention in two index point bins, based on the distance to July cutoff. The solid lines are fitted values of probability of grade retention from a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

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

Panel A: Any Crime

Panel B: Violent Crimes

Panel C: Property Crimes

Panel D: 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 4: The Number of Adult Crime Convictions by Distance to the July Test Cutoff
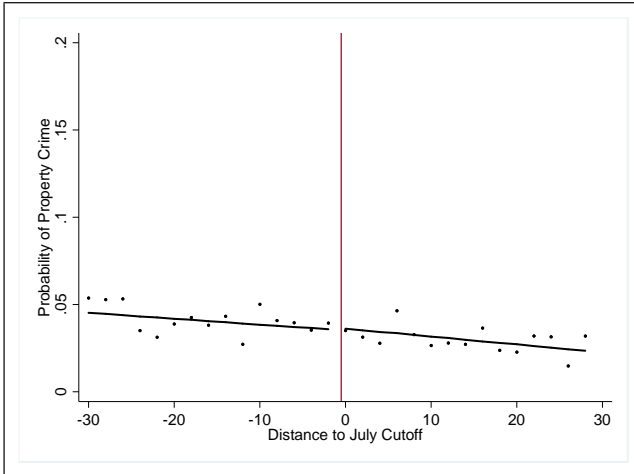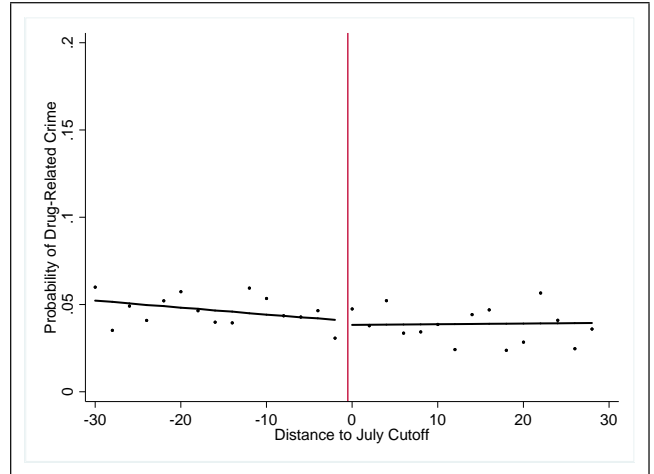
Panel A: Any Crime

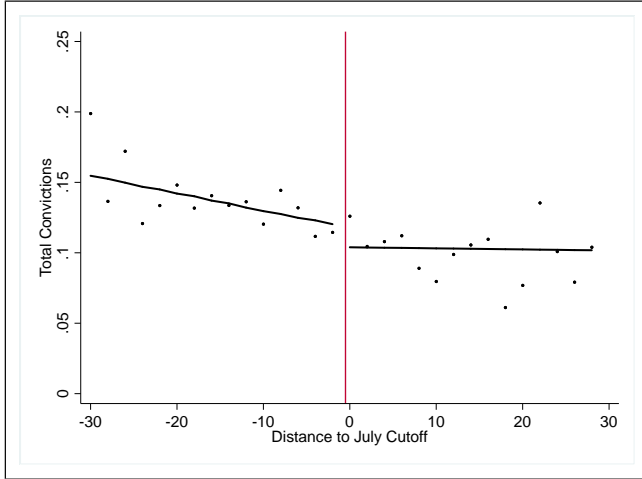Panel B: Violent Crimes

Panel C: Property Crimes

Panel D: Drug-Related Crimes
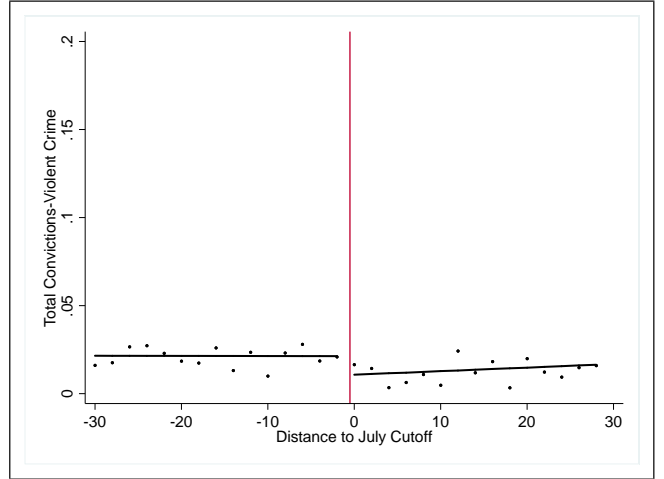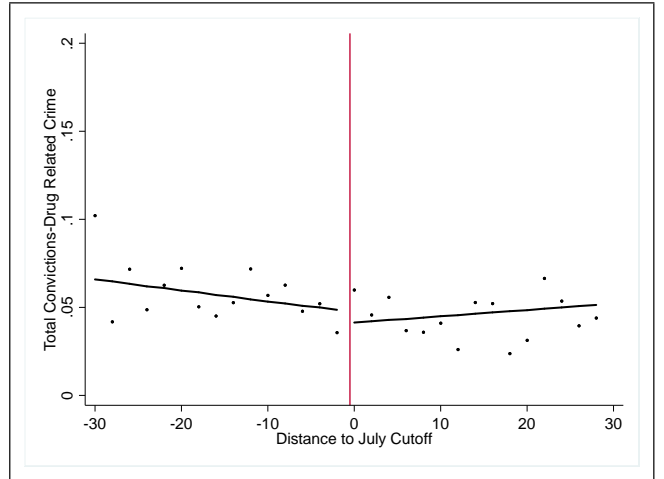

The vertical line denotes -0.5 points to the left of the July promotional cutoff (centered at 0). Each circle represents the unconditional mean of the number of adult crime convictions, overall and by type, in two index point bins based on the distance to July cutoff. The solid lines are fitted values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

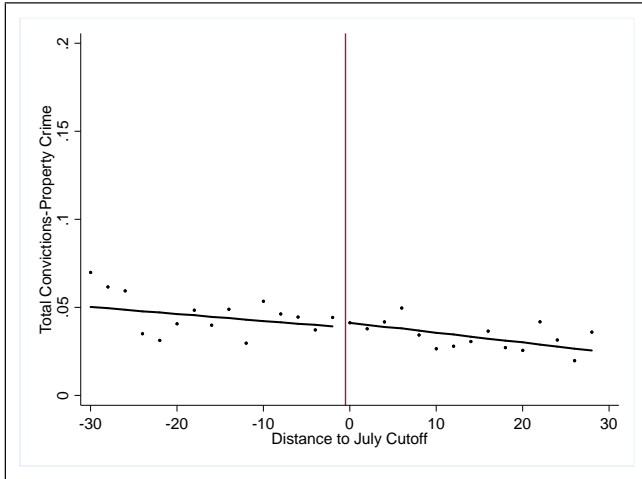**Figure 5: Placebo Estimates of the Effect of Failing a July Promotional Cutoff on the Probability of Adult Criminal Conviction**

Panel A: Violent Crimes
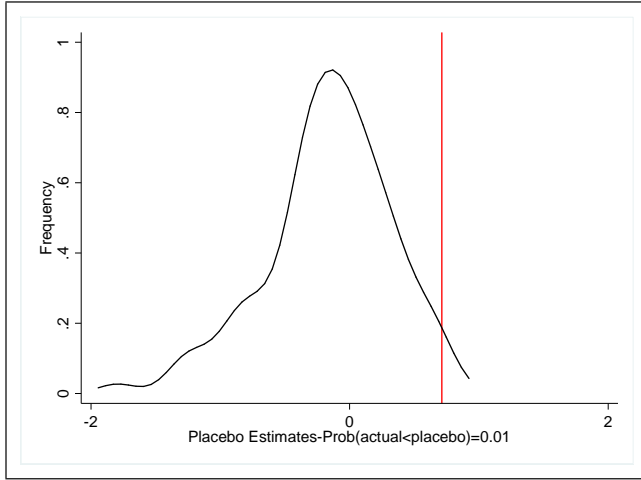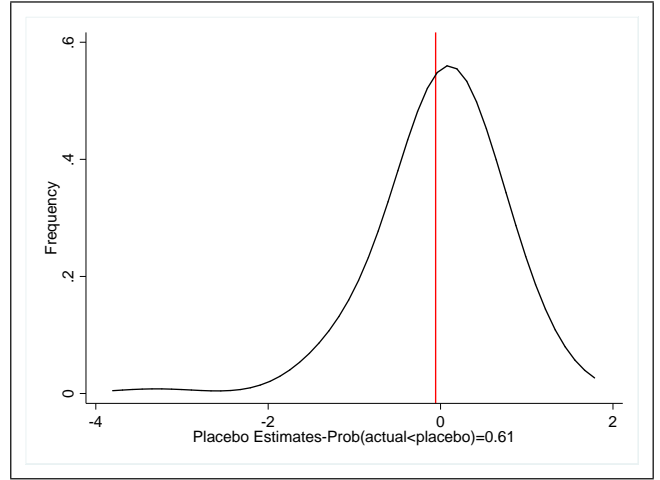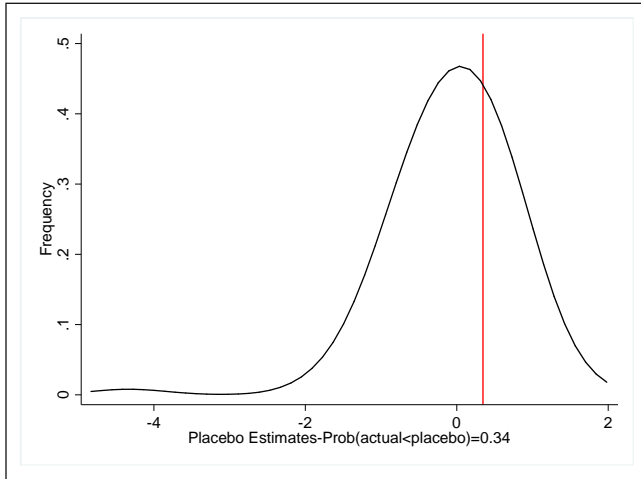
Panel B: Property Crimes

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 6: Placebo Estimates of the Effect of Failing a July Promotional Cutoff on the Number of Adult Criminal Convictions**

Panel A: Violent Crimes
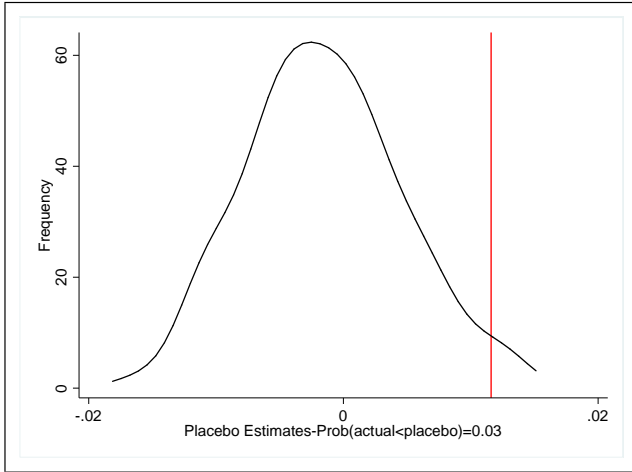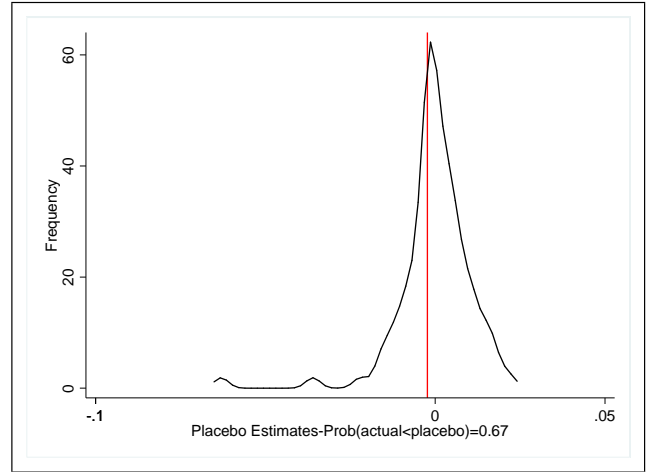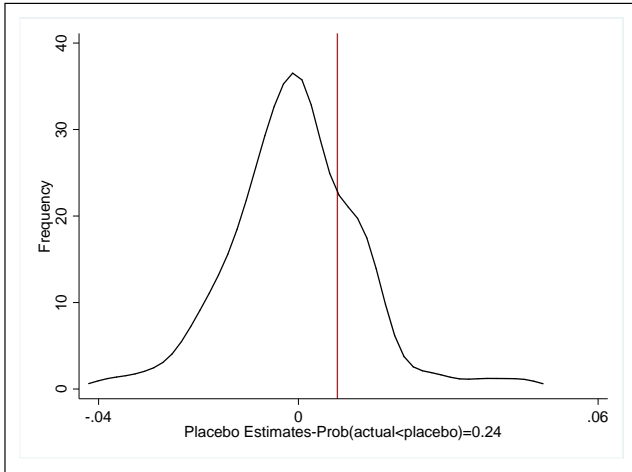
Panel B: Property Crimes

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 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 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						
Adult Conviction	0.115	0.319	0.110	0.313	0.119	0.324
Adult Crime Type:						
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
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 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)	7th Grade Composite Test Score (6)
Failed July Promotion Cutoff	-0.003 (0.007)	0.011 (0.014)	-0.011 (0.014)	0.016 (0.013)	-0.017 (0.014)	0.006 (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 7th 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 8th Grade		
(1)	(2)	(3)
Failed July Promotion Cutoff		
0.677*** (0.018)	0.678*** (0.018)	0.676*** (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 the Likelihood of Adult Crime Conviction (in Percent)

	Reduced Form			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Convicted of Any Crime (x100)						
Failed July Promotion Cutoff	0.687	0.846	0.750			
	(0.875)	(0.864)	(0.886)			
Grade Retention				1.015	1.248	1.109
				(1.290)	(1.270)	(1.283)
Panel B: Convicted of a Violent Crime (x100)						
Failed July Promotion Cutoff	0.668*	0.713**	0.648**			
	(0.352)	(0.351)	(0.261)			
Grade Retention				0.987*	1.052**	0.959**
				(0.518)	(0.516)	(0.379)
Panel C: Convicted of a Property Crime (x100)						
Failed July Promotion Cutoff	-0.072	-0.057	-0.025			
	(0.569)	(0.573)	(0.587)			
Grade Retention				-0.107	-0.084	-0.037
				(0.841)	(0.845)	(0.851)
Panel D: Convicted of a Drug-Related Crime (x100)						
Failed July Promotion Cutoff	0.254	0.346	0.254			
	(0.598)	(0.586)	(0.598)			
Grade Retention				0.375	0.511	0.376
				(0.883)	(0.862)	(0.867)
Sample Size	14,728	14,728	14,728	14,728	14,728	14,728
Controls:						
Test Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	Yes	No	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	Yes

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. Dependent variables take the value of one if the student is convicted as an adult at age 25 or younger. All outcome variables are 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 5: Reduced Form and IV Estimates of Test Failure and Grade Retention on the Number of First-time Criminal Convictions by Age 25

	Reduced Form			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Any Crime (Mean=0.136)						
Failed July Promotion Cutoff	0.015	0.017	0.018			
	(0.012)	(0.012)	(0.012)			
Grade Retention				0.022	0.025	0.027
				(0.018)	(0.018)	(0.018)
Panel B: Violent Crimes (Mean=0.020)						
Failed July Promotion Cutoff	0.011**	0.012**	0.011**			
	(0.005)	(0.005)	(0.005)			
Grade Retention				0.016**	0.017**	0.017**
				(0.007)	(0.007)	(0.007)
Panel C: Property Crimes (Mean=0.043)						
Failed July Promotion Cutoff	-0.002	-0.002	-0.001			
	(0.007)	(0.007)	(0.007)			
Grade Retention				-0.004	-0.003	-0.002
				(0.010)	(0.010)	(0.010)
Panel D: Drug Related Crimes (Mean=0.060)						
Failed July Promotion Cutoff	0.007	0.008	0.008			
	(0.008)	(0.008)	(0.008)			
Grade Retention				0.010	0.012	0.011
				(0.011)	(0.011)	(0.011)
Sample Size	14,728	14,728	14,728	14,728	14,728	14,728
Controls:						
Test Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	Yes	No	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	Yes

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. Dependent variables are the number of convictions up to age 25. Standard errors clustered at the school level are in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 6: Mechanism - The Effect of Grade Retention on High School Peer Composition

	Missing High School Data (1)	Percent of Special Ed Students in High School (2)	Percent of White Students in High School (3)	Average 9th Grade Math Standardized Test Score (4)
Failed July Promotion Cutoff	0.210***			
	(0.015)			
Grade Retention		-0.001	-0.027	-0.062**
		(0.004)	(0.017)	(0.027)
Sample Size	15,501	11,115	11,115	11,115

Authors' estimation of equation (4) as described in the text. Missing high school data occur because students do not enroll in a public high school in Louisiana. 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 7: Mechanism - Regression Discontinuity Estimates of Grade Retention on Absenteeism and Disciplinary Incidents

	Missing Outcome Data (1)	Elapsed Time Relative to July Exam Year 1 (2)	Year 2 (3)	Year 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 8: 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*** (2.334)	0.528 (0.811)	0.335 (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 9: The Effect of Grade Retention on Adult Criminal Convictions, by Age

	Age of Adult Crime≥19 (1)	Age of Adult Crime≥20 (2)
Panel A: Any Violent Crime Conviction (x100)		
Grade Retention	0.835* (0.477)	0.673* (0.402)
Panel B: Any Property Crime Conviction (x100)		
Grade Retention	-0.327 (0.736)	-0.151 (0.684)
Panel C: Any Drug-Related Crime Conviction (x100)		
Grade Retention	0.306 (0.764)	0.269 (0.750)
Sample Size	14,728	14,728

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. The dependent variable in Column (1) takes the value of one if the individual is 19 years old or older at the time of the offense, while in Column (2) the dependent variable is an indicator equal to 1 if the individual is 20 years old or older at the time of the offense. Both outcome variables are 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 10: The Effect of Retention on Adult Criminal Convictions Away from the Cutoff
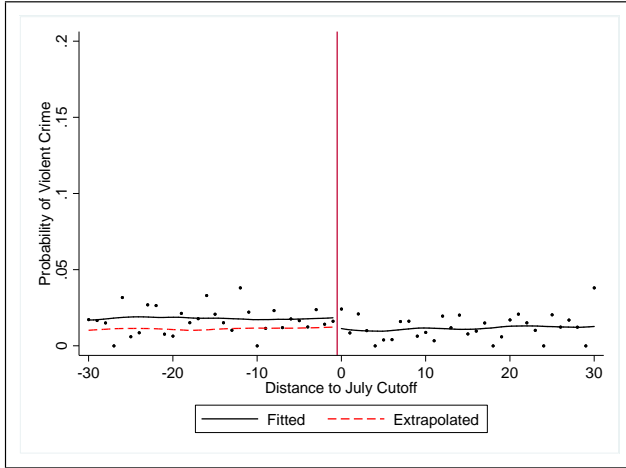
	Treatment Effect on the Treated (Extensive Margin) (1)	Treatment Effect on the Treated (Intensive Margin) (2)	Treatment Effect on the Untreated (Extensive Margin) (3)	Treatment Effect on the Untreated (Intensive Margin) (4)
Panel A: Violent Crime				
Grade Retention	1.055***	0.013***	0.920***	0.010***
	(0.263)	(0.003)	(0.283)	(0.003)
Panel B: Property Crime				
Grade Retention	1.423***	0.017***	1.481***	0.017***
	(0.448)	(0.005)	(0.505)	(0.006)
Panel C: Drug-Related Crime				
Grade Retention	1.010	0.016*	0.904	0.014*
	(0.700)	(0.009)	(0.564)	(0.007)
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 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 using a nonparametric block bootstrap at the school level with 500 replications are in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.
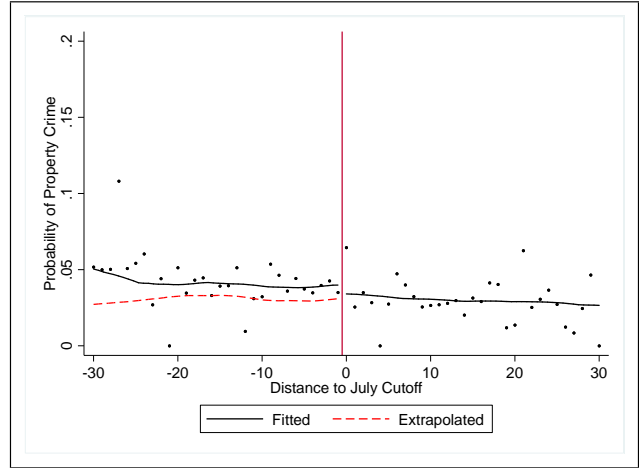
Online Appendix: Not for Publication

Figure A-1: Linear Reweighting Estimates of the Counterfactual-Treatment Effect on the Treated (Extensive Margin)
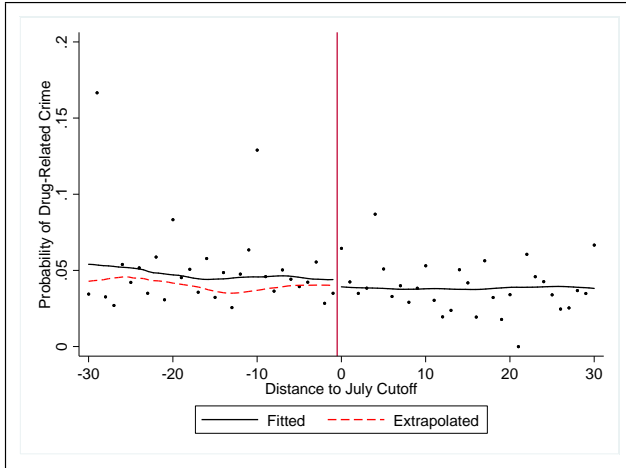
Panel A: Violent Crimes
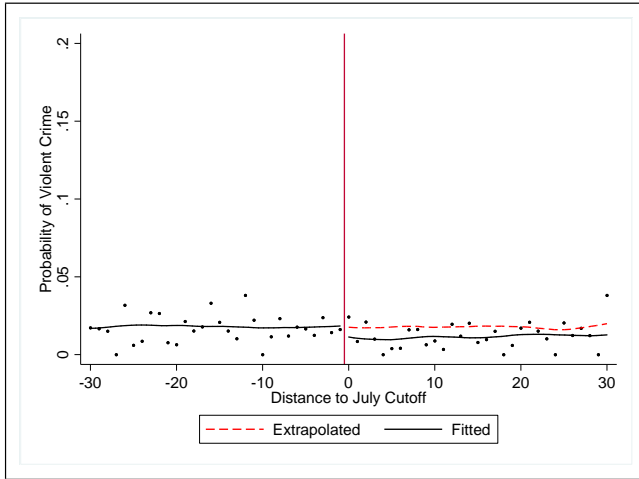

Panel B: Property Crimes

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-2: Linear Reweighting Estimates of the Counterfactual-Treatment Effect on the Untreated (Extensive Margin)

Panel A: Violent Crimes

Panel B: Property Crimes
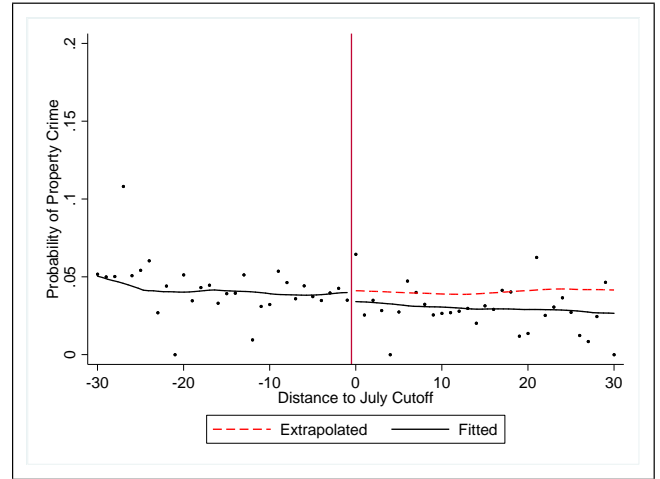

Panel C: Drug-Related Crimes

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-3: Linear Reweighting Estimates of the Counterfactual-Treatment Effect on the Treated (Intensive Margin)
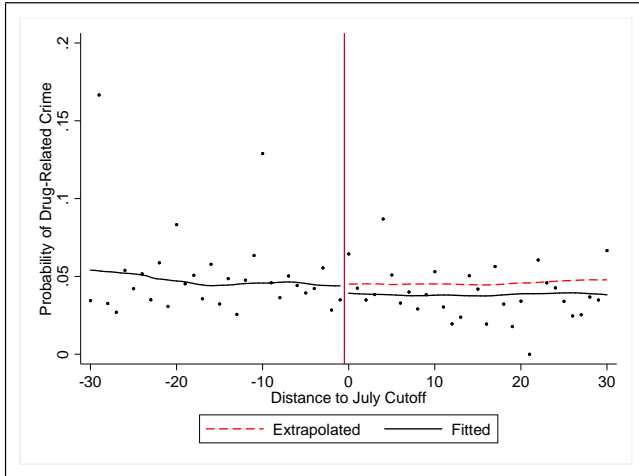
Panel A: Violent Crimes
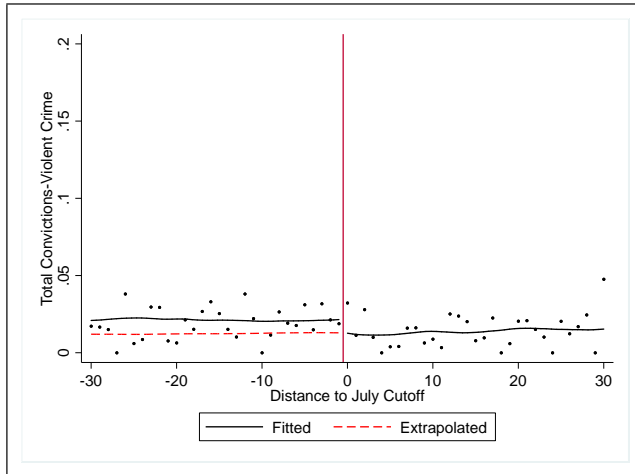

Panel B: Property Crimes

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-4: Linear Reweighting Estimates of the Counterfactual-Treatment Effect on the Untreated (Intensive Margin)

Panel A: Violent Crimes

Panel B: Property Crimes
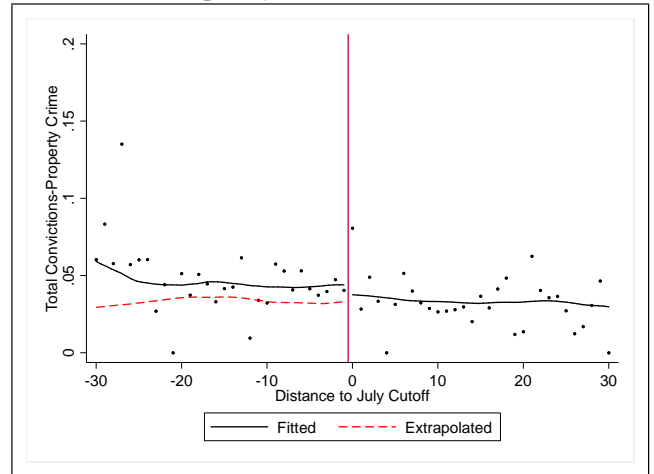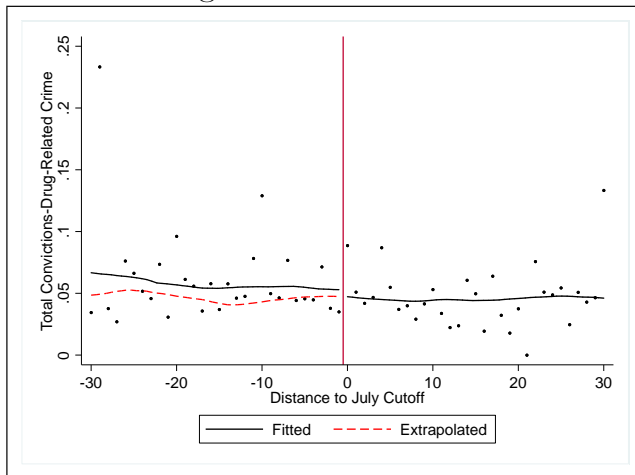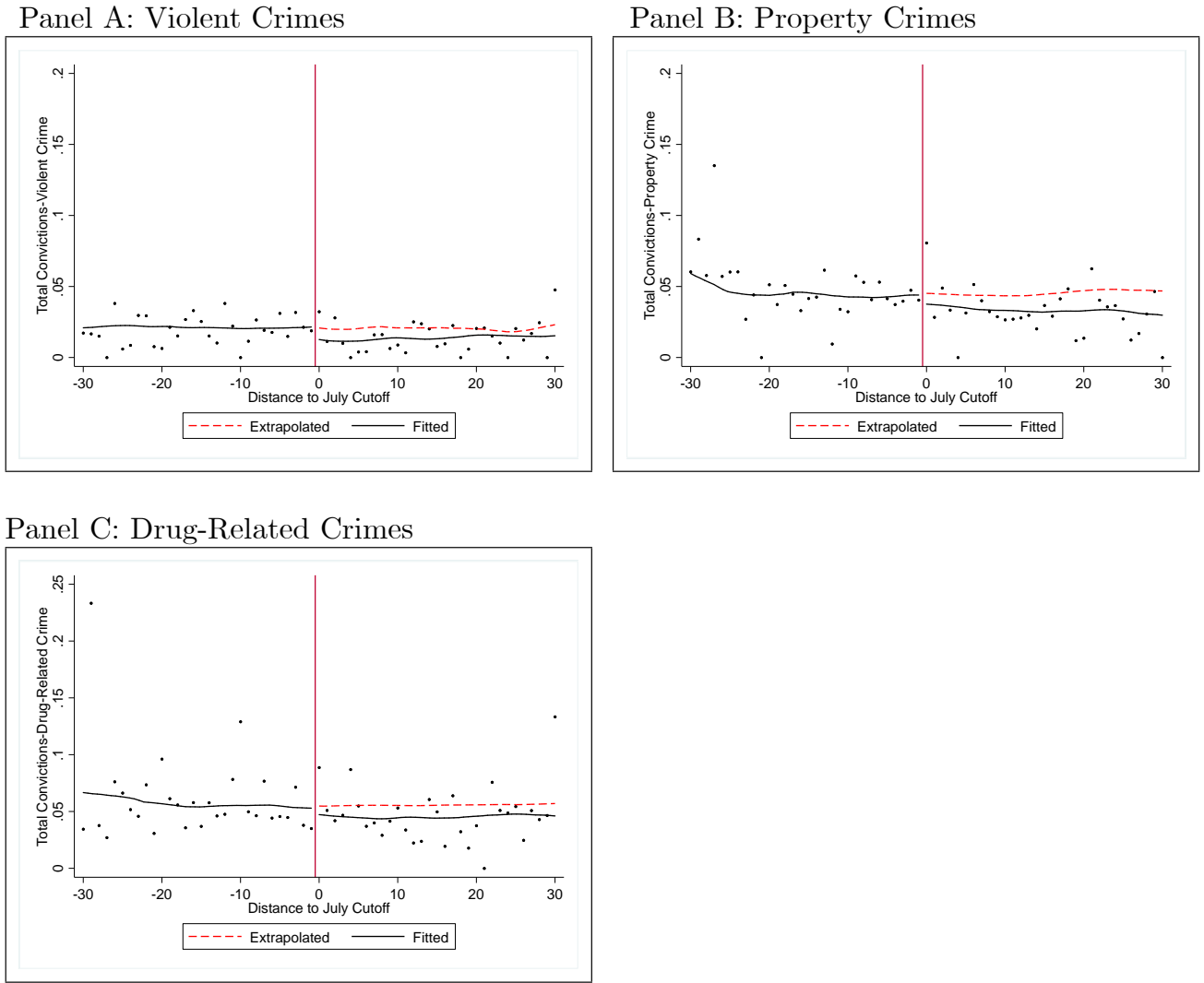

Panel C: Drug-Related Crimes

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.

Table A-1: 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)	Linear Spline Index=[-15,15]	Linear Spline Index=[-45,45]	Quadratic Spline Index=[-60,60]	Quadratic Spline (Global)
	(1)	(2)	(3)	(4)	(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, and immigration status. 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-2: 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, and immigration status. 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 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 Baseline Test Scores (3)	Standard Errors Clustered at the Index Level (4)
Panel A: Convicted of a Violent Crime (x100)				
Grade Retention	1.178**	1.397**	1.080**	1.052***
	(0.528)	(0.612)	(0.526)	(0.404)
Panel B: Convicted of a Property Crime (x100)				
Grade Retention	-0.067	-0.499	0.058	-0.084
	(0.876)	(0.973)	(0.850)	(0.658)
Panel C: Convicted of a Drug-Related Crime (x100)				
Grade Retention	0.374	0.730	0.510	0.511
	(0.937)	(0.947)	(0.860)	(0.783)
Sample Size	14,728	10,199	14,378	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 7th grade composite scores into the specifications. Covariates include indicators for gender, race, free/reduced lunch and immigrant status. All outcome variables are multiplied by 100 to obtain percent values. Standard errors in parentheses are clustered at the school level in columns (1)-(3) and at the index (running variable) level in column (4): * 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 Number of Adult Crime Convictions

	Number of Convictions Up to Age 28 (1)	Drop Hurricane Katrina/Rita Regions (2)	Add Baseline Test Scores (3)	Standard Errors Clustered at the Index Level (4)
Panel A: Number of Violent Crime Convictions				
Grade Retention	0.018**	0.020**	0.018**	0.017***
	(0.007)	(0.008)	(0.007)	(0.005)
Panel B: Number of Property Crime Convictions				
Grade Retention	-0.003	-0.009	-0.000	-0.003
	(0.010)	(0.011)	(0.010)	(0.008)
Panel C: Number of Drug-Related Crime Convictions				
Grade Retention	0.010	0.013	0.012	0.012
	(0.011)	(0.011)	(0.011)	(0.012)
Sample Size	14,728	10,199	14,378	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 7th grade composite scores into the specifications. Covariates include indicators for gender, race, free/reduced lunch and immigrant status. Standard errors in parentheses are clustered at the school level in columns (1)-(3) and at the index (running variable) level in column (4): * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A-5: 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 fixed effects. Standard errors in parentheses are clustered at the school level: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A-6: 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 (1.651)	0.262 (0.267)	3.042 (8.492)	-0.264*** (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 immigrant status. Standard errors clustered at the school level are in parentheses: * significant at 10%, ** significant at 5%, *** significant at 1%.