Spare the Rod? The Effect of No Child Left Behind on Failing Schools *

Preliminary and Incomplete–Comments Welcome!

Jane Cooley† and Jeffrey Tracynski‡

March 14, 2011

Abstract

Under No Child Left Behind, schools face sanctions after repeatedly failing to meet proficiency targets. In this paper, we explore school responses to failing to meet proficiency targets. Using data on North Carolina public elementary schools, our primary strategy for isolating an effect of failing is a regression discontinuity design that compares schools around the threshold for passing. We find evidence that failing schools improve performance, particularly when they are facing sanctions for failing. We also find that improvements do not come at the expense of high-achievers, but may derive at the expense of improvement for the lowest-achievers. We also explore whether schools target resources toward failing subgroups of students. We then consider the extent to which the disparities in estimated effects across sanction statuses can be attributed to a direct effect of the sanctions, particularly looking at how effects vary by the intensity of the sanction.

Keywords: accountability, sanctions, school incentives

*We would like to thank Karl Scholz, Chris Taber, Jean Francois Houde, John Micklewright, and participants at numerous seminars for helpful comments and suggestions. We are also grateful to the North Carolina Education Research Data Center for providing the data. Cooley’s work was supported by the University of Wisconsin’s Institute for Research on Poverty and Graduate School Fellowships. All remaining errors are our own.

†Department of Economics, University of Wisconsin-Madison, Madison, WI 53706; jcooley@ssc.wisc.edu.

‡Department of Economics, University of Wisconsin-Madison, Madison, WI 53706; traczynski@wisc.edu.
1 Introduction

The No Child Left Behind (NCLB) Act of 2002 marked a new era in federally mandated accountability for public schools in the United States. NCLB requires that schools meet proficiency targets each year in order to be designated as making adequately yearly progress (AYP). With NCLB up for reauthorization, considerable questions remain about the effect of NCLB’s particular structure of accountability on failing schools.

In this paper, we explore the effect of failing AYP on school performance in North Carolina. A key challenge in addressing this question is that schools who fail AYP are also likely to have worse performance than schools that pass, even in the absence of failure. We bring new evidence to bear on the effect of failing AYP in North Carolina, using a regression discontinuity approach. We compare schools just above the threshold for passing to schools just below. The intuition is that schools around the threshold are similar, except that random measurement error in test scores causes some to meet AYP and some to fail.

A key question in the literature is how accountability affects traditionally disadvantaged students, and whether benefits to targeted students derive at the cost to other students (see Ballou and Springer [2008], Neal and Schanzenbach [2010], Krieg [2008], Ladd and Lauen [2010], among others). Similar to prior studies, we explore the differential impact of failing AYP by prior achievement level of the student for North Carolina, and in particular whether schools focus on marginal students, leaving the lowest performers behind, as evidenced in Neal and Schanzenbach [2010].

AYP is determined not only based on overall proficiency, but also requires that schools meet proficiency counts in up to 10 subgroups, including traditionally disadvantaged racial and socioeconomic subgroups. The subgroup rules are an important feature of NCLB, but have not been explored much in the literature. In part, this is because many studies on the effects of accountability rely on evidence from states, such as Florida and Texas, that had accountability prior to NCLB, and these pre-existing systems did not include the subgroup rules. As we show in our data, the subgroup rules, particularly disabled and limited English proficient, appear to be the dominant reason that schools fail. We bring new evidence on the effect of subgroup rules, particular whether schools focus resources toward failing students in failing subgroups, perhaps at the detriment of failing students in other subgroups.

Another key contribution of our paper is then to delve into the mechanisms of the school’s response to failing AYP. For instance, failing AYP can affect schools simply through the stigma effect of failing or fear of future dissolution as a result of repeated
failure. However, NCLB also entails a particular structure of sanctions that schools face after repeated failure. While the literature has thought carefully about the potential merits and pitfalls of accountability based on proficiency counts, considerably less is known about how sanctions affect school performance. The particular structure of sanctions mandated by NCLB, school choice for students in failing schools and provision of extra tutoring, could either improve or hurt the ability of failing schools to meet AYP standards in subsequent years.

We do this first by comparing schools who face sanctions after failing AYP to those who fail AYP absent sanctions. Evidence of heterogeneous responses across sanction statuses might begin to suggest that sanctions play a role in the response of failing schools. However, this evidence alone is not conclusive, in particular because it remains difficult to determine whether the estimated effect is in fact a result of sanctions rather than an effect of repeated failure. We attempt to detect an effect of sanctions by determining how effects differ by schools based on the intensity of treatment, such as the number of school options available to students in failing schools.

We take advantage of panel data on test scores for students in North Carolina public schools to isolate an effect of failing AYP on the growth in student achievement. Applying regression discontinuity techniques to our question is not immediately straightforward because the of the multidimensionality implicit in the subgroup rules that determine whether a school makes AYP. Past studies that have attempted to apply regression discontinuity in this multi-dimensional setting have focused on the subgroup that is closest to the threshold for passing to define the distance from passing [Bacolod et al., 2009, Ahn and Vigdor, 2009]. We develop a new approach which counts the number of test passes a school would need in order to make AYP. In contrast to previous methods, this takes into account the fact that a school that fails to make AYP in 5 subgroups by 10% is different than a school that fails to make AYP by 10% in 1 subgroup. In our context, we find that this produces much sharper predictions of AYP status.

While very useful for approximating random assignment for schools on the threshold for passing, a potential criticism of the regression discontinuity approach is that it only detects effects for schools on the margin of passing. This could understate the overall effect of failing AYP, if schools just above the margin still work hard to avoid failure in subsequent years and thus face similar incentives. On the other hand, if there is a direct effect of sanctions, it may be the right approach for detecting this effect relative to the basic effect of failing (or being close to failing). On the other hand, using regression discontinuity could overstate the effect of failing AYP and/or sanctions, if failing schools close to the margin have the most hope of passing in subsequent years and thus work
harder relative to schools well below the margin for passing.

Thus, we supplement our regression discontinuity analysis with a difference-in-difference approach, which exploits the fact that only Title I schools (schools receiving federal funds) are subject to sanctions. Report cards for all schools (independent of Title I status) indicate whether the school failed to make AYP. Because of variation in levels of poverty across districts, there is a great deal of heterogeneity in designation of Title I schools across districts (i.e., equally poor schools could be Title I schools in one district and not Title I schools in another). Our difference-in-difference strategy uses non-Title I schools that would have faced sanctions in a given year given their record of repeatedly failing AYP as a control for Title I schools that actually face sanctions. We also includes a rich set of student and school characteristics, along with school fixed effects, to help soak up any remaining heterogeneity in the two types of schools that might affect their ability to produce student achievement growth.

Previous studies have considered the effect of state and district-level accountability systems on schools and achievement in a variety of contexts (e.g., Reback [2008], Jacob [2005], Neal and Schanzenbach [2010], Rockoff and Turner [2008], Figlio and Rouse [2006], and Chakrabarti [2010]). While these studies provide important insight into the potential effects of accountability, less is known about the effect of NCLB. Studying the effect of NCLB is important given the unique structure of sanctions and the subgroup rules, which are not necessarily shared in these other settings. The number of studies on NCLB is growing, though their focus is substantially different from ours. Dee and Jacob [2009] consider differential responses of states to NCLB based on whether the had prior accountability policies in place, using variation in pre-accountability standards across states and differential jumps in achievement pre-/post-NCLB. Ballou and Springer [2008] focuses on the distributional effect of NCLB in 7 midwestern states, exploiting variation in accountability across grades. Krieg [2008] explores distributional responses in Washington State, comparing student achievement before and after NCLB is enacted. Ladd and Lauen [2010] also use panel data on student performance before and after NCLB to determine the distributional effect of NCLB in North Carolina.

Building on the literature that has investigated the effect of NCLB, we have three key contributions. First, we apply a new approach to the identification of the effect of failing AYP, using regression discontinuity and difference-in-difference. This helps to eliminate other time-varying unobservables that may confound the effect of NCLB. Second, we also explore directly the effect of the subgroup rules. Third, we explore the direct effect of sanctions on the ability of (or incentives for) failing schools to respond to NCLB pressure.
Ahn and Vigdor [2009] is most similar in spirit to our paper. They also use a regression discontinuity approach to student the effect of NCLB in North Carolina and also try to separate out a differential effect of being in sanctions versus failing AYP absent sanctions. Our regression discontinuity approach differs in terms of how we define the distance from passing, as described above and appears to produce sharper predictions. While they focus on overall achievement, our analysis expands on their findings by focusing on distributional effects and subgroup rules. We also expand on their analysis by looking for a direct effect of sanctions (as separate from an effect of repeated failure) through the intensity of treatment analysis. We also explore heterogeneity in responses across failing schools, in part by supplementing the regression discontinuity with the difference-in-difference strategy.

The paper proceeds as follows. Section 2 provides background on school accountability in North Carolina. Section 3 describes the data and accountability in North Carolina, along with more details on the structure of sanctions. In Section 4, we describe our estimation strategy for the regression discontinuity and difference-in-difference strategies. Section 5 presents our findings for the overall effect of failing AYP and distributional effects. Section 6 explores mechanisms of response and in particular whether effects vary by sanction status. We also attempts to uncover whether the estimated differential effects for schools facing sanctions derives from the sanctions themselves or other mechanisms, such as repeated failure or stigma effects. Section 7 concludes.

2 Background on Accountability

NCLB’s school accountability standards are based on proficiency counts. Proficiency in North Carolina public schools is defined as students performing at achievement level 3 or higher on the state standardized exams. The states set proficiency targets, which determine the percentage of students who should test at the proficient level. To make AYP, schools must satisfy these proficiency targets in each of 10 subgroups: white, black, Hispanic, Native American, Asian, multiracial, economically disadvantaged, Limited English Proficient (LEP), and students with disabilities along with the overall student population.¹

While much of the literature on accountability has focused on the important question of how to effectively measure school performance, i.e., through achievement growth,

¹These proficiency targets apply only to full academic year students, those who have attended school at least 140 days prior to Spring testing.
proficiency counts or other methods,\(^2\), an important question remains as to the effect of different types of incentives tied to meeting or failing to meet the standards. The rewards or sanctions attached to failing standards can be equally important as the standards themselves. In part, this is why it may be difficult to generalize from findings on state systems that used proficiency counts but with other rewards/sanctions to NCLB. Thus, understanding both the direct effect of failing AYP and the differential effect of the sanctions imposed under NCLB can be critical to evaluating the effectiveness of the policy.

Prior research shows that teachers and schools are very responsive to incentives, such as providing salary bonuses for meeting standards.\(^3\) Sanctions differ sharply from these bonus mechanisms in that, instead of being rewarded for good performance, schools are penalized for failing to meet given standards.

Under NCLB, schools who repeatedly fail to meet standards face a series of escalating sanctions. Figure 1 describes the sanction schedule for schools in North Carolina that fail to make AYP. Schools enter in year 1 of Title I Improvement after failing to make AYP for the same subject two years in a row. At this point, schools are required to give students the chance to attend other non-failing schools in the district and are required to provide transportation. Choice extends to all students and capacity constraints are not a valid reason not to permit choice. However, disadvantaged students get priority in choosing schools to attend. Once students choose to attend another school, they are allowed to stay there regardless of whether their original school exits Title I improvement status. This may be an important part of parents’ decisions given switching costs and that parents may be more reluctant to send their children to a school on a temporary basis.

On the one hand, school choice could encourage the school to work harder to keep its students, thereby improving student performance. On the other, if the better students leave, the low-performing school may have an even harder time meeting AYP standards in the next year. Likewise, if there are no viable school choice options within the district, the sanction would have no direct consequence. A considerable number of studies have investigated the effect of school choice in other settings, but the evidence is mixed.\(^4\) It is not clear from the existing literature whether the direct effect of the NCLB sanctions would exert positive pressure on lower-performing schools.

\(^2\)See Neal and Schanzenbach [2010], among others
\(^3\)For instance, studies have investigated how teachers respond to policies such as pay for performance or salary bonuses for meeting standards. See Ahn [2009], Vigdor [2009] for the North Carolina context; Figlio and Kenny [2007], Ladd [1999], Lary [2009], Glewwe et al. [2003], among others.

\(^4\)See Hoxby [2003] for an overview, Figlio and Rouse [2006], Chakrabarti [2010], among many others.)
After failing for a third consecutive year in the same subgroup and subject, schools enter year 2 of Title I improvement. At this point, schools must offer students the opportunity to receive extra tutoring or provide after-school programs for failing students, (referred to as supplemental education services or SES in the literature). Students are given the opportunity to either participate in the extra tutoring programs or to exercise choice by moving to a non-failing school. Relatively little is known about the effectiveness of these extra tutoring services. Often transportation is not provided to the after-school or tutoring program, and this has limited take up in many cases, according to Zimmer et al. [2007] and others. Furthermore, the types of programs, providers, access and opportunities vary considerably across districts. The studies that evaluate the effect of these supplemental educational services find at best mixed results [Zimmer et al., 2007, Heinrich et al., 2010]. Schools facing these sanctions could have higher student performance as a direct result of SES, could suffer from the redirection of resources away from classrooms.

After continued failure, a school is eventually restructured. While the form of this can vary, school are put under new governance and rapid change occurs. Thus, differential responses of schools from repeated failure could be in response to fear of being restructured or the stigma attached to failure, rather than the sanction per se. This will motivate our intensity of treatment analysis in Section 6. Finally, it is noteworthy that to exit Title I Improvement status, a school must pass the failing subject and subgroup for two years in a row.

2.1 Accountability in North Carolina

The federal government’s AYP standards were layered on top of a well-developed testing and accountability system in North Carolina. The state standards were implemented in 1996/97 and evaluated school quality based primarily on growth, measured by schools’ improvement in achievement from one year to the next. Under state accountability, teachers are rewarded with bonuses ranging from $500 to $1500 if schools meet growth targets. This system of bonuses remained in effect after the introduction of NCLB.

There are several important distinctions between the federal and state accountability standards. First, state standards are based primarily on growth in student achievement rather than proficiency counts. Second, state standards only consider overall growth and do not have targets for different subgroups of students. Third, the schools and teachers receive rewards for meeting state standards, rather than sanctions for failing to meet federal standards. Finally, state standards apply to all schools, whereas federal
standards apply only to Title I schools, which receive funds from the federal government as discussed further in Section 3.

Its considerable history with accountability prior to NCLB makes North Carolina a good context for studying the effect of sanctions. However, a potential problem is that schools are facing two types of incentives, and it may be difficult to separate an effect of sanctions imposed by AYP from the effect of the growth standards. Our regression discontinuity approach isolates an effect of sanctions on marginal schools, given that the probability of meeting growth standards did not experience a jump at the same threshold. Furthermore, our approach of using non-Title I schools as a control also helps isolate the effect of sanctions, given that non-Title I schools face the same state standards but not the sanctions for failing to meet federal standards. However, we also test the generalizability of our findings by considering whether estimated responses to AYP change after conditioning on whether the school met growth standards.⁵

3 Data

3.1 School Report Cards

We collect school report card data from the North Carolina Public Schools website (accrpt.ncpublicschools.org/docs/), which contains detailed information on how the schools perform on accountability standards. The state designs math and reading tests and designates a performance level on the test, achievement level 3, above which students are deemed proficient in the tested subject. Table 1 describes the average proficiency rates by each subgroup and by reading and math for schools where those subgroups of student exist.⁶ Whites and Asians have the highest proficiency rates in reading, 0.86 and 0.85 respectively, and math, 0.85 and 0.89. Disabled and LEP students have the lowest proficient rates in both subjects, between 0.50 and 0.56.

The state also designates target proficiency rates that determine the percentage of students in each subgroup who need to pass the exam in order for the school to make AYP. Under NCLB, the goal is that states are working towards 100% proficiency over time. Schools that do not meet proficiency targets in a given subgroup have several other ways that they can still make AYP. First, the safe harbor provision permits a

⁵Ladd and Lauen [2010] focuses more directly on how responses of schools to NCLB differ from responses to North Carolina’s growth standards, though using a different identification approach, as discussed above.

⁶This includes schools that do not they have sufficient number of students of a given subgroup to count toward AYP.
subgroup to meet its proficiency target if the proficiency rate increases by 10% relative to the prior year and shows a 0.1 percentage point increase in attendance rates up to the 90% threshold for that subgroup. Second, a 95% confidence interval is applied for each subgroup beginning in 2004, to help account for error from small sample sizes. We discuss precisely how this confidence interval is calculated in Section 4. Importantly, confidence intervals cannot be combined with safe harbor.⁷

Table 2 describes the percentage of schools that make AYP in a given year and subject. While the percentage passing varies considerably across time, the smallest value is 45% in 2007 and the largest is 81% in 2004. The high pass rate in 2004 coincides with the introduction of confidence intervals. The dip in the following year coincides with an increase in proficiency targets from 68.9 to 76.7 in reading and from 74.6 to 81 in math. Prior to 2006, when the math test was reformulated, schools on average were more likely to fail AYP in reading rather than math. In 2006, this switched, with schools more likely to fail math than reading.

These data describe not only whether a school met AYP in reading and/or math, but how they passed AYP, as shown in Tables 3 and 4. We focus on the reading statistics, though similar patterns hold for math. A school must have at least 40 students in a given subgroup for that subgroup to be counted toward AYP, ⁸ and the first column shows the number of schools where this is true, i.e., where the given subgroup binds. Note that the free/reduced price lunch subgroup binds in almost all school-years, 7608 out of 7954. Multi-racial binds in the least school years, only 71. After that, American Indian and Asian are the least common subgroups to bind, 230 to 210, respectively.

The second column show the percentage of schools who meet the AYP standard for the given subgroup, given that the subgroup is binding. By far, the most difficult targets to meet are for the disabled students and LEP. Only 45% of schools pass in the disabled category and 55% for LEP. By far, schools are most likely to pass in the white subgroup, just under 100% of schools. The next 4 columns further describe the schools who met AYP for a given subgroup, how the AYP requirement was satisfied, i.e, the “regular way” of meeting the target, safe harbor, confidence intervals and growth. The targets that appear to be more challenging to meet, such as LEP or disabled, are much more likely to be met through safe harbor or confidence intervals. This is not surprising as these groups

⁷These facts are taken primarily from information on the North Carolina public schools website, http://www.ncpublicschools.org/docs/nclb/abcayp/overview/aypoverview.pdf.
⁸To help minimize the strategic manipulation of the test pool by schools, as evidenced in Figlio [2006] and others, at least 95% of the students must be tested in each subgroup. These participation rates apparently apply to the entire student population, rather than only full academic year students, which are used for proficiency targets.
may both be smaller (which is the motivation for confidence intervals) and movement
toward the target (which is the motivation of safe harbor) is easier to achieve than the
target itself.

Table 5 provides a breakdown of the number of schools that face each type of sanction by year. Recalling that schools need to fail AYP two years in a row, the first year that most schools can face year 1 sanctions is in 2005. 109 (out of 1327) schools face year 1 sanctions in 2005, 71 in 2006 and 116 in 2007. By 2006, 59 schools are in year 2 sanctions and 51 in 2007.

3.2 Title I Schools

While all schools have report cards available online that include whether they met AYP standards or not, only Title I schools face sanctions for repeated failure. To be eligible to receive Title I funds, schools must have either (1) at least 35% low-income students (students receiving free/reduced price lunch) or a (2) percentage of low-income students at least as high as the district’s overall percentage, whichever is the lower of the 2. Districts rank schools by poverty and serve schools until they have no more funds. Schools with 75% or more free/reduced price lunch must be served.9

Schools can operate two kinds of Title I programs—schoolwide or targeted assistance. As the names suggest, targeted assistance programs target resources to lower-income students, whereas schoolwide programs can use Title I funds for programs that benefit the entire student population. To qualify for schoolwide Title I funds, the school needs to have at least 40% low-income students. Only one third of schools who are eligible for Title I funds are funded nationwide. Title I funds can be used for expenses such as hiring teachers to reduce class size, tutoring, computer lab, teacher assistants, parental involvement activities, purchase of materials, and pre-kindergarten programs, among others.10

Title I schools make up about 76% of our sample of elementary and middle schools in North Carolina and serve 72% of the students. The above rules suggest that designation of a Title I school may vary greatly depending on the district and how poor the school is relative to other schools in the district. Figure B compares the percentage free/reduced price lunch in Title I and non-Title I schools. There is considerable overlap in the two distributions, in particular between the ranges of 55% and 75% free/reduced price lunch.

9Federal funds are distributed to district based, generally speaking, on decennial census poverty counts.
10Information on North Carolina’s Title I funding is taken from http://www.ncpublicschools.org/program-monitoring/titleIA/.
About half of the schools between 25% and 35% free/reduced price lunch have higher poverty than the district average. It is not clear how the others qualify to receive Title I funds, though there are grandfather rules and other exemptions that allow schools to maintain their Title I status.

Table 6 describes the percentage of Title I and non-Title I schools that make AYP in a given year. In every year, a higher percentage of Title I schools pass AYP than non-Title I schools. Since Title I schools also serve the most disadvantaged students in a district, this evidence potentially supports the intuition that non-Title I schools do not respond as much to AYP standards, perhaps because they do not face any sanctions for failing. Thus, non-Title I schools may provide a good control group for Title I schools, in the sense that they are facing state growth standards but not sanctions, as discussed below. They also could help control for the stigma effect of failing AYP in the absence of sanctions.

3.3 Student data

We combine school report card data with longitudinal administrative data on student performance in North Carolina public schools. The student level data follows students as long as they attend North Carolina public schools. The analysis focuses on the period after NCLB, 2002/03 to 2006/07, and elementary and middle schools, grades 3 to 8, when students are required to take end of grade exams in both reading and mathematics. These tests are designed to measure their grade level proficiency. We standardize test scores to have mean 0 and standard deviation 1 for each grade-year. Along with the raw achievement measures, North Carolina defines achievement levels that are increasing in achievement. Students performing at achievement levels 3 or 4 are designated as proficient, whereas students performing at achievement levels 1 or 2 do not meet proficiency standards.

Each student record is further linked to a grade within an identifiable school in an identifiable district, which we use to link to the school report card data. Included in the student data are background characteristics, such as race, sex, and parental education. Column 1 of Table 7 presents summary statistics for our sample at the student level. Overall, approximately 12% of students in reading and 14% in math do not meet the cutoff for passing, scoring below achievement level 3. 48% of the students receive free or reduced price lunch. Black students make up 30% of our sample. Hispanics, the next largest minority, are only 7%.

\footnote{When there are discrepancies in the student’s reported race over time, we take the most frequently reported value.}
4 Identification and Estimation

Let $Y_{igst}$ denote achievement on either math or reading for student $i$ in grade $g$ in school $s$ and academic year $t$. Let $X_{it}$ denote individual characteristics and $Z_{st}$ the characteristics of the school. Let $FailAYP_{st-1}$ denote whether a school failed to make AYP in $t-1$. We begin our analysis by restricting our sample to Title I schools and focusing on their response to failing AYP, as they are the schools that are held accountable for the federal standards. The basic specification would estimate the effect of sanctions on achievement as

$$Y_{igst} = \beta_0 + \beta_1 Y_{i,g-1,s-1,t-1} + \beta_2 FailAYP_{s,t-1} + X_{it}\beta_3 + Z_{st}\beta_4 + \alpha_g + \alpha_t + \alpha_s + \xi_{ist},$$

(1)

where $\alpha_g$, $\alpha_t$, $\alpha_s$ denote time, grade and school fixed effects.

Schools that fail AYP are likely to differ from schools that pass in unobservable ways, suggesting that a simple OLS regression would be biased toward finding a negative effect of failing AYP. We address this first by including time-varying school observable composition measures and the time-invariant unobservable school fixed effects. Identification in the baseline specification then relies on variation within schools over time. Year and grade fixed effects help control for normal variation in achievement over time and across grades.

However, even with school fixed effect, remaining time-varying unobservables, such as changes in North Carolina policies that differentially affect lower-performing schools, may bias our estimates of the effect of failing AYP. This helps to motivate our regression discontinuity approach. Unobservable factors affecting student test scores, such as teacher or school quality or mean reversion of student performance, are likely continuously related to the school’s prior performance. However, the status of failing AYP (and potentially associated sanctions) are discontinuously applied to schools that do not have enough proficient students in each binding subgroup. By comparing schools around the threshold for passing, we exploit the logic that test scores are measured with error, so schools around the threshold for passing are equivalent except for the shock of failing AYP.

Applying a regression discontinuity design to this setting is not immediately straightforward because AYP status is determined by a number of factors, as discussed above, and so is a multidimensional problem. Prior studies, such as Ahn and Vigdor [2009] and Bacolod et al. [2009], have used the minimum distance from passing across subgroups. We collapse the AYP determination into a single index based roughly on the number of
students who would need to pass on order to meet the target. We prefer this measure because a school that fails by 1 student in 10 subgroups would face different pressures than a school that fails by 1 student in 1 subgroup, and these types of schools would be treated equivalently under previous methods.

We begin by constructing the number of students in each binding subgroup/subject combination that must test as proficient to reach the school’s target percentage proficient. To find the target percentage proficient for each subgroup/subject combination within a school, we need to account for the confidence interval and safe harbor rules. Let \( j \) denote subgroups and \( k \) subjects (reading or math). \( P_{jkst} \) is the percent of students in subgroup \( j \) and subject \( k \) in school \( s \) in year \( t \) who test as proficient. \( T_k \) is the target proficiency level set by the state for subject \( k \), and \( C_{jkst} \) is the appropriately defined one-sided 95% confidence interval. A school \( s \) is deemed to make its proficiency target in year \( t \) for subgroup \( j \) in subject \( k \) if

\[
P_{jkst} \geq \bar{P}_{jkst}
\]

where

\[
\bar{P}_{jkst} = \min \{ T_k, T_k - C_{jkst}, P_{jkst-1} + 0.1 \cdot (1 - P_{jkst-1}) \}.
\]

This expression shows the three possible ways for a school to make its proficiency target - the “regular” way, where \( P_{jkst} \geq T_k \); the “confidence interval” way, where \( P_{jkst} \geq T_k - C_{jkst} \); and the “safe harbor” way, where \( P_{jkst} \geq P_{jkst-1} + 0.1 \cdot (1 - P_{jkst-1}) \).\(^{12}\)

As discussed above, Tables 3 and 4 show the number of schools for which each subgroup binds, as well as the percentage of these subgroups that pass in each way.

We then determine the number of students who must test as proficient in order for a school to meet its target in a given subgroup/subject combination. This is given by

\[
TP_{jkst} = \bar{P}_{jkst} \cdot N_{jkst}
\]

where \( N_{jkst} \) is the number of full academic year students. We analogously find the actual number of students deemed proficient as

\[
TP_{jkst} = P_{jkst} \cdot N_{jkst}.
\]

We then aggregate across subgroups and subjects to create a single measure of the

\(^{12}\text{There is also a fourth way schools can pass their proficiency target. Students new to the North Carolina school system who do not test as proficient can be counted as proficient if they are on track to become proficient within 4 years of entry. This is known as the “growth standard” method of passing. Since these students are simply included in the count of proficient students, this method does not create a separate target.}\)
distance each school is from passing.

\[ TestPasses_{st} = \begin{cases} 
\min_{jk} \{ TP_{jkst} - TP_{jkst} \} & \text{if } TP_{jkst} \geq TP_{jkst} \forall j, k \\
\sum_{jk} 1 \cdot \{ TP_{jkst} < TP_{jkst} \} (TP_{jkst} - TP_{jkst}) & \text{otherwise}
\end{cases} \]

where \( 1 \cdot \{ TP_{jkst} < TP_{jkst} \} \) is an indicator function. Since a school must have more than its target number of proficient students in every binding subgroup/subject combination in order to make AYP, \( TestPasses_{st} \) is weakly positive when a school passes AYP and negative otherwise. As constructed, \( TestPasses_{st} \) is the minimum number of test passes that could change the AYP status of a school. For a school with a positive value of \( TestPasses_{st} \), this number represents the number of test passes a school has above the target number in the subgroup/subject combination closest to the target. When \( TestPasses_{st} \) is negative, it represents the sum of the test passes a school needs to reach its targets in every category it is currently failing. Thus, \( TestPasses_{st} \) is our measure of how far a away a school is from making AYP.

Under this measure, a discontinuity in whether or not a school makes AYP exists where \( TestPasses_{st} \) is 0. A school with a zero or positive value of \( TestPasses_{st} \) has a very high probability of making AYP, while a school with a negative value has a small probability of making AYP. In reality, all schools with \( TestPasses_{st} \geq 0 \) should make AYP, and all schools with \( TestPasses_{st} < 0 \) should not make AYP. Figure 4 shows how the probability of a school making AYP changes discontinuously as \( TestPasses_{st} \) becomes non-negative, along with a quadratic separately fitted to the points on each side of the discontinuity. While there is a significant discontinuity at \( TestPasses_{st} = 0 \), our measure does not perfectly predict AYP. This is due to either measurement error in the variables used to construct \( TestPasses_{st} \) (such as not knowing the confidence interval cutoff precisely) or a failure to precisely account for the rules governing the AYP system. To the best of our knowledge, there is no reason to suspect that the errors in predicting AYP status are systematic in a way that would bias our estimates. For this reason, a fuzzy regression discontinuity design would be more appropriate for our context.\(^{13}\)

Thus, to estimate the effect of failing AYP using our regression discontinuity approach, we instrument for \( FailAYP_{st-1} \) with our prediction of whether the school failed AYP, i.e., and indicator \( 1 \{ TestPasses_{st-1} < 0 \} \). We also account for the distance from passing by including \( Testpasses_{st-1} \) as a control variable in the regression. Formally,

\(^{13}\)See Hahn et al. [2001] and Imbens and Lemieux [2008].
equation (1) becomes

\[ Y_{ist} = \beta_0 + \beta_1 Y_{i,g-1,s-1,t-1} + \beta_2 \text{FailAYP}_{s,t-1} + X_{it}\beta_3 + Z_{st}\beta_4 \\
+ \beta_5 \text{Testpasses}_{st-1} + \alpha_g + \alpha_t + \alpha_s + \xi_{ist}. \] (2)

We estimate this using local linear regression and a uniform kernel, restricting the sample to Title I schools that are progressively closer to the threshold for passing.

We test whether schools just below and just above the threshold for passing are similar in observables. Figures 5, 6, 7 show that there are no jumps in various control variables around the threshold for passing, including the percentage of nonwhite, free/reduced price lunch students or the number of subgroups counting toward AYP. As we see no jumps in observables, this lends some support that there are no disparities in unobservables around the threshold that might bias our estimates of the effect of failing AYP.

While we begin by considering the effect of failing AYP on average student achievement, as specified in equation (3), a perhaps even more interesting question is whether school reallocate their attention among students as a result of failing AYP. In particular, we might expect that marginal students, as evidenced in Neal and Schanzenbach [2010] and elsewhere, may receive more attention as a result of failing AYP relative to other students in the school. Let \( AchLev_{i,g,s,t-1} \) be a vector of indicators for whether a student is at achievement level 1, 2, or 3 in the prior year. Recall that students need to score at or above achievement level 3 to be proficient in the subject. We might expect that schools have most incentive to focus attention on students at achievement level 2, as these students score just below the threshold for passing, and thus achievement level 1 students (the lowest performers) might be left behind. We will test whether there is evidence of this by estimating a specification that interacts the effect of failing AYP with the prior achievement level of the student.

We then consider whether estimated effects differ by sanction status, as follows

\[ Y_{igst} = \alpha_0 + \alpha_1 Y_{i,g-1,s-1,t-1} + \alpha_2 \text{FailAYP}_{s,t-1} + \alpha_3 \text{Sanction1}_{st} + \alpha_4 \text{Sanction2}_{st} \\
+ X_{it}\alpha_5 + Z_{st}\alpha_6 + \alpha_7 \text{Testpasses}_{st-1} + \alpha_g + \alpha_t + \alpha_s + \xi_{ist}, \] (3)

where \( \text{FailAYP}_{s,t-1} \) denotes a school that failed AYP in \( t-1 \) but does not face sanctions, \( \text{Sanction1}_{st} \) denotes whether the schools is in Year 1 sanctions and \( \text{Sanction1}_{st} \) Year 2 (or higher) sanctions. We also consider how these effects vary by achievement level.

If we see a different effect of sanctions relative to failing AYP absent sanctions, this
may provide evidence that sanctions have a direct effect on schools. We explore this further by comparing how the “effect” of sanctions varies by the intensity of treatment, as discussed in Section 6.

4.1 Difference-in-Differences

A potential shortcoming of the regression discontinuity is that it may either understate or overstate the average effect of failing AYP. It may understate the effect if schools just above the threshold exert additional effort to keep from failing in subsequent years. This may be particularly true for estimating the effect of failing AYP when schools do not face sanctions. However, if sanctions create an additional effect on failing schools’ ability to meet AYP, the estimate could be closer to the effect of sanctions net of the incentive to avoid failing (given it is comparable on either side of the threshold). On the other hand, the regression discontinuity could overstate the effect of failing relative to the average failing school, if schools on the margin are better able to meet AYP standards in subsequent years.

Fortunately, we have an alternative estimator that allows us to consider the effects for the average treated school. To do this, we exploit the fact that AYP status is reported for all schools in North Carolina, but only Title I schools face sanctions. Figure 3 shows the raw density of reading and math achievement for the subset of Title I schools who face year 1 sanctions and non-Title I schools who would face year 1 sanctions given their repeated failure of AYP prior to year 1 sanctions. Even without conditioning on schools fixed effects and other controls, the achievement densities look fairly similar, though it is true the non-Title I schools are generally higher achieving than Title I schools.

If the treatment effect of sanctions is constant, we are effectively over-identified with our difference-in-difference and regression discontinuity identification strategies. Figure 8 shows how the gains in achievement vary with $TestPasses_{st-1}$ for the subsample of Title I schools who face year 1 sanctions in $t$ and non-Title I schools who would have faced year 1 sanctions given their record of failing AYP. The gains are conditional on percentage free/reduced price lunch and percentage nonwhite to attempt to remove some of the heterogeneity that we control for in our regressions. We see much larger jumps at the threshold for passing in both reading and math for Title I schools. We also observe a smaller jump for non-Title I schools, but in neither case is this statistically significantly different from 0. We also observe that schools below the threshold make larger gains than schools above the threshold, consistent with a positive response of schools to year 1 sanctions. Furthermore, the responsiveness of schools decreases for schools further below
the threshold for passing, thus suggesting larger responses for marginal schools. For reading, the trends with respect to distance from passing look fairly similar for Title I and non-Title I schools, which provides more support for the validity of using non-Title I schools as a control. For math, the gains look slightly steeper for non-Title I schools, which may lead to some concern. However, it is important to recall that the regression includes other controls, such as school fixed effects, which may pick up some of this remaining heterogeneity.

We re-estimate equation (3) as a difference-in-difference estimator as follows

$$Y_{igt} = \beta_0 + \beta_1 Y_{i,g-1,s-1,t-1} + \beta_2 \text{FailAYP}_{s,t-1} + \beta_3 \text{FailAYP}_{st-1} \times \text{TitleI}_{st}$$

$$+ \beta_4 \text{TitleI}_{st} + X_{it}\beta_0 + Z_{st}\beta_0 + \alpha_g + \alpha_t + \alpha_s + \xi_{ist}, \quad (4)$$

where $\text{TitleI}_{st}$ is an indicator for whether the school is a Title I school. Our parameter of interest is $\beta_3$. It captures the average additional effect of failing AYP on Title I schools. The level effect of failing AYP ($\beta_2$) controls for time-varying factors that may be correlated with AYP failure. It also controls for stigma associated with failure in non-Title I schools.

5 Results

5.1 Failing AYP

Tables 8 and 9 present the results for the effect of failing AYP in reading and math respectively. Column (1) reports results for the whole sample of schools. The estimated effect of failing AYP is 0.02 in reading and 0.03 in math. While the regressions do control for unobservable time-invariant characteristics that may affect the schools’ ability to make AYP, a remaining concern is time-invariant unobservables which may be correlated with AYP status and also affect the school’s ability to produce student gains. When we restrict schools to be within 5 \(\text{Testpasses}_{st-1}\) from the threshold for passing, we find a larger positive effect in schools failing AYP in reading (0.05) and a weakly negative effect in math (-0.01), though neither is statistically significantly different from 0. There is a tradeoff between restricting the sample to schools closer to the threshold, which will be more similar in terms of unobservables and statistical power. When we expand to within 10 \(\text{Testpasses}_{st-1}\) from the threshold for passing, the effect for reading and math are both larger and statistically significantly different from 0, 0.06 in reading and 0.05 in math. We also estimate the effect for schools within 15 \(\text{Testpasses}_{st-1}\), and here the
effects are only slightly larger than for the whole sample in reading and not statistically significant in math. However, it is worth noting that this sample is about two-thirds of the size of the whole sample.

Table 10 considers how the effects of failing AYP vary based on the prior achievement level of the student in the given subject. On the one hand, a potential set of reforms for meeting AYP standards could involve teaching to the test or other changes in curriculum, which could lead to improvement for all students. On the other hand, it could be that schools respond by shifting resources to lower achievers at the expense of higher achievers. Or, they may focus even more particularly on students at achievement level 2 (students just below the threshold) at the expense of the lowest achievers, achievement level 1 students. Panel A reports findings for reading and Panel B for math. In both cases, for the whole sample (Column 1) the effect of failing AYP is positive for the excluded category (achievement level 4 or the highest achievers) 0.04 in reading and 0.05 in math. Achievement level 3 students benefit less in reading and math, -0.02 and -0.04. For math, this pretty much washes out the positive effect for those just above the threshold. In math, it appears that students below the threshold have comparable gains after failing AYP as the highest achievers (i.e., it is only achievement level 3 students that do not benefit). For reading, both achievement level 1 and 2 students gain less relative to the highest achievers, with the lowest achievers apparently experiencing losses of about 0.04 of a standard deviation.

The question is whether these distributional effects are driven by time-varying unobservables. Columns (2), (3) and (4) consider our regression discontinuity results within 5, 10 and 15 \( TestPasses_{t-1} \) from the threshold for passing respectively. Now, we see some evidence in both reading and math in the restricted samples, that the lowest-achievers (achievement level 1 students) experience losses when schools fail AYP, at least relative to the highest achievers and sometimes net losses. The findings for achievement level 2 and 3 students (those below and above the threshold) cannot reject that these students have similar gains to the highest achievers. Thus, overall the findings from the regression discontinuity do suggest that the response to failing AYP may be deriving at the expense of the lowest achievers, but achievement level 2, 3 and 4 students experience comparable gains.

5.2 Difference-in-difference

Table 11 compares findings for the effect of failing AYP when we use our difference-in-difference approach. Column (1) and (4) present the average effect of failing AYP for
reading and math, respectively. Only the interactions with Title I schools are reported, i.e., $\beta_3$ in equation (4). The estimated effect of failing AYP is 0.01 in reading and 0.02 in math. These estimates support the intuition that the regression discontinuity may overstate the results of failing AYP by focusing on marginal schools. They are also somewhat smaller (at least for reading) than estimates on the whole sample of Title I schools used in the fixed effect regression above. This supports the intuition that non-Title I schools may help control for unobservable factors that are not picked up with just school fixed effects.

Columns (2) and (5) estimate the results broken out by a student’s prior achievement level, as in Table 10. The distributional effects look quite different than what we predict with our regression discontinuity approach. For reading, we still find the negative effect on the lowest achievers, but much larger in magnitude -0.10. There is a small positive effect for students around the threshold, with slight larger positive effects on students just below 0.03 for achievement level 2 compared to 0.02 for achievement level 3. There is no effect of the highest achievers. For math, there is a small positive effect on the highest achievers (0.02), which also extends to achievement level 2 and 3 students, but is smaller in magnitude than the regression discontinuity results. However, the big difference is that the lowest achievers gain even more, 0.06 of a standard deviation, relative to the other students. If non-Title I schools provide an adequate control group, this disparity in distributional effects could derive through differences in the composition of students across marginal schools and the average failing Title I school. Future work will continue to investigate this disparity.

We also want to consider how students in failing subgroups fare relative to students in non-failing subgroups. In some cases, we might expect that schools reallocate resources disproportionately to failing subgroups, potentially at the expense of students in non-failing subgroups. This seems particularly feasible in the case of schools failing in limited English proficient or disabled categories, where the type of instruction might be targeted to these students at the exclusion of other subgroups. Columns (3) and (6) estimate these effect of failing AYP by whether the student is in a failing subgroup. The effect of non-failing subgroups is 0 across reading and math and 0.02 and 0.03 respectively for being in a failing subgroup for math and reading. This gives some evidence that schools are targeting resources toward failing subgroups. We plan to explore this also in the regression discontinuity framework. However, because failing subgroups are only defined for schools below the threshold for passing, this type of comparison is less straightforward in the regression discontinuity framework.
5.3 Effect of Sanctions

As discussed above, failing schools face increasing incentives to meet AYP with repeated failure. In particular, we want to see whether schools respond differently to being under sanctions versus failing AYP absent sanctions. Table 12 considers effects separately for schools facing Year 1 sanctions (school choice), Year 2 sanctions (choice or extra tutoring for failing students) or if they fail AYP absent sanctions. As discussed above, while an argument could be made that schools face similar pressures to pass AYP above and below the threshold, sanctions clearly discontinuously affect failing schools.

Column 1 considers first the whole sample of Title I schools and findings a positive effect on growth in student performance for schools failing AYP absent sanction, 0.02 in reading and 0.03 in math. However, the effects appear to be much larger when schools face sanctions. The estimated effects of Year 1 sanctions are 0.04 in reading and 0.09 in math. For Year 2 sanctions, the effects are 0.06 in reading and 0.09 in math.

We see similar patterns when we restrict the schools to be within 5, 10 and 15 TestPasses_{st−1} from the threshold. As in findings for the overall effect of failing AYP (in Table 8 and 9), the effects are larger when we focus on schools around the threshold for passing. Focusing on Column (3) that restricts the sample to be within 10 test passes, we find significant effects of failing AYP absent sanctions, 0.08 for reading and 0.10 for math. For schools facing year 1 sanctions, the effect is 0.09 in reading and 0.14 in math. The effect of year 2 sanctions are similar in magnitude to year 1, though not statistically significantly different from 0, probably due to small sample sizes. Interestingly, though there is still some evidence of a larger effect of sanctions relative to failing AYP absent sanctions, it is weaker within narrower bandwidths relative to the whole sample and we could not reject that the coefficients are the same in the restricted sample.

It is also interesting to consider whether the distributional effects across students’ prior achievement levels vary by the level of sanctions. The distributional effect of failing AYP absent sanctions could be quite different than the effect of sanctions. For instance, the effect of the Year 1 sanction of school choice might depend on who is choosing (or threatening) to exercise choice. Failing students get priority, but even higher achievers can choose to leave. Year 2 sanctions offer the potential of extra tutoring services that are targeted to failing students. This could help low-achievers, potentially at the expense of high-achievers if resources are redistributed away from them to pay for the tutoring.

Table 13 considers how distributional effects for failing schools vary by sanction status and prior achievement level of the student. Focusing on Column (3) and (7), which present results for reading and math for schools within 10 test passes bandwidth, we find that schools failing AYP absent sanctions show positive effects of 0.08 in math and...
0.09 in reading for the highest achievers. There is some evidence that achievement level 3 students benefit slightly more in math (relative to the high achievers) by 0.02. The effect for achievement level 2 students is not statistically significantly different from the positive effect on the highest achievers. The lowest achievers appear to lose out in math (-0.10) but not in reading. Recall that our findings for the overall effect of failing AYP in Table 10 found that if anything the lowest achievers experienced losses relative to higher-achieving students.

Turning to the results for year 1 sanctions, the effect for the highest achievers is slightly larger in reading (0.10) than for failing AYP absent sanctions (0.08) and considerably larger in math (0.17 compared to 0.09). Under Year 1 sanctions, we now find that the lowest achievers experience relative losses in reading of -0.16, whereas the achievement level 2 and 3 students appear experience gains that are not statistically significantly different from the highest achievers. This mirrors the findings of the overall effect of failing AYP. However, the results for math are more difficult to interpret. The point estimates suggest that the lowest achievers may experience some relative losses, though the estimates are not statistically significantly different from 0. Achievement level 2 students experience comparable gains to the highest achievers. However, the odd finding is that achievement level 3 students also experience relative losses of -0.07. This finding is robust across the different bandwidths. Thus, schools appear to be preserving the math performance of the highest achievers and students just below the threshold, while sacrificing students just above and students far below. Section 6 considers who is exercising choice in response to Year 1 sanctions, which may shed some light on these findings.

The distributional effects for Year 2 sanctions look fairly different. While again there is a positive effect on the highest achievers (0.18 in reading and 0.14 in math), all student at all other achievement levels experience smaller gains relative to the high achievers. Students below the threshold experience sizable relative losses of 0.12 for achievement level 1 in reading and -0.18 for achievement level 2. Thus, the overall net effect for students below the threshold is approximately 0. For math, achievement level 1 students experience larger relative losses of -0.15 leading to a 0 net effect. Achievement level 2 students experience smaller relative losses of -0.09, leaving a net gain of 0.05 in math. Achievement level 3 students experience about half the gains of achievement level 4 students (0.08 for reading and 0.07 for math). Because Year 2 sanctions target resources toward low-achievers through extra tutoring, these results are particularly surprising. It could be that the extra tutoring services are not helping as much as prior initiatives to help students below the threshold. This is consistent with evidence of low levels of
take-up among failing students and low levels of effectiveness, as shown in Zimmer et al. [2007]. However, students also have the potential to exercise choice in response to year 2 sanctions. It is more difficult to determine the potential distributional effects of choice.

6 Mechanisms

Overall, these findings provide some evidence that the response of failing schools varies by sanction status. There are several potential reasons this could be the case.

First, there could simply be time-varying treatment effects of failing AYP. It could be that schools that progressively fail to meet AYP work progressively harder to pass, leading to different effects from repeated failure. Second, the treatment effects may differ across sanction statuses simply because of heterogeneity in responses of schools to sanctions. For instance, schools that face year 2 sanctions during our sample period are likely to be worse (by the AYP metric at least) in terms of their response to failure than schools who only face year 1 sanctions or schools that fail AYP but do not face sanctions. These schools may be less equipped to respond to sanctions and thus have different responses to sanctions than the average failing school. Our regression discontinuity results do present some evidence of this sort of heterogeneity.

Third, the treatment effects may differ because of a direct effect of sanctions, i.e., school choice or extra tutoring services drive the positive results. We can test this hypothesis for school choice first by seeing who exercises choice and then by comparing results by the intensity of treatment, i.e., how many choices are available in our context. We explore school choice as a response to year 1 and 2 sanction statuses directly in Section 6.1. Section 6.2 considers how treatment effects of sanctions vary by the degree of choice available. Because we lack data on the take-up of supplemental education services, we cannot examine this directly as an effect of year 2 sanctions.

6.1 Choice

Our evidence suggests that year 1 sanctions have a positive effect on student performance. While all students have the opportunity to exercise choice in sanctioned schools, the lowest performing and highest poverty children have first priority for transfer options and free transportation when funding is limited. Thus, one reason that sanctioned schools improve could be that they are losing their lowest-achieving students or being held accountable for fewer subgroups. Parents must be notified the summer after a school fails to make AYP and will face sanctions in $t + 1$ and be given the option of sending
their child to another school in the district that is not subject to sanctions, if available.14

We can directly examine whether students in sanctioned schools exercise choice and which students leave. Let $Choice_{st+1}$ indicate whether a student attends a different (non-sanctioned) school in the same district in year $t + 1$. We model how choice varies with sanctions as

$$Choice_{st+1} = \gamma_0 + Sanction_{st+1}\gamma_1 + Sanction_{st+1} \times TitleI_{st} \gamma_2 + X_{it} \gamma_3 + Z_{st} \gamma_4 + \delta_g + \delta_t + \delta_s + \epsilon_{st+1}. \quad (5)$$

Similarly to above, using non-Title I schools that would have faced sanctions helps control for any differential mobility due to failing AYP but not due to the sanction itself. Grade fixed effects help control for natural movement across grades.

The model also controls for student demographics, namely indicators for 8 of the potential subgroups that are monitored under NCLB, with white being the excluded category. School characteristics include the percentages of students belonging to each of these subgroups. Schools likely differ also in unobservable ways that are correlated with student movement, such as worse schools having more student turnover. We control for these unobservable time-invariant differences with school fixed effects. Because of the large number of fixed effects, we choose to estimate this as a linear probability model rather than as a probit or logit. Finally, year fixed effects capture changes in mobility over time which may not be related with sanctions.

$Sanction_{st+1}$ is our variable of interest and denotes whether students will face year 1, year 2+ sanctions if they stay in the same school. Recall that as sanctions escalate, districts must continue to provide choice, but parents may choose to request supplemental education services instead under year 2 sanctions and higher. Similarly to above, for non-Title I schools this captures whether these schools would face sanctions in $t + 1$ given their repeated failure in the same subject in prior years.

Column (1) of Table (14) shows some evidence that students in Title I schools are somewhat more likely to exercise choice under year 1 sanctions ($0.0057$), but not under year 2 sanctions. The effect of sanctions on school choice should also depend on the number of options available in the district. Column (2) compares interacts the effect of sanctions with the number of choices (i.e., number of non-sanctioned school/grades available to the student) in the district. We find that the effect of both year 1 and year 2 sanctions increases in the number of choices available to the student. Interestingly, we find a positive coefficient on year 1 sanctions when there are 0 choices available (the level term). This may result from moves to another district, which could include charter

---

14Capacity constraints are not a valid excuse for the district not to provide options.
schools. In cases where no options are available, by our measure, districts are required to provide options by reaching agreement with a neighboring district. For year 2 sanctions, there is evidence that students are less likely to change schools when there are no options available, by our measure. This discrepancy may be reconciled if these types of moves out of the system are more likely to occur early.

The effects of sanctions on school choice are quite small in magnitude and fairly comparable across sanction statuses. With 10 choices in the district, students in Title I schools under year 1 or 2 sanctions are about 1% more likely to change schools. With 20 choices, they are about 2% more likely to change schools.

Given that some students are changing schools in response to sanctions, an even more interesting question is who is moving. We find (not reported) that students who performed below the threshold in either reading or math are less likely to exercise choice under sanctions than students above the threshold. However, some of this is due to different numbers of options available to the students. When we account for the number of options available, we find that students below the threshold are in fact more likely to exercise choice under year 1 sanctions, though the magnitudes are quite small. Even accounting for number of options, there is no evidence that the students below the threshold are more likely to exercise choice than higher achieving students under year 2 sanctions. This may make sense, as extra tutoring services are available only for these students and they may opt to take up the extra tutoring rather than changing schools in year 2 of sanctions.

Together these results suggest that to some extent students are responding to sanctions by exercising choice, and higher-achieving students are leaving sanctioned schools on average more than low-achievers. Part of the relatively large positive effect we see for high achievers under Year 1 or 2 sanctions may then be efforts of failing schools not to lose their highest achievers.

These findings also suggest that our previous results may actually understate the response of schools to sanctions on average, given exit of higher achievers. To determine whether exit and changes in student composition are a meaningful bias downward on our estimates of the effect of sanctions, we re-estimate equation (3) in the restricted sample of students who do not appear to change schools in response to sanctions. The findings regarding the effect of sanctions do not change significantly, in part because movers comprise such a small percentage of the estimation sample, a little over 4%.
6.2 Effects by Intensity of Treatment

While the evidence above shows that the degree of choice in response to sanctions is fairly small in magnitude, we might still expect that the response of schools to sanctions is larger in districts with more choice (if in fact the choice sanction is having a direct effect on schools). In fact, the small degree of choice could be a result of effective responses of sanctioned schools that help them to keep students despite their repeated failure to make AYP.

To determine whether choice sanctions have a direct effect on schools, Table 15 interacts sanctions with the number of options available to see if there are larger effects in districts with more choices. For reading, there is some evidence that the effect is increasing in the number of choices, particularly for year 1 sanctions. For math, this does not appear to be the case. Thus, the results present mixed findings across subjects for an effect of sanctions increasing in the number of choices, which makes it difficult to interpret. Future work will also consider whether the effect differs across achievement levels and whether there are non-linearities.

To consider whether the effects differ for the average failing school, we also consider whether there is evidence of differential effects by intensity of treatment using our difference-in-difference strategy. In particular, Table 16 considers whether the response to sanctions increases in the number of options available in the districts. Columns (1) and (3) consider the average effects on reading and math of being a Title I school under sanctions in a district with more choice. While there is no effect for reading, we find weak evidence that the effect of year 1 sanctions is increasing in math with the number of options available, by 0.005. Students in districts with 0 options still improve achievement by 0.0155 on average. When there are 10 options available, the gains are 0.02 of a standard deviation, and 0.026 with 20 options available.\footnote{The median number of choices is 9; the 75\textsuperscript{th} percentile is 18.}

We find more evidence that the response of schools to year 2 sanctions varies by the degree of choice. This could result if schools observe parents exiting when they face year 1 sanctions and try to respond to prevent exit under year 2 sanctions. For reading, the point estimate is not statistically different from 0 when no options are available, but the effect interacted with number of choices is 0.0014. For math, the effect with 0 choices is actually negative, -0.026, but positive as choices are available, 0.0026. With 10 choices, this translates into small gains of 0.006 in reading and small losses of -0.005 in math. With 20 choices, this means larger gains of 0.02 in reading and 0.016 in math. This provides some evidence that the estimated effect varies with the degree of choice, and
may therefore be a direct result of sanctions. The fact that the response of failing AYP absent sanctions does not vary with the degree of choice also supports this hypothesis.

Columns (2) and (4) break out results by whether the student is a failing student in a failing subgroup \((\text{TargStud})\). With 10 options for the student in the district, Title I schools facing year 1 sanctions improve achievement of target students by 0.03 in reading and 0.05 in math. With 20 options, the gains are 0.05 of a standard deviation in reading and math. These gains are more significant in magnitude. Interestingly, with year 1 sanctions the gains are weakly decreasing in math by the degree of choice for target students.

For year 2 sanctions, reading gains increase relatively more for target students with the number of options available, whereas there is no evidence of an additional effect of being a target student in math. With 10 options, target students in Title I schools gain 0.06 in reading and approximately 0 in math. With 20 options, target students gain 0.08 of a standard deviation in reading and 0.03 in math.

Overall, these results suggest some substantial variation in treatment effects of sanctions by the intensity of treatment, the degree of choice available in the district. Generally, it appears that more choice leads to larger gains in student performance in Title I schools who are sanctioned, particularly for target students in reading.

7 Conclusion

We use a regression discontinuity approach to determine the effect of failing AYP on school performance. We find improvement in student performance in schools that fail AYP. These effects are generally comparable across students of different achievement levels, with exceptions of marginal losses or no gains for the lowest achieving students. For higher achieving students (including those just below the threshold for passing), the effects are fairly large in magnitude, ranging from 0.09 to 0.17 in magnitude.

We also consider whether responses differ across whether the school faces sanctions or fails AYP absent sanctions. We find that the gains in achievement are generally larger when schools face sanctions, relative to failing AYP absent sanctions. Furthermore, the distributional effects appear to be quite different across sanction statuses, providing some suggestive evidence of potentially a direct effect of sanctions.

To determine the potential for sanctions (particularly the choice sanction) to have a direct effect, we also consider the degree of choice exercised in sanctioned schools. We find that only a little over 4% of students in sanctioned schools exercise choice. If anything, it appears that lower-achievers are generally not more likely to take advantage
of choice than higher achievers. However, we do find some evidence that the effect of sanctions increases substantially with the number of choices available, suggesting that the choice incentive may be driving some of the positive responses to sanctions that we observe in the data.
References


Tom Ahn and Jacob Vigdor. Does no child left behind have teeth? examining the impact of federal accountability sanctions in north carolina. Working Paper, October 2009.


## A Tables

### Table 1: School Proficiency Rates by Subgroup and Subject

<table>
<thead>
<tr>
<th></th>
<th>Reading Mean</th>
<th>Reading SD</th>
<th>Math Mean</th>
<th>Math SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Students</td>
<td>0.7886</td>
<td>0.1471</td>
<td>0.7744</td>
<td>0.1566</td>
</tr>
<tr>
<td>White</td>
<td>0.8642</td>
<td>0.1052</td>
<td>0.8512</td>
<td>0.1102</td>
</tr>
<tr>
<td>Black</td>
<td>0.6916</td>
<td>0.1685</td>
<td>0.6535</td>
<td>0.1913</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.6447</td>
<td>0.1811</td>
<td>0.6679</td>
<td>0.1601</td>
</tr>
<tr>
<td>Am. Ind.</td>
<td>0.6882</td>
<td>0.171</td>
<td>0.6808</td>
<td>0.1879</td>
</tr>
<tr>
<td>Asian</td>
<td>0.85</td>
<td>0.1333</td>
<td>0.8923</td>
<td>0.0898</td>
</tr>
<tr>
<td>Multi</td>
<td>0.7732</td>
<td>0.1675</td>
<td>0.7363</td>
<td>0.1187</td>
</tr>
<tr>
<td>FRL</td>
<td>0.7106</td>
<td>0.1589</td>
<td>0.698</td>
<td>0.1725</td>
</tr>
<tr>
<td>LEP</td>
<td>0.5159</td>
<td>0.2031</td>
<td>0.5646</td>
<td>0.1910</td>
</tr>
<tr>
<td>Disabled</td>
<td>0.5009</td>
<td>0.1876</td>
<td>0.5184</td>
<td>0.2026</td>
</tr>
</tbody>
</table>

Data Source: school report cards Average proficiency rates for schools from 2002/03 to 2006/07.

### Table 2: Percentage of Schools Passing AYP

<table>
<thead>
<tr>
<th>Year</th>
<th>Overall</th>
<th>Reading</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>0.5706</td>
<td>0.6036</td>
<td>0.7380</td>
</tr>
<tr>
<td>2004</td>
<td>0.8080</td>
<td>0.8259</td>
<td>0.9029</td>
</tr>
<tr>
<td>2005</td>
<td>0.6652</td>
<td>0.7135</td>
<td>0.7764</td>
</tr>
<tr>
<td>2006</td>
<td>0.4830</td>
<td>0.6509</td>
<td>0.5928</td>
</tr>
<tr>
<td>2007</td>
<td>0.4508</td>
<td>0.7221</td>
<td>0.4880</td>
</tr>
</tbody>
</table>

Sample: Title I Schools; Source: school report cards
### Table 3: School AYP Status, by Subgroup: Reading

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>No. with Binding</th>
<th>% Passing</th>
<th>Out of schools who pass that subgroup: % Regular</th>
<th>% Safe</th>
<th>%CI</th>
<th>% Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Students</td>
<td>7954</td>
<td>0.9296</td>
<td>0.9143</td>
<td>0.0226</td>
<td>0.0629</td>
<td>0.0003</td>
</tr>
<tr>
<td>White</td>
<td>6625</td>
<td>0.9982</td>
<td>0.9909</td>
<td>0.0002</td>
<td>0.009</td>
<td>0</td>
</tr>
<tr>
<td>Black</td>
<td>4978</td>
<td>0.7993</td>
<td>0.6869</td>
<td>0.0834</td>
<td>0.2292</td>
<td>0.0005</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1569</td>
<td>0.7247</td>
<td>0.5629</td>
<td>0.1284</td>
<td>0.3078</td>
<td>0.0009</td>
</tr>
<tr>
<td>Am. Ind.</td>
<td>230</td>
<td>0.8043</td>
<td>0.6486</td>
<td>0.0973</td>
<td>0.2541</td>
<td>0</td>
</tr>
<tr>
<td>Asian</td>
<td>210</td>
<td>0.9857</td>
<td>0.9372</td>
<td>0.0145</td>
<td>0.0483</td>
<td>0</td>
</tr>
<tr>
<td>Multi</td>
<td>71</td>
<td>1</td>
<td>0.9859</td>
<td>0</td>
<td>0.0141</td>
<td>0</td>
</tr>
<tr>
<td>FRL</td>
<td>7608</td>
<td>0.84</td>
<td>0.7534</td>
<td>0.0699</td>
<td>0.1759</td>
<td>0.0008</td>
</tr>
<tr>
<td>LEP</td>
<td>864</td>
<td>0.5439</td>
<td>0.2123</td>
<td>0.4671</td>
<td>0.3206</td>
<td>0</td>
</tr>
<tr>
<td>Disabled</td>
<td>3285</td>
<td>0.4473</td>
<td>0.287</td>
<td>0.4491</td>
<td>0.2368</td>
<td>0</td>
</tr>
</tbody>
</table>

Source: school report cards Years: 2002/03-2006/07

### Table 4: School AYP Status, by Subgroup: Math

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>No. with Binding</th>
<th>% Passing</th>
<th>Out of schools who pass that subgroup: % Regular</th>
<th>% Safe</th>
<th>%CI</th>
<th>% Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Students</td>
<td>7954</td>
<td>0.8669</td>
<td>0.7859</td>
<td>0.1594</td>
<td>0.0505</td>
<td>0.0041</td>
</tr>
<tr>
<td>White</td>
<td>6629</td>
<td>0.9652</td>
<td>0.9128</td>
<td>0.0519</td>
<td>0.0347</td>
<td>0.0006</td>
</tr>
<tr>
<td>Black</td>
<td>4985</td>
<td>0.7025</td>
<td>0.5497</td>
<td>0.3415</td>
<td>0.1059</td>
<td>0.0029</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1575</td>
<td>0.7098</td>
<td>0.5197</td>
<td>0.2987</td>
<td>0.1771</td>
<td>0.0045</td>
</tr>
<tr>
<td>Am. Ind.</td>
<td>230</td>
<td>0.7174</td>
<td>0.6</td>
<td>0.297</td>
<td>0.103</td>
<td>0</td>
</tr>
<tr>
<td>Asian</td>
<td>211</td>
<td>0.9953</td>
<td>0.9333</td>
<td>0.0429</td>
<td>0.0238</td>
<td>0</td>
</tr>
<tr>
<td>Multi</td>
<td>71</td>
<td>0.8732</td>
<td>0.5806</td>
<td>0.1129</td>
<td>0.3065</td>
<td>0</td>
</tr>
<tr>
<td>FRL</td>
<td>7615</td>
<td>0.7778</td>
<td>0.6044</td>
<td>0.2998</td>
<td>0.0917</td>
<td>0.0041</td>
</tr>
<tr>
<td>LEP</td>
<td>860</td>
<td>0.6198</td>
<td>0.3114</td>
<td>0.5009</td>
<td>0.1782</td>
<td>0.0094</td>
</tr>
<tr>
<td>Disabled</td>
<td>3301</td>
<td>0.4793</td>
<td>0.3001</td>
<td>0.5003</td>
<td>0.18</td>
<td>0.0013</td>
</tr>
</tbody>
</table>

Source: school report cards Years: 2002/03-2006/07
### Table 5: School Sanctions

<table>
<thead>
<tr>
<th>Sanction Level</th>
<th>Year 2003</th>
<th>Year 2004</th>
<th>Year 2005</th>
<th>Year 2006</th>
<th>Year 2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1326</td>
<td>1315</td>
<td>1210</td>
<td>1190</td>
<td>1104</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>12</td>
<td>109</td>
<td>71</td>
<td>116</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>59</td>
<td>51</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>47</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>N</td>
<td>1328</td>
<td>1328</td>
<td>1327</td>
<td>1327</td>
<td>1325</td>
</tr>
</tbody>
</table>

Source: school report cards

### Table 6: Percentage of Schools Passing AYP by Title I Status

<table>
<thead>
<tr>
<th>Year</th>
<th>Title I</th>
<th>Non-Title I</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>0.5706</td>
<td>0.4311</td>
</tr>
<tr>
<td>2004</td>
<td>0.8080</td>
<td>0.6597</td>
</tr>
<tr>
<td>2005</td>
<td>0.6652</td>
<td>0.5286</td>
</tr>
<tr>
<td>2006</td>
<td>0.4830</td>
<td>0.3798</td>
</tr>
<tr>
<td>2007</td>
<td>0.4508</td>
<td>0.3844</td>
</tr>
</tbody>
</table>

Source: school report cards
Table 7: Summary Statistics 2003-06 (Student Level)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Full Sample</th>
<th>Failed AYP†</th>
<th>Met AYP†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>Mean</td>
</tr>
<tr>
<td>Reading</td>
<td>0.4961</td>
<td>0.7910</td>
<td>0.3798</td>
</tr>
<tr>
<td>Ach. Level 1</td>
<td>0.0224</td>
<td>0.1478</td>
<td>0.0301</td>
</tr>
<tr>
<td>Ach. Level 2</td>
<td>0.0977</td>
<td>0.2968</td>
<td>0.1249</td>
</tr>
<tr>
<td>Ach. Level 3</td>
<td>0.4427</td>
<td>0.4967</td>
<td>0.4648</td>
</tr>
<tr>
<td>Ach. Level 4</td>
<td>0.4373</td>
<td>0.4961</td>
<td>0.3802</td>
</tr>
<tr>
<td>Math</td>
<td>0.6223</td>
<td>0.7503</td>
<td>0.4834</td>
</tr>
<tr>
<td>Ach. Level 1</td>
<td>0.0265</td>
<td>0.1606</td>
<td>0.0450</td>
</tr>
<tr>
<td>Ach. Level 2</td>
<td>0.1134</td>
<td>0.3171</td>
<td>0.1650</td>
</tr>
<tr>
<td>Ach. Level 3</td>
<td>0.3698</td>
<td>0.4828</td>
<td>0.4040</td>
</tr>
<tr>
<td>Ach. Level 4</td>
<td>0.4899</td>
<td>0.4999</td>
<td>0.3854</td>
</tr>
<tr>
<td>FRL</td>
<td>0.4799</td>
<td>0.4996</td>
<td>0.5429</td>
</tr>
<tr>
<td>Am. Indian</td>
<td>0.0169</td>
<td>0.1290</td>
<td>0.0248</td>
</tr>
<tr>
<td>Asian</td>
<td>0.0183</td>
<td>0.1340</td>
<td>0.0178</td>
</tr>
<tr>
<td>Black</td>
<td>0.2939</td>
<td>0.4555</td>
<td>0.3691</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.0664</td>
<td>0.2489</td>
<td>0.0825</td>
</tr>
<tr>
<td>Multiracial</td>
<td>0.0256</td>
<td>0.1579</td>
<td>0.0259</td>
</tr>
<tr>
<td>White</td>
<td>0.5790</td>
<td>0.4937</td>
<td>0.4799</td>
</tr>
<tr>
<td>LEP</td>
<td>0.0363</td>
<td>0.1869</td>
<td>0.0450</td>
</tr>
<tr>
<td>N</td>
<td>499322</td>
<td>182823</td>
<td>316499</td>
</tr>
</tbody>
</table>

† Summary statistics correspond to the characteristics of the school in the year that they failed or met AYP standards.
<table>
<thead>
<tr>
<th></th>
<th>Whole (1)</th>
<th>Within 5 (2)</th>
<th>Within 10 (3)</th>
<th>Within 15 (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FailAYP(_{t-1})</td>
<td>0.0188***</td>
<td>0.049</td>
<td>0.0632***</td>
<td>0.0274*</td>
</tr>
<tr>
<td></td>
<td>(0.00682)</td>
<td>(0.0404)</td>
<td>(0.0231)</td>
<td>(0.0145)</td>
</tr>
<tr>
<td>TestPasses(_{t-1})</td>
<td>-9.03e-05</td>
<td>0.000445</td>
<td>0.00256</td>
<td>0.000258</td>
</tr>
<tr>
<td></td>
<td>(0.000133)</td>
<td>(0.00382)</td>
<td>(0.00161)</td>
<td>(0.000812)</td>
</tr>
<tr>
<td>Parent HS Degree</td>
<td>0.0974***</td>
<td>0.0991***</td>
<td>0.0937***</td>
<td>0.0977***</td>
</tr>
<tr>
<td></td>
<td>(0.00477)</td>
<td>(0.00820)</td>
<td>(0.00619)</td>
<td>(0.000554)</td>
</tr>
<tr>
<td>Parent 4-year+</td>
<td>0.247***</td>
<td>0.244***</td>
<td>0.242***</td>
<td>0.246***</td>
</tr>
<tr>
<td></td>
<td>(0.00559)</td>
<td>(0.00890)</td>
<td>(0.00708)</td>
<td>(0.00635)</td>
</tr>
<tr>
<td>Nonwhite</td>
<td>-0.105***</td>
<td>-0.105***</td>
<td>-0.100***</td>
<td>-0.101***</td>
</tr>
<tr>
<td></td>
<td>(0.00325)</td>
<td>(0.00536)</td>
<td>(0.00399)</td>
<td>(0.00349)</td>
</tr>
<tr>
<td>Male</td>
<td>-0.0336***</td>
<td>-0.0298***</td>
<td>-0.0306***</td>
<td>-0.0313***</td>
</tr>
<tr>
<td></td>
<td>(0.00185)</td>
<td>(0.00321)</td>
<td>(0.00244)</td>
<td>(0.00218)</td>
</tr>
<tr>
<td>Free/reduced price lunch</td>
<td>-0.0714***</td>
<td>-0.0727***</td>
<td>-0.0731***</td>
<td>-0.0734***</td>
</tr>
<tr>
<td></td>
<td>(0.00213)</td>
<td>(0.00365)</td>
<td>(0.00275)</td>
<td>(0.00244)</td>
</tr>
<tr>
<td>Limited English Prof.</td>
<td>-0.0107*</td>
<td>0.000190</td>
<td>-0.0108</td>
<td>-0.0107</td>
</tr>
<tr>
<td></td>
<td>(0.00605)</td>
<td>(0.0102)</td>
<td>(0.00789)</td>
<td>(0.00695)</td>
</tr>
<tr>
<td>Disabled</td>
<td>-0.253***</td>
<td>-0.236***</td>
<td>-0.233***</td>
<td>-0.236***</td>
</tr>
<tr>
<td></td>
<td>(0.00527)</td>
<td>(0.00942)</td>
<td>(0.00676)</td>
<td>(0.00580)</td>
</tr>
<tr>
<td>No. of Binding Subgroups</td>
<td>-0.00148</td>
<td>0.000315</td>
<td>-0.00177</td>
<td>-0.00147</td>
</tr>
<tr>
<td></td>
<td>(0.00109)</td>
<td>(0.00227)</td>
<td>(0.00162)</td>
<td>(0.00143)</td>
</tr>
<tr>
<td>% Free/reduced price lunch</td>
<td>0.0559</td>
<td>0.0522</td>
<td>0.0796</td>
<td>0.0809</td>
</tr>
<tr>
<td></td>
<td>(0.0424)</td>
<td>(0.0876)</td>
<td>(0.0622)</td>
<td>(0.0534)</td>
</tr>
<tr>
<td>% LEP</td>
<td>0.0226</td>
<td>0.162</td>
<td>-0.132</td>
<td>-0.0666</td>
</tr>
<tr>
<td></td>
<td>(0.0916)</td>
<td>(0.210)</td>
<td>(0.155)</td>
<td>(0.120)</td>
</tr>
<tr>
<td>% White</td>
<td>0.0949</td>
<td>-0.254</td>
<td>-0.0754</td>
<td>0.00454</td>
</tr>
<tr>
<td></td>
<td>(0.0771)</td>
<td>(0.222)</td>
<td>(0.118)</td>
<td>(0.0910)</td>
</tr>
<tr>
<td>% Disabled</td>
<td>0.212***</td>
<td>-0.0224</td>
<td>0.171</td>
<td>0.181**</td>
</tr>
<tr>
<td></td>
<td>(0.0774)</td>
<td>(0.160)</td>
<td>(0.106)</td>
<td>(0.0889)</td>
</tr>
<tr>
<td>N</td>
<td>462425</td>
<td>142749</td>
<td>240697</td>
<td>317225</td>
</tr>
<tr>
<td>R(^2)</td>
<td>0.599</td>
<td>0.608</td>
<td>0.603</td>
<td>0.601</td>
</tr>
</tbody>
</table>

Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%

Regressions include controls for prior achievement, year, school, and grade fixed effects along with a constant. The years include 2004/06, with 2003 for lags. Column (2) includes observations within 5 students from the margin for passing |TestPasses\(_{st-1}\) | ≤ 5, and similarly for “within 10” and “within 15” for columns (3) and (4).
Table 9: Effect of Failing AYP: Math

<table>
<thead>
<tr>
<th></th>
<th>Whole</th>
<th>Within 5</th>
<th>Within 10</th>
<th>Within 15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>FailAYP&lt;sub&gt;t−1&lt;/sub&gt;</td>
<td>0.0268***</td>
<td>-0.00957</td>
<td>0.0546*</td>
<td>0.0205</td>
</tr>
<tr>
<td></td>
<td>(0.00825)</td>
<td>(0.0464)</td>
<td>(0.0301)</td>
<td>(0.0194)</td>
</tr>
<tr>
<td>TestPasses&lt;sub&gt;t−1&lt;/sub&gt;</td>
<td>-0.000264*</td>
<td>-0.00573</td>
<td>0.00175</td>
<td>-0.000873</td>
</tr>
<tr>
<td></td>
<td>(0.00137)</td>
<td>(0.0435)</td>
<td>(0.00205)</td>
<td>(0.00104)</td>
</tr>
<tr>
<td>Parent HS Degree</td>
<td>0.0625***</td>
<td>0.0536***</td>
<td>0.0593***</td>
<td>0.0643***</td>
</tr>
<tr>
<td></td>
<td>(0.00397)</td>
<td>(0.00708)</td>
<td>(0.00552)</td>
<td>(0.00486)</td>
</tr>
<tr>
<td>Parent 4-year+</td>
<td>0.216***</td>
<td>0.205***</td>
<td>0.213***</td>
<td>0.216***</td>
</tr>
<tr>
<td></td>
<td>(0.00489)</td>
<td>(0.00854)</td>
<td>(0.00667)</td>
<td>(0.00579)</td>
</tr>
<tr>
<td>Nonwhite</td>
<td>-0.0804***</td>
<td>-0.0736***</td>
<td>-0.0755***</td>
<td>-0.0760***</td>
</tr>
<tr>
<td></td>
<td>(0.00298)</td>
<td>(0.00506)</td>
<td>(0.00378)</td>
<td>(0.00332)</td>
</tr>
<tr>
<td>Male</td>
<td>-0.00494***</td>
<td>-0.00488</td>
<td>-0.00804***</td>
<td>-0.00546***</td>
</tr>
<tr>
<td></td>
<td>(0.00175)</td>
<td>(0.00320)</td>
<td>(0.00244)</td>
<td>(0.00210)</td>
</tr>
<tr>
<td>Free/reduced price lunch</td>
<td>-0.0585***</td>
<td>-0.0615***</td>
<td>-0.0601***</td>
<td>-0.0587***</td>
</tr>
<tr>
<td></td>
<td>(0.00205)</td>
<td>(0.00372)</td>
<td>(0.00276)</td>
<td>(0.00236)</td>
</tr>
<tr>
<td>Limited English Prof.</td>
<td>0.0427***</td>
<td>0.0331***</td>
<td>0.0354***</td>
<td>0.0390***</td>
</tr>
<tr>
<td></td>
<td>(0.00552)</td>
<td>(0.00942)</td>
<td>(0.00744)</td>
<td>(0.00646)</td>
</tr>
<tr>
<td>Disabled</td>
<td>-0.177***</td>
<td>-0.175***</td>
<td>-0.174***</td>
<td>-0.177***</td>
</tr>
<tr>
<td></td>
<td>(0.00438)</td>
<td>(0.00862)</td>
<td>(0.00595)</td>
<td>(0.00513)</td>
</tr>
<tr>
<td>No. of Binding Subgroups</td>
<td>0.000595</td>
<td>0.00509*</td>
<td>0.00275</td>
<td>0.00192</td>
</tr>
<tr>
<td></td>
<td>(0.00155)</td>
<td>(0.00303)</td>
<td>(0.00215)</td>
<td>(0.00193)</td>
</tr>
<tr>
<td>% Free/reduced price lunch</td>
<td>0.121*</td>
<td>0.0914</td>
<td>0.182**</td>
<td>0.0812</td>
</tr>
<tr>
<td></td>
<td>(0.0677)</td>
<td>(0.118)</td>
<td>(0.0795)</td>
<td>(0.0768)</td>
</tr>
<tr>
<td>% LEP</td>
<td>0.220*</td>
<td>0.778***</td>
<td>0.230</td>
<td>0.284*</td>
</tr>
<tr>
<td></td>
<td>(0.128)</td>
<td>(0.292)</td>
<td>(0.193)</td>
<td>(0.158)</td>
</tr>
<tr>
<td>% White</td>
<td>0.0974</td>
<td>-0.294</td>
<td>-0.0226</td>
<td>0.0427</td>
</tr>
<tr>
<td></td>
<td>(0.0964)</td>
<td>(0.223)</td>
<td>(0.132)</td>
<td>(0.108)</td>
</tr>
<tr>
<td>% Disabled</td>
<td>0.0564</td>
<td>-0.0634</td>
<td>0.0862</td>
<td>0.0718</td>
</tr>
<tr>
<td></td>
<td>(0.109)</td>
<td>(0.255)</td>
<td>(0.172)</td>
<td>(0.142)</td>
</tr>
<tr>
<td>N</td>
<td>461888</td>
<td>142586</td>
<td>240447</td>
<td>316892</td>
</tr>
<tr>
<td>R²</td>
<td>0.644</td>
<td>0.652</td>
<td>0.651</td>
<td>0.650</td>
</tr>
</tbody>
</table>

Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%

Regressions include controls for prior achievement, year, school, and grade fixed effects along with a constant. The years include 2004/06, with 2003 for lags. Column (2) includes observations within 5 students from the margin for passing $|TestPasses_{st-1}| \leq 5$, and similarly for “within 10” and “within 15” for columns (3) and (4).
Table 10: Effect of Failing AYP By Prior Achievement Level

<table>
<thead>
<tr>
<th></th>
<th>Whole (1)</th>
<th>Within 5 (2)</th>
<th>Within 10 (3)</th>
<th>Within 15 (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Reading</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FailAYP$_{t-1}$</td>
<td>0.0356***</td>
<td>0.0554</td>
<td>0.0679***</td>
<td>0.0350**</td>
</tr>
<tr>
<td></td>
<td>(0.00746)</td>
<td>(0.0415)</td>
<td>(0.0232)</td>
<td>(0.0149)</td>
</tr>
<tr>
<td>FailAYP$<em>{t-1}$×Ach 1$</em>{t-1}$</td>
<td>-0.0727***</td>
<td>-0.0815*</td>
<td>-0.0504</td>
<td>-0.0456*</td>
</tr>
<tr>
<td></td>
<td>(0.0199)</td>
<td>(0.0426)</td>
<td>(0.0311)</td>
<td>(0.0258)</td>
</tr>
<tr>
<td>FailAYP$<em>{t-1}$×Ach 2$</em>{t-1}$</td>
<td>-0.0426***</td>
<td>0.0101</td>
<td>-0.00444</td>
<td>-0.0158</td>
</tr>
<tr>
<td></td>
<td>(0.00993)</td>
<td>(0.0215)</td>
<td>(0.0149)</td>
<td>(0.0128)</td>
</tr>
<tr>
<td>FailAYP$<em>{t-1}$×Ach 3$</em>{t-1}$</td>
<td>-0.0164***</td>
<td>-0.00999</td>
<td>-0.00230</td>
<td>-0.00552</td>
</tr>
<tr>
<td></td>
<td>(0.00534)</td>
<td>(0.0103)</td>
<td>(0.00726)</td>
<td>(0.00659)</td>
</tr>
<tr>
<td><strong>Panel B: Math</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FailAYP$_{t-1}$</td>
<td>0.0454***</td>
<td>-0.00934</td>
<td>0.0576*</td>
<td>0.0250</td>
</tr>
<tr>
<td></td>
<td>(0.00890)</td>
<td>(0.0463)</td>
<td>(0.0306)</td>
<td>(0.0198)</td>
</tr>
<tr>
<td>FailAYP$<em>{t-1}$×Ach 1$</em>{t-1}$</td>
<td>-0.00860</td>
<td>-0.104*</td>
<td>-0.0719*</td>
<td>-0.0397</td>
</tr>
<tr>
<td></td>
<td>(0.0261)</td>
<td>(0.0577)</td>
<td>(0.0407)</td>
<td>(0.0342)</td>
</tr>
<tr>
<td>FailAYP$<em>{t-1}$×Ach 2$</em>{t-1}$</td>
<td>-0.0113</td>
<td>0.0271</td>
<td>0.00917</td>
<td>0.0163</td>
</tr>
<tr>
<td></td>
<td>(0.0109)</td>
<td>(0.0221)</td>
<td>(0.0158)</td>
<td>(0.0137)</td>
</tr>
<tr>
<td>FailAYP$<em>{t-1}$×Ach 3$</em>{t-1}$</td>
<td>-0.0405***</td>
<td>0.00702</td>
<td>-0.00651</td>
<td>-0.0130</td>
</tr>
<tr>
<td></td>
<td>(0.00705)</td>
<td>(0.0146)</td>
<td>(0.0105)</td>
<td>(0.00899)</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>462425</td>
<td>142749</td>
<td>240697</td>
<td>317225</td>
</tr>
<tr>
<td><strong>R²</strong></td>
<td>0.603</td>
<td>0.611</td>
<td>0.607</td>
<td>0.605</td>
</tr>
</tbody>
</table>

Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%

Regressions include controls for prior achievement, prior achievement level dummies, year, school, and grade fixed effects along with a constant and demographic controls as shown in Tables 8 and 9. The years include 2004/06, with 2003 for lags. Column (2) includes observations within 5 students from the margin for passing $|TestPasses_{st-1}| \leq 5$, and similarly for “within 10” and “within 15” for columns (3) and (4).
Table 11: Effect of Failing AYP: Difference-in-Differences

<table>
<thead>
<tr>
<th></th>
<th>Reading</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>FailAYP&lt;sub&gt;t-1&lt;/sub&gt;×Title I</td>
<td>0.0102** (0.00506)</td>
<td>0.00758 (0.00515)</td>
</tr>
<tr>
<td>FailAYP&lt;sub&gt;t-1&lt;/sub&gt;×Ach&lt;sub&gt;1&lt;/sub&gt;&lt;sub&gt;t-1&lt;/sub&gt;×Title I</td>
<td>-0.103*** (0.0230)</td>
<td>0.0617** (0.0245)</td>
</tr>
<tr>
<td>FailAYP&lt;sub&gt;t-1&lt;/sub&gt;×Ach&lt;sub&gt;2&lt;/sub&gt;&lt;sub&gt;t-1&lt;/sub&gt;×Title I</td>
<td>0.0272** (0.0112)</td>
<td>0.0107 (0.0103)</td>
</tr>
<tr>
<td>FailAYP&lt;sub&gt;t-1&lt;/sub&gt;×Ach&lt;sub&gt;3&lt;/sub&gt;&lt;sub&gt;t-1&lt;/sub&gt;×Title I</td>
<td>0.0149*** (0.00576)</td>
<td>0.00872 (0.00716)</td>
</tr>
<tr>
<td>FailAYP&lt;sub&gt;t-1&lt;/sub&gt;×FailingSubgroup&lt;sub&gt;t-1&lt;/sub&gt;×Title I</td>
<td></td>
<td>0.0200*** (0.00688)</td>
</tr>
<tr>
<td>N</td>
<td>1357375</td>
<td>1357375</td>
</tr>
<tr>
<td>R²</td>
<td>0.616</td>
<td>0.618</td>
</tr>
</tbody>
</table>

Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% Table presents only interactions with Title I schools, where non-Title I schools act as a control group. Regressions also control for level effect of failing AYP and a dummy for whether the school is a Title I school, along with controls for prior achievement, prior achievement level dummies, year, school, and grade fixed effects along with a constant and demographic controls as shown in Tables 8 and 9. The years include 2004/06, with 2003 for lags. Columns (2) and (5) include controls for the effect of failing AYP in the prior year interacted with the achievement level of the students. Column (3) and (6) include controls for the effect of failing AYP interacted with the student being in a failing subgroup in t – 1.
### Table 12: Effect of Failing AYP and Sanctions

<table>
<thead>
<tr>
<th></th>
<th>Whole (1)</th>
<th>Within 5 (2)</th>
<th>Within 10 (3)</th>
<th>Within 15 (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Reading</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FailAYP&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>0.0164**</td>
<td>0.0674*</td>
<td>0.0847***</td>
<td>0.0240</td>
</tr>
<tr>
<td></td>
<td>(0.00762)</td>
<td>(0.0366)</td>
<td>(0.0238)</td>
<td>(0.0170)</td>
</tr>
<tr>
<td>Sanction 1</td>
<td>0.0407***</td>
<td>0.0944</td>
<td>0.0922**</td>
<td>0.0546</td>
</tr>
<tr>
<td></td>
<td>(0.0130)</td>
<td>(0.0638)</td>
<td>(0.0452)</td>
<td>(0.0334)</td>
</tr>
<tr>
<td>Sanction 2</td>
<td>0.0637***</td>
<td>0.0902</td>
<td>0.0944</td>
<td>0.0563</td>
</tr>
<tr>
<td></td>
<td>(0.0159)</td>
<td>(0.0881)</td>
<td>(0.0612)</td>
<td>(0.0395)</td>
</tr>
<tr>
<td>N</td>
<td>462425</td>
<td>142749</td>
<td>240697</td>
<td>317225</td>
</tr>
<tr>
<td>R²</td>
<td>0.599</td>
<td>0.608</td>
<td>0.603</td>
<td>0.601</td>
</tr>
<tr>
<td><strong>Panel B: Math</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FailAYP&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>0.0345***</td>
<td>0.0147</td>
<td>0.0985***</td>
<td>0.0531***</td>
</tr>
<tr>
<td></td>
<td>(0.00933)</td>
<td>(0.0410)</td>
<td>(0.0314)</td>
<td>(0.0197)</td>
</tr>
<tr>
<td>Sanction 1</td>
<td>0.0903***</td>
<td>0.0394</td>
<td>0.143**</td>
<td>0.127**</td>
</tr>
<tr>
<td></td>
<td>(0.0184)</td>
<td>(0.0734)</td>
<td>(0.0645)</td>
<td>(0.0493)</td>
</tr>
<tr>
<td>Sanction 2</td>
<td>0.0917***</td>
<td>-0.0955</td>
<td>0.0910</td>
<td>0.0842</td>
</tr>
<tr>
<td></td>
<td>(0.0232)</td>
<td>(0.0928)</td>
<td>(0.0818)</td>
<td>(0.0550)</td>
</tr>
<tr>
<td>N</td>
<td>461888</td>
<td>142586</td>
<td>240447</td>
<td>316892</td>
</tr>
<tr>
<td>R²</td>
<td>0.644</td>
<td>0.652</td>
<td>0.651</td>
<td>0.649</td>
</tr>
</tbody>
</table>

Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%

Regressions include controls for prior achievement, year, school, and grade fixed effects along with a constant and demographic controls as shown in Tables 8 and 9. Unlike the previous sets of regressions that focus on an overall effect of failing AYP for all schools that fail, $\text{FailAYP}_{t-1}$ denotes here schools that failed AYP in $t - 1$, but do not face sanctions in $t$. The years include 2004/06, with 2003 for lags. Column (2) includes observations within 5 students from the margin for passing $|\text{TestPasses}_{st-1}| \leq 5$, and similarly for “within 10” and “within 15” for columns (3) and (4).
Table 13: Effect of Failing AYP and Sanctions by Prior Achievement Level

<table>
<thead>
<tr>
<th></th>
<th>Reading</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Whole</td>
<td>Within 5</td>
</tr>
<tr>
<td>FailAYP(_{t-1})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0255***</td>
<td>0.0665*</td>
</tr>
<tr>
<td></td>
<td>(0.00853)</td>
<td>(0.0390)</td>
</tr>
<tr>
<td>FailAYP(<em>{t-1}) × Ach 1(</em>{t-1})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0440*</td>
<td>-0.0255</td>
</tr>
<tr>
<td></td>
<td>(0.0226)</td>
<td>(0.0551)</td>
</tr>
<tr>
<td>FailAYP(<em>{t-1}) × Ach 2(</em>{t-1})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0252**</td>
<td>0.0322</td>
</tr>
<tr>
<td></td>
<td>(0.0113)</td>
<td>(0.0275)</td>
</tr>
<tr>
<td>FailAYP(<em>{t-1}) × Ach 3(</em>{t-1})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.00585</td>
<td>-0.00564</td>
</tr>
<tr>
<td></td>
<td>(0.00631)</td>
<td>(0.0138)</td>
</tr>
<tr>
<td>Sanction 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0670***</td>
<td>0.0990</td>
</tr>
<tr>
<td></td>
<td>(0.0138)</td>
<td>(0.0604)</td>
</tr>
<tr>
<td>Sanction 1 × Ach 1(_{t-1})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.117***</td>
<td>-0.214***</td>
</tr>
<tr>
<td></td>
<td>(0.0366)</td>
<td>(0.0868)</td>
</tr>
<tr>
<td>Sanction 1 × Ach 2(_{t-1})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0683***</td>
<td>-0.000819</td>
</tr>
<tr>
<td></td>
<td>(0.0192)</td>
<td>(0.0413)</td>
</tr>
<tr>
<td>Sanction 1 × Ach 3(_{t-1})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0182*</td>
<td>0.00300</td>
</tr>
<tr>
<td></td>
<td>(0.0109)</td>
<td>(0.0187)</td>
</tr>
<tr>
<td>Sanction 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.141***</td>
<td>0.176*</td>
</tr>
<tr>
<td></td>
<td>(0.0173)</td>
<td>(0.104)</td>
</tr>
<tr>
<td>Sanction 2 × Ach 1(_{t-1})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.134***</td>
<td>-0.191*</td>
</tr>
<tr>
<td></td>
<td>(0.0403)</td>
<td>(0.112)</td>
</tr>
<tr>
<td>Sanction 2 × Ach 2(_{t-1})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.146***</td>
<td>-0.163***</td>
</tr>
<tr>
<td></td>
<td>(0.0224)</td>
<td>(0.0595)</td>
</tr>
<tr>
<td>Sanction 2 × Ach 3(_{t-1})</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0859***</td>
<td>-0.137**</td>
</tr>
<tr>
<td></td>
<td>(0.0150)</td>
<td>(0.0554)</td>
</tr>
</tbody>
</table>

N               | 462425 | 142749   | 240697   | 317225   | 461888 | 142586   | 240447   | 316892    |
R\(^2\)          | 0.603  | 0.607    | 0.605    | 0.650    | 0.650  | 0.657    | 0.655    | 0.654     |

Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%; Regressions include controls for prior achievement, prior achievement level dummies, year, school, and grade fixed effects along with a constant and demographic controls as shown in Tables 8 and 9. Unlike the previous sets of regressions that focus on an overall effect of failing AYP for all schools that fail, \(\text{FailAYP}_{t-1}\) denotes here schools that failed AYP in \(t-1\), but do not face sanctions in \(t\). The years include 2004/06, with 2003 for lags. Column (2) includes observations within 5 students from the margin for passing \(|\text{TestPasses}_{st-1}| \leq 5\), and similarly for “within 10” and “within 15” for columns (3) and (4).
Table 14: Effect of Sanctions on School Choice

<table>
<thead>
<tr>
<th>Sanction $1_{t+1}$ $\times$ Title I</th>
<th>0.00571***</th>
<th>0.00477***</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.00114)</td>
<td>(0.00167)</td>
</tr>
<tr>
<td>Choices $\times$ Sanction $1_{t+1}$ $\times$ Title I</td>
<td>0.00084***</td>
<td>0.00312***</td>
</tr>
<tr>
<td></td>
<td>(9.25e-05)</td>
<td>(5.56e-05)</td>
</tr>
<tr>
<td>Sanction $2_{t+1}$ $\times$ Title I</td>
<td>-0.00104</td>
<td>-0.00399*</td>
</tr>
<tr>
<td></td>
<td>(0.00154)</td>
<td>(0.00230)</td>
</tr>
<tr>
<td>Choices $\times$ Sanction $2_{t+1}$ $\times$ Title I</td>
<td>0.00119***</td>
<td>0.00532***</td>
</tr>
<tr>
<td></td>
<td>(0.00163)</td>
<td>(0.000112)</td>
</tr>
<tr>
<td>Choices in District</td>
<td>0.00149***</td>
<td>0.00149***</td>
</tr>
<tr>
<td></td>
<td>(3.38e-05)</td>
<td>(4.25e-05)</td>
</tr>
<tr>
<td>Choices $\times$ Title I</td>
<td>0.00150***</td>
<td>0.00150***</td>
</tr>
<tr>
<td></td>
<td>(0.00119)</td>
<td>(0.00153)</td>
</tr>
<tr>
<td>Sanction $1_{t+1}$</td>
<td>0.00208***</td>
<td>0.00331***</td>
</tr>
<tr>
<td></td>
<td>(0.000752)</td>
<td>(0.00114)</td>
</tr>
<tr>
<td>Sanction $2_{t+1}$</td>
<td>0.00392***</td>
<td>0.00721***</td>
</tr>
<tr>
<td></td>
<td>(0.000954)</td>
<td>(0.00153)</td>
</tr>
<tr>
<td>Title I</td>
<td>0.00246**</td>
<td>-0.0176***</td>
</tr>
<tr>
<td></td>
<td>(0.00119)</td>
<td>(0.00153)</td>
</tr>
<tr>
<td>American Indian</td>
<td>0.0111***</td>
<td>0.0132***</td>
</tr>
<tr>
<td></td>
<td>(0.00166)</td>
<td>(0.00197)</td>
</tr>
<tr>
<td>Asian</td>
<td>-0.0118***</td>
<td>-0.0140***</td>
</tr>
<tr>
<td></td>
<td>(0.00129)</td>
<td>(0.00153)</td>
</tr>
<tr>
<td>Black</td>
<td>0.0166***</td>
<td>0.0201***</td>
</tr>
<tr>
<td></td>
<td>(0.000470)</td>
<td>(0.000560)</td>
</tr>
<tr>
<td>Disabled</td>
<td>0.00809***</td>
<td>0.0100***</td>
</tr>
<tr>
<td></td>
<td>(0.000514)</td>
<td>(0.000610)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.00435***</td>
<td>-0.00527***</td>
</tr>
<tr>
<td></td>
<td>(0.000860)</td>
<td>(0.00102)</td>
</tr>
<tr>
<td>LEP</td>
<td>0.00426***</td>
<td>0.00503***</td>
</tr>
<tr>
<td></td>
<td>(0.00109)</td>
<td>(0.00127)</td>
</tr>
<tr>
<td>FRP lunch</td>
<td>0.0273***</td>
<td>0.0324***</td>
</tr>
<tr>
<td></td>
<td>(0.000387)</td>
<td>(0.000459)</td>
</tr>
<tr>
<td>Multiracial</td>
<td>0.0118***</td>
<td>0.0132***</td>
</tr>
<tr>
<td></td>
<td>(0.00111)</td>
<td>(0.00128)</td>
</tr>
</tbody>
</table>

Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. The dependent variable is a dummy for whether a student changes schools in the year that a school would face sanctions. Regressions also school demographics, corresponding to the individual demographics reports, along with year, school, and grade fixed effects and a constant. The years include 2003/06, with 2002 for lags. Column (2) considers whether more students move when there are more choices of non-failing school/grades in the district.
Table 15: Effect of Failing AYP and Sanctions by Intensity (Number of Choices)

<table>
<thead>
<tr>
<th></th>
<th>Whole (1)</th>
<th>Within 5 (2)</th>
<th>Within 10 (3)</th>
<th>Within 15 (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FailAYP(_{t-1})</td>
<td>0.00590</td>
<td>0.0263</td>
<td>0.0585*</td>
<td>0.0142</td>
</tr>
<tr>
<td></td>
<td>(0.00996)</td>
<td>(0.0594)</td>
<td>(0.0298)</td>
<td>(0.0200)</td>
</tr>
<tr>
<td>FailAYP(_{t-1}) × No. of Choices</td>
<td>0.000745*</td>
<td>-0.000507</td>
<td>0.000506</td>
<td>0.000644</td>
</tr>
<tr>
<td></td>
<td>(0.000425)</td>
<td>(0.00102)</td>
<td>(0.000653)</td>
<td>(0.000499)</td>
</tr>
<tr>
<td>Sanction 1</td>
<td>0.000178</td>
<td>-0.417</td>
<td>-0.0169</td>
<td>-0.0447</td>
</tr>
<tr>
<td></td>
<td>(0.0178)</td>
<td>(0.530)</td>
<td>(0.0846)</td>
<td>(0.0547)</td>
</tr>
<tr>
<td>Sanction 1 × No. of Choices</td>
<td>0.00351***</td>
<td>0.0537</td>
<td>0.00719</td>
<td>0.00749**</td>
</tr>
<tr>
<td></td>
<td>(0.00129)</td>
<td>(0.0573)</td>
<td>(0.00503)</td>
<td>(0.00323)</td>
</tr>
<tr>
<td>Sanction 2</td>
<td>0.0481**</td>
<td>-0.357</td>
<td>0.195</td>
<td>-0.00449</td>
</tr>
<tr>
<td></td>
<td>(0.0211)</td>
<td>(0.573)</td>
<td>(0.137)</td>
<td>(0.0558)</td>
</tr>
<tr>
<td>Sanction 2 × No. of Choices</td>
<td>0.000760</td>
<td>0.0693</td>
<td>-0.0222</td>
<td>0.00320</td>
</tr>
<tr>
<td></td>
<td>(0.00156)</td>
<td>(0.0975)</td>
<td>(0.0176)</td>
<td>(0.00500)</td>
</tr>
<tr>
<td>N</td>
<td>462425</td>
<td>142749</td>
<td>240697</td>
<td>317225</td>
</tr>
<tr>
<td>R(^2)</td>
<td>0.599</td>
<td>0.606</td>
<td>0.603</td>
<td>0.601</td>
</tr>
</tbody>
</table>

Panel B: Math

<table>
<thead>
<tr>
<th></th>
<th>Whole (1)</th>
<th>Within 5 (2)</th>
<th>Within 10 (3)</th>
<th>Within 15 (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FailAYP(_{t-1})</td>
<td>0.0285**</td>
<td>0.0222</td>
<td>0.106***</td>
<td>0.0428*</td>
</tr>
<tr>
<td></td>
<td>(0.0129)</td>
<td>(0.0558)</td>
<td>(0.0412)</td>
<td>(0.0246)</td>
</tr>
<tr>
<td>FailAYP(_{t-1}) × No. of Choices</td>
<td>0.000452</td>
<td>-0.00141</td>
<td>-0.000540</td>
<td>0.000680</td>
</tr>
<tr>
<td></td>
<td>(0.000560)</td>
<td>(0.00181)</td>
<td>(0.000949)</td>
<td>(0.000655)</td>
</tr>
<tr>
<td>Sanction 1</td>
<td>0.0707***</td>
<td>0.131</td>
<td>0.218**</td>
<td>0.119*</td>
</tr>
<tr>
<td></td>
<td>(0.0243)</td>
<td>(0.310)</td>
<td>(0.109)</td>
<td>(0.0699)</td>
</tr>
<tr>
<td>Sanction 1 × No. of Choices</td>
<td>0.000747</td>
<td>-0.0122</td>
<td>-0.00490</td>
<td>0.000107</td>
</tr>
<tr>
<td></td>
<td>(0.00119)</td>
<td>(0.0351)</td>
<td>(0.00417)</td>
<td>(0.00281)</td>
</tr>
<tr>
<td>Sanction 2</td>
<td>0.0578**</td>
<td>0.133</td>
<td>0.304*</td>
<td>0.0832</td>
</tr>
<tr>
<td></td>
<td>(0.0291)</td>
<td>(0.345)</td>
<td>(0.158)</td>
<td>(0.0700)</td>
</tr>
<tr>
<td>Sanction 2 × No. of Choices</td>
<td>0.00266</td>
<td>-0.0464</td>
<td>-0.0248</td>
<td>-0.00103</td>
</tr>
<tr>
<td></td>
<td>(0.00194)</td>
<td>(0.0658)</td>
<td>(0.0200)</td>
<td>(0.00395)</td>
</tr>
<tr>
<td>N</td>
<td>461888</td>
<td>142586</td>
<td>240447</td>
<td>316892</td>
</tr>
<tr>
<td>R(^2)</td>
<td>0.644</td>
<td>0.652</td>
<td>0.650</td>
<td>0.649</td>
</tr>
</tbody>
</table>

Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% Number of choices denotes the number of non-failing options in the district. Regressions include controls for prior achievement, number of choices, year, school, and grade fixed effects along with a constant and demographic controls as shown in Tables 8 and 9. Unlike the previous sets of regressions that focus on an overall effect of failing AYP for all schools that fail, \(\text{FailAYP}_{t-1}\) denotes here schools that failed AYP in \(t-1\), but do not face sanctions in \(t\). The years include 2004/06, with 2003 for lags. Column (2) includes observations within 5 students from the margin for passing \(|\text{TestPasses}_{st-1}| ≤ 5\), and similarly for “within 10” and “within 15” for columns (3) and (4).
Table 16: Effect of Sanctions by Intensity (Number of Choices)

<table>
<thead>
<tr>
<th>Interaction</th>
<th>Reading</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sanction 1 × Title I</td>
<td>-0.00175 (0.00517)</td>
<td>0.0155*** (0.00466)</td>
</tr>
<tr>
<td></td>
<td>0.00464 (0.00534)</td>
<td>0.0124** (0.00482)</td>
</tr>
<tr>
<td>Sanction 1 × Choices × Title I</td>
<td>0.000378 (0.000332)</td>
<td>0.000543* (0.000790***</td>
</tr>
<tr>
<td></td>
<td>0.001035 (0.000332)</td>
<td>0.000790*** (0.000300)</td>
</tr>
<tr>
<td>Sanction 1 × TargStud × Title I</td>
<td>0.00382 (0.0129)</td>
<td>0.0481*** (0.0117)</td>
</tr>
<tr>
<td>Sanction 1 × Choices × TargStud × Title I</td>
<td>0.00233*** (0.000818)</td>
<td>-0.00152 (0.000739)</td>
</tr>
<tr>
<td>Sanction 2 × Title I</td>
<td>-0.00767 (0.00778)</td>
<td>-0.0257*** (0.00702)</td>
</tr>
<tr>
<td></td>
<td>-0.00197 (0.00842)</td>
<td>-0.0230*** (0.00760)</td>
</tr>
<tr>
<td>Sanction 2 × Choices × Title I</td>
<td>0.00140 (0.000694)</td>
<td>0.00206*** (0.000627)</td>
</tr>
<tr>
<td></td>
<td>0.00129* (0.000565)</td>
<td>0.00211*** (0.000627)</td>
</tr>
<tr>
<td>Sanction 2 × TargStud × Title I</td>
<td>0.0152 (0.0163)</td>
<td>-0.00974 (0.0147)</td>
</tr>
<tr>
<td>Sanction 2 × Choices × TargStud × Title I</td>
<td>0.00220* (0.00130)</td>
<td>0.00118 (0.00117)</td>
</tr>
<tr>
<td>FailAYP_{t-1} × Title I</td>
<td>0.00834** (0.00334)</td>
<td>0.0140*** (0.00890***</td>
</tr>
<tr>
<td></td>
<td>0.0100*** (0.00345)</td>
<td>0.00890*** (0.00312)</td>
</tr>
<tr>
<td>FailAYP_{t-1} × Choices × Title I</td>
<td>-0.000101 (0.000161)</td>
<td>0.000220 (0.000169)</td>
</tr>
<tr>
<td></td>
<td>-1.94e-05 (0.000169)</td>
<td>0.000423*** (0.000152)</td>
</tr>
<tr>
<td>FailAYP_{t-1} × TargStud × Title I</td>
<td>0.0353*** (0.00822)</td>
<td>0.0609*** (0.00744)</td>
</tr>
<tr>
<td>FailAYP_{t-1} × Choices × TargStud × Title I</td>
<td>-0.000498 (0.000385)</td>
<td>-0.00117*** (0.000348)</td>
</tr>
</tbody>
</table>

N: 1357375 1357375 1355388 1355388
R²: 0.618 0.619 0.685 0.686

Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Schools designated as failing AYP are defined as failing AYP in year \(t-1\) but not facing sanctions in year \(t\). Choices denotes the number of options for a student in the district for a non-sanctioned schools with the relevant grade. Columns (1) and (3) correspond to the same regression as in Table 12, but with controls for Choices along with the interactions with Title I schools shown in the table. Only interactions with Title I schools reported, though regressions also controls for a level effect of Sanction and FailAYP. Columns (2) and (4) break out the results by TargStud. A target student is defined as a failing student (below achievement level 3) in a subgroup that failed AYP in the prior year. These regressions also control for TargStud interacted with Sanction and FailAYP dummies and Choices. All regressions include the demographic controls described in Table 12, along with controls for the student’s prior year achievement level, the number of subgroups that count toward AYP status, a dummy for whether the school is a Title I school, and year, school, and grade fixed effects along with a constant. The years include 2003/06, with 2002 for lags.
B Figures

Figure 1: Title I Improvement

<table>
<thead>
<tr>
<th>After...</th>
<th>then...</th>
<th>and implements...</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 year of not making Adequate Yearly Progress (AYP) in mathematics or reading/language arts.</td>
<td>the school addresses challenge areas.</td>
<td>no sanctions under NCLB.</td>
</tr>
<tr>
<td>2 consecutive years of not making AYP in the same subject.</td>
<td>the school enters Year 1 of Title I Improvement at the beginning of the next school year</td>
<td>school choice, unless the school is in a pilot district offering supplemental educational services as the first year option, and receives technical assistance.</td>
</tr>
<tr>
<td>3 years of not making AYP in the same subject.</td>
<td>the school enters Year 2 of Title I Improvement at the beginning of the next school year</td>
<td>school choice, supplemental educational services and receives technical assistance.</td>
</tr>
<tr>
<td>4 years of not making AYP in the same subject.</td>
<td>the school enters Year 3 of Title I Improvement at the beginning of the next school year</td>
<td>school choice, supplemental educational services, corrective action and receives technical assistance.</td>
</tr>
<tr>
<td>5 years of not making AYP in the same subject.</td>
<td>the school enters Year 4 of Title I Improvement at the beginning of the next school year</td>
<td>school choice, supplemental educational services, deviates a plan for restructuring and receives technical assistance.</td>
</tr>
<tr>
<td>6 years of not making AYP in the same subject.</td>
<td>the school enters Year 5 of Title I Improvement at the beginning of the next school year</td>
<td>school choice, supplemental educational services, restructuring and receives technical assistance.</td>
</tr>
</tbody>
</table>

NOTES: Title I schools enter Improvement after two consecutive years of not making Adequate Yearly Progress (AYP) in the same subject (reading/language arts or mathematics). Title I schools exit Improvement after two consecutive years of making AYP in the subject that identified them for Improvement. If a school makes AYP in the identifying subject in any one year after entering Improvement, it does not move to the next level of sanctions in the next school year. For every year a school in Improvement does not make AYP in the identifying subject, it moves to the next level of sanctions. It is possible for a school to exit Improvement for one subject, while entering into or remaining in Improvement based on the other subject.

Source: [http://www.ncpublicschools.org/docs/nclb/includes/nclbtitle1chart.pdf](http://www.ncpublicschools.org/docs/nclb/includes/nclbtitle1chart.pdf)
Figure 2: Percent Free/Reduced Price Lunch in Title 1 vs. Non-Title I Schools

Figure 3: Densities of Achievement Pre-Year 1 Sanctions

Figure 4: Probability of Making AYP vs. Testpasses
Figure 5: % Nonwhite Students vs. Testpasses

Figure 6: % FRL Students vs. Testpasses

Figure 7: Number of AYP targets vs. Testpasses
Figure 8: Achievement Gains by Forcing Variable (Testpasses)

The gains condition on percent free/reduced price lunch and percent nonwhite.