The Effects of Open Enrollment under 'No Child Left Behind'

Jessica Wagner

University of Toronto*

June 1, 2022

Abstract

This paper assesses whether expansion of school choice through the open enrollment provision of 'No Child Left Behind' (NCLB) was effective in shifting students away from low performing schools in California. Under NCLB, schools failing to make 'Adequate Yearly Progress' (AYP) for two or more years were forced to offer students the option to transfer to AYP-passing schools. Using the discontinuous assignment to treatment based on a school achievement threshold, I identify the causal effect of being sanctioned with open enrollment on subsequent enrollment growth. I uncover small and statistically insignificant enrollment declines for schools that marginally failed to make AYP and were sanctioned with school choice for one year, and statistically significant declines in enrolment growth of 2.7 percentage points for schools that had school choice for a second year. Evaluation of heterogeneous treatment effects, and a case study of a large urban district, suggest that due to supply constraints transfer take-up was low even in the presence of high-performing nearby schools. The results highlight the limitations of top-down school choice policies that fail to expand access to quality schooling options.

*I thank my supervisor, Robert McMillan, for his invaluable guidance and support throughout this project. Thank you also to Natalie Bau, Gustavo Bobonis, Loren Brandt, Elizabeth Dhuey, Yao Luo, Jordi Mondria, Carolyn Pitchik, and participants in the Second Year Paper study group and ECO 4060 seminar for helpful comments and discussion.

1 Introduction

Large gaps in educational achievement persist across socio-economic groups in the United States and, despite the efforts of policy makers, chronically under-performing schools remain prevalent in many areas of the country. School choice policies, particularly involving open enrollment, are increasingly employed to address this issue (Betts & Loveless, 2005; Grady et al., 2010). Open enrollment allows students to attend public schools outside their mandatory geographical boundary, thus de-linking neighbourhoods and schools and expanding options for parents. Some implementations of open enrollment have been heavily studied, such as programs in Boston and New York which use matching algorithms to assign students to their preferred schools.[1] Yet this type of assignment mechanism is employed in only a handful of large urban school districts. More common implementations of open enrollment, such as the federally instituted 'No Child Left Behind' school choice policy, have been studied far less. My research seeks to fill this gap.

This paper identifies the effect on student enrollment patterns of offering students in low performing schools the opportunity to transfer to higher performing schools within their district. The enactment of the 'No Child Left Behind' (NCLB) Act in 2002 introduced a provision forcing schools that failed to make adequate yearly progress (AYP) in academic achievement for two or more consecutive years to offer transfers to other in-district schools. This law represents the largest instance of open enrollment in the US to date. The locally applied expansion of school choice came with transportation support and was lifted only if schools subsequently *met* AYP requirements for two consecutive years. The provision presents a useful opportunity for addressing self-selection issues as the sanction applied discontinuously to schools on either side of an achievement cut-off. I use this treatment threshold in a regression discontinuity design to estimate the causal effect of the open enrollment sanction on subsequent enrollments. Further, an examination of heterogeneous effects of the policy along with a case study of the largest district in the state uncover suggestive evidence of constraints faced by students and education providers under expanded choice.[2]

While a substantial body of work has investigated the accountability aspects of NCLB and their impact on achievement,[3] less attention has been paid to understanding the role of the law's

[1]See, e.g., Avery and Pathak (2021) for a summary of this research. In general, algorithms which match students and schools based on elicited rankings have been shown to increase welfare and boost achievement. However, depending on the match process they may be susceptible to manipulation by sophisticated students.

[2]The evidence is suggestive in that I am unable to identify causal estimates of heterogeneous treatment effects or parameters in the demand or supply functions for school spaces.

[3]See Dee and Jacob (2011), Gilraine (2018), Hemelt (2011), Krieg (2008, 2011), Macartney et al. (2015), Neal and

sanctions. Failing schools faced regulatory actions intended both to extend a lifeline to students in inadequate schools and to incentivize their educators via penalties.[4] In theory, the school choice sanction giving students the opportunity to transfer provides better educational opportunities to students in weak schools. Further, it puts pressure on educators to increase productivity or risk losing students, and with them, funding. Prior work has found that students and parents respond to expanded school choice by seeking information on school quality (Lovenheim & Walsh, 2018) and with demand for high achieving schools (Avery & Pathak, 2021; Deming et al., 2014; Hastings & Weinstein, 2008). However, in practice there may be limited options for switching to better schools, and districts may lack the resources or sufficient incentives to provide enhanced schooling options. The present study investigates the extent to which students utilized the school choice option and seeks to uncover whether supply-side constraints prevented them from doing so. This research has implications both for policy design as well as for understanding the preferences of students and local education agencies over school assignment.

This study examines the effects of school choice under NCLB in the state with the largest public school system in the US: California. Using a panel of school-level data from the California Department of Education, I compare changes in enrollment of similar schools across the achievement cut-off for making AYP, where schools above the cut-off just failed to achieve AYP and schools below it just passed. If school choice succeeds in empowering parents and creating competitive pressure for schools, one would expect to observe enrollment losses for schools that marginally failed in the year following failure. However, I find statistically insignificant declines in enrollment at schools whose students were offered transfers for the first time. This result holds for a subset of schools that had nearby higher performing schools to which students might have transferred, and for student subgroups that were among the lowest performers in their school as well as those who were the highest performers.

To investigate whether transfers under the program are simply delayed, I examine the second year in which schools are subject to expanded choice. To avoid confounding with the effects of additional sanctions, I limit the sample to school that just passed AYP and therefore are still subject to choice but not additional sanctions. This has the added benefit of limiting anticipatory effects of additional future sanctions. Comparing these schools to unsanctioned schools, it does appear as though transfer take-up was realized in the second year of the program. Schools with

Schanzenbach (2010), and Wong et al. (2009)

[4] For a detailed description of the full schedule of accountability sanctions see Section 2.1.

choice saw enrollment growth of 2.7 percentage points lower than schools without. I find some suggestive evidence for the existence of supply constraints, as stronger enrollment adjustments were not observed for schools which had better in-district alternatives and would therefore be expected to have higher demand for out-transfers. In a case study using student-level data from the a large California district, few students take up open enrollment transfers in either the first or second year in which they are available, likely due to the lack of nearby schools that offer seats to transfer students. These findings highlight the importance of implementation and enforcement if school choice policies are to overcome the status quo and improve educational opportunities for disadvantaged students.

I contribute to earlier work that examines the achievement effects of school choice and open enrollment, as well as to analyses of NCLB sanctions. Open enrollment policies present identification challenges as there is potential for endogenous selection both into the decision to institute the policy and in the decision of students to make use of it. Further, general equilibrium effects may permeate as students react to a new school assignment system and these are difficult to disentangle. Many studies of open enrollment circumvent these issues by relying on random variation generated by lotteries to allocate seats at oversubscribed schools (Bruhn, 2020; Deming et al., 2014; Hastings et al., 2006). As acknowledged by these papers, selection into the lottery itself limits the generalizability of results. Additionally, this identification strategy lacks scope for identifying how the supply of practical school options contributes to the effectiveness of these programs. My analysis of the NCLB choice sanction addresses both selection and spillovers by exploiting a quasi-random and locally applied policy change to examine both the demand and supply response under expanded school choice.

This paper is also closely related to work that examines the school choice sanction of NCLB. The challenge in isolating the lone effect of any one NCLB sanction is that they are applied as a series of escalating penalties when schools continue to fail consecutive years of AYP. Ahn and Vigdor (2014) estimate the effects of each NCLB sanction on student achievement growth, exploiting a reversal in the ordering of sanctions in some North Carolina school districts to remove the impacts of past sanctions and the threat of upcoming ones. They find that only the final sanction involving school restructuring had statistically significant impacts on achievement. With ample data points surrounding the AYP threshold in a given year, I expand on Ahn and Vigdor's analysis with precise estimates of the responsiveness to early sanctions. The inhibited take-up I document in the first year of choice provides a potential explanation for their null result, and stronger effects in the second year may suggest that test score effects should be revisited.

Other related work studying NCLB has found that failing AYP at any point, unconditional on the number of consecutive fails, increased out-transfers by about 10 students in North Carolina schools, or 10% of a standard deviation in the typical transfer rate (Holbein, 2016). Research examining the choice 'threat'[5] in Wisconsin finds some evidence of proficiency rate improvements in marginally failing schools and some correlational evidence suggesting that the effects are stronger when schools are exposed to more local competition in the form of nearby passing schools (Chakrabarti, 2014). Papers focusing on first-year failure show that test score gains are attributable to increased teacher effort, as evidenced in North Carolina administrative data (Macartney et al., 2015), and nation-wide surveys (Reback et al., 2014), but do not distinguish which accountability sanctions induce teachers to improve performance.

The remainder of the paper proceeds as follows: Section 2 presents background on NCLB and its implementation in California, Section 3 describes the data I use, Section 4 outlines the empirical methodology, Section 5 presents the results, Section 6 explores the robustness of estimates and alternative estimation strategies, Section 7 examines heterogeneous treatment effects, and Section 8 concludes.

2 Background

2.1 No Child Left Behind Act

'No Child Left Behind' (NCLB) was a major US federal education reform implemented in academic year (AY) 2001/02 as an amendment to the 'Elementary and Secondary School Act' in place since 1965. The policy was designed with the goal of closing educational gaps between advantaged and disadvantaged students using two main methods: increasing accountability and increasing school choice. Accountability was instituted by setting testing standards and penalizing schools – and by association, teachers – for failing to bring students of all demographic and socio-economic subgroups up to those standards. School choice made more school options available to students in the event that their school did not meet NCLB testing standards, with the possibility of the school itself

[5]The first year of NCLB failure is often referred to as inducing the 'choice threat' in the literature, although in truth schools who fail once are subject to the threat of all sanctions, with school choice being the most imminent.

restructuring as a private charter if failure persisted.

The rationale for accountability is to put pressure on schools and teachers to increase educational productivity. By collecting and publicizing standardized test results, the policy creates stigma for poorly performing schools which could result in the loss of students, and with them funding. Further, sanctions imposed under NCLB legislation, described below, directly penalize schools and teachers through potential school restructuring and personnel changes. The added dimension of school choice should in theory accelerate enrollment losses at poorly performing schools, thereby intensifying the pressure on educators.

NCLB mandated that all states conduct annual standardized tests in mathematics and English Language Arts (ELA) for grades 3 to 8 and one of grades 10 to 12. States were required to define the test performance level at which students were deemed proficient in each subject. Every year, a state-wide target proficiency rate for each subject was set, such that the target proficiency rate would reach 100% for both subjects by 2014. This rate dictates the share of students in every school (combining grades) and in every demographic and socio-economic subgroup within each school who must achieve proficiency. Schools that meet these criteria are said to make 'Adequate Yearly Progress' (AYP). AYP status is forfeited if schools fail to meet at least a 95% test participation rate, in both subjects, for the student population as a whole as well as in eligible subgroups. Subgroups are required to make AYP only if they are sufficiently large. It is up to the individual state to determine the cut-off in subgroup size for AYP eligibility.

If schools receiving Title I funds[6] – over half of all schools nationwide – fail to make AYP or have test participation rates below 95% for any eligible subgroup, they face educational sanctions. Sanctions escalate in severity for every additional year of failure, as follows:

[6]Title I funds provide financial assistance to school districts and schools with high numbers or shares of children from low-income families. In California, they represented about 1% of the total state education budget in AY 2017/18 (California Department of Education, 2018).

Year of failure	Sanction
1	Threatened Status (no sanction)
2	Allow within-district transfers
3	Provide private tutoring and supplemental services
4	Do at least one of the following:
	Replace some staff
	Alter the curriculum
	Decrease authority of the principal
	Appoint an outside expert
	Increase the length of the school day/year
5	Design a school restructuring plan
6	Implement the school restructuring plan.

Schools facing sanctions must meet AYP criteria for two consecutive years before sanctions are lifted. Schools are required to use a share of Title I funds to cover transportation costs from within-district transfers and the cost of private tutoring services. If schools fail to comply with the law, the federal Department of Education has the right to withhold Title I funds. However, few school districts were sanctioned in the first several years of the law, despite complaints that schools were not providing adequate access to choice (Hoff, 2006). The legislation does not outline the process by which districts should offer transfers, except to dictate that they must notify parents in a timely manner[7] and provide at least one option to transfer to a passing school in the district. If passing schools are not available, they are directed to cooperate "to the extent practicable" with other districts to accept their transfer students. In practice, much is left to the discretion of districts. Whether they follow these rules is of particular interest to this study.

2.2 NCLB in California

NCLB was introduced in California three years after the state had initiated its own accountability system with the 'Public Schools Accountability Act' of 1999. The earlier program publicized school performance through the 'Average Performance Index' (API) but did not directly penalize schools that had poor test results. NCLB bore many similarities to this earlier program: mandated testing for most grades (2 to 11), test reporting at the subgroup level[8], and parental notification when

[7]Notification of transfer options must be given at least before the first day of school.

[8]California's subgroups include African American, American indian, asian, Hispanic, Filipino, Pacific Islander, white, economically disadvantaged, disabled, and English learner. Economically disadvantaged students are those eligible for free or reduced prices lunches or neither of whose parents had higher than a high school degree.

schools fell below a performance threshold. Thus, California was well-positioned to implement many NCLB rules quite smoothly, and was able to determine which schools passed or failed in AY 2001/02. Despite this, the state set minimum subgroup sizes and defined proficiency in such a way that over 60% of schools failed in the first year of NCLB. In later years, they applied for and were granted permission to use 'safe harbor' exceptions,[9] special rules for the disability subgroup,[10] and the ability to use 2- and 3-year averages to compute proficiency rates. The use of these exemptions grew over time, with many schools keeping sanctions at bay despite performing near the margin and facing increasingly high standards (Polikoff & Wrabel, 2013). The high failure rates in the first several years of NCLB in California, combined with the increasing use of alternatives to the achievement-based determination of AYP, provide additional motivation to focus on the earliest enactment of the school choice sanction for analysis.

Under the precursory accountability law, all districts were required to post their schools' performance data online and send parents detailed reports of standardized test results at their child's school. Incorporating NCLB mandates, districts were also required to send a notification letter informing parents if the school failed AYP and describing the consequences of this. Sample letters describe the dimension(s) in which schools failed (test-subgroup pairs), the student's right to apply for private tutoring or transfer under the law, available transfer options – which were often none, and what the school planned to do to improve. These letters were sent in the summer, and were at times sent, perhaps intentionally, quite late (Hoff, 2006).

California also operated an open enrollment policy pre-dating NCLB. It mandated that schools accept transfer applicants if they had sufficient space, but did not provide funds for transporting students away from their neighbourhood schools. Districts had discretion over transfers under the law, and the extent to which the state enforced the policy is unclear. Betts et al. (2006) examine administrative data from the San Diego Unified School District and find that 6% of students transferred under the open enrollment policy in 2003. Given that NCLB choice comes with transportation funding, it is perhaps reasonable to expect even more transfer demand under the new policy. However, in 2005 a group of parents filed a complaint against San Diego Unified and Compton Unified Districts, alleging that they had been given notice of the option to transfer only days

[9] Safe harbor allows schools to pass AYP even if they fell below the proficiency rate for a subgroup, as long as that subgroup reduced the share of non-proficient students by 10% or more in one or both subjects. This 10% reduction was allowed to be measured with a 75% confidence interval.

[10] In the disability subgroup, schools were permitted to inflate proficiency rates by 20 percentage points to determine AYP if this was the only subgroup that did not meet the target.

before the school year began and were provided inadequate options to do so (Hoff, 2006). Separately, a 2003 case study by the Center for Education Policy found little take up of school choice in 15 districts across the country (EdSource, 2004). This raises the possibility that districts curtailed transfers either by failing to provide suitable school choice options or by failing to notify parents of their options in a timely manner, both requirements under the law.

In their defence, school administrators argued that there were simply not sufficient school alternatives in some districts and the law did not make provisions, funding or otherwise, for the expansion or contraction of schools to accommodate transfers. It is not clear whether Title I funds were indeed withheld from schools or districts that did not expend the mandated amount on transportation services for transfers, or implement other aspects of the law as required. To date there has been no systematic analysis of the extent to which schools provided, and students took up, the option to transfer schools under NCLB in California.

In AY 2010/11, California implemented the 'Open Enrollment Act.' The Act stipulated that a list of 1000 schools with the lowest API[11] would be compiled by the State Superintendent at the end of each school year and be published no later than September 1. Students in these schools could then apply for transfers no later than January 1 of the next calendar year for the subsequent September enrollment date. The Act outlines detailed procedures for districts to follow regarding transfer applications absent under NCLB, though NCLB transfers were still available. The Act stipulates that "A school district of residence shall not adopt any other policies that in any way prevent or discourage pupils from applying for a transfer to a school district of enrollment" (Section 48355(b)) and also, "Communications to parents or guardians by districts regarding the open enrollment options provided by this article shall be factually accurate and not target individual parents or guardians or residential neighborhoods on the basis of a child's actual or perceived academic or athletic performance or any other personal characteristic" (Section 48355(c)). This is suggestive of a recognition of prior issues with open enrollment implementation that are as yet undocumented in the literature. I limit my analysis to the years prior to AY2010/11, when the additional overlapping open enrollment policy was introduced. The Act itself could be an interesting subject for future work.

[11] This excluded charter schools and restricted the distribution across Elementary, Middle and High Schools to mimic that of the bottom decile of schools. Additionally, single districts could have no more than 10% of their schools on the list (rounded up to the nearest integer if number of schools is not divisible by 10).

3 Data

I use school- and subgroup-level data spanning AY 2001/02 to 2009/10 shared publicly by the California Department of Education (CDE).[12] The AYP data describe the number of testable students[13] enrolled in each subgroup, the number taking the mathematics and ELA tests in each subgroup, the number scoring 'proficient or above' on each test in each subgroup, and the proficiency rates required under NCLB in each subject. The data also flag whether schools met all NCLB criteria and contain some information regarding exemptions. While the AYP data are not reported at the grade level, I also obtain annual data on enrollment by gender, grade and ethnicity published by CDE.[14] Schools are included in the study sample if they served any grade from 2 to 8. Few high schools fail NCLB standards, as the policy was mainly targeted at improving performance at lower grade levels, motivating the focus on primary education. School-year observations in which enrollment grew more than double or shrank by more than half are suppressed as these likely capture expansions and contractions associated with the opening and closing of schools and not the policy of expanded choice. This paper focuses on the years AY2001/02 to AY2004/05 in order to understand the effect of school choice under NCLB as it was first imposed in AY2003/04, as well as how students and schools responded to AYP failure in the years preceding and just after the implementation of school choice.

I supplement the school-level data with student administrative data from the an anonymous large urban school district in the state. These data span the years AY2002/03 to AY2009/10 and link students to their enrolled school each year. I use this information to flag whether a student transferred to another school before graduating. Though I am not able to discern whether a student took advantage of NCLB-mandated school choice to leave their current school, the data allow for more precise measurement of student mobility responses to policy changes. The data also describe student characteristics that determine NCLB subgroups such as ethnicity, socio-economic disadvantage, and English learner status. Lastly, they include individual student test scores each year, which I normalize to the state-level mean and standard deviation.

Table 1 presents summary statistics of the school-level state data describing schools in Panel A and districts in Panel B, and the District student-level data in Panel C. The first two rows of the

[12]Retrieved from https://www.cde.ca.gov/re/pr/aypdatafiles.asp.

[13]Only students in grades 2 to 11 who were continuously enrolled from October to the time of the test in Spring were considered "testable" when determining whether schools met the required test participation rates.

[14]Retrieved from https://www.cde.ca.gov/ds/ad/filesenr.asp.

table demonstrate that test enrollment is a subset of total school enrollment since only grades 2 to 11 are tested. The largest racial groups in California schools are Hispanic and white students, and the average school has more than half of students who are considered socio-economically disadvantaged. While enrollment was declining at the typical school in this period, total enrollment in the state and total number of schools was increasing (not shown) reflecting a shift toward smaller schools. The average district had 7 schools, and schools within a district were typically quite close together and fairly similar in their performance in ELA and mathematics.[15] Schools in the District in the same period were larger on average, and typically the majority of students were Hispanic. Students were also more likely to be socio-economically disadvantaged in this population, and enrollment was declining at a greater rate than elsewhere in the state. In a given year, the average school saw about 7% of students transfer to another school in the district before reaching graduation. Notably, in both English and Math, the District displays substantial variation in the academic achievement across its schools, as shown by the dissimilarity index which is about 3-fold higher than in the rest of the state.

4 Empirical Method

Given that the assignment to treatment (enforced school choice) is based on school performance, comparisons of means across AYP status would likely yield biased estimates of the effect on enrollment.[16] To identify the causal effect of the policy, I implement a regression discontinuity (RD) design, comparing schools that marginally failed the AYP achievement criteria to those that marginally passed, as determined using the lowest performing eligible subgroup. Since failure of any eligible subgroup on either test subject results in failure of the school as a whole, I use the lowest performing eligible subgroup to determine if schools are on the margin of failing AYP. I define the proficiency gap of school s at time t as:

$$X_{st} = min_{g,j}\{A^g_{stj} - T_{tj}\}_{g,j} \quad s.t. \quad E^g_{st} \geq \bar{E}_{st},$$

where g indexes subgroup, $j \in \{M, E\}$ indexes the test subject (Mathematics or ELA), A^g_{stj} is the proficiency rate in test j for subgroup g in school s and academic year t and T_{tj} is the target

[15] The dissimilarity index dictates the share of students that would have to be moved across schools in a district in order for all schools to have the same proficiency rate.

[16] For example, English learners are a fast-growing student demographic and struggle to meet testing standards, which could bias naïve estimates toward a positive relationship between AYP failure and enrollment growth.

proficiency rate. California sets the minimum eligible subgroup size, $\bar{E}_{st}$, according to the following piece-wise function of school enrollment:

$$\bar{E}_{st} = \begin{cases} 50, & \text{if } E_{st} < 334 \\ 0.15E_{st}, & \text{if } E_{st} \in [334, 666] \\ 100, & \text{if } E_{st} > 666. \end{cases}$$

The proficiency gap, X_{st}, serves as a running variable in the RD specification and proficiency failure, $F_{st} = \mathbb{1}\{X_{st} < 0\}$, captures the intent-to-treat assignment. Due to testing participation requirements and safe harbor exceptions, proficiency-based failure does not perfectly predict AYP failure.[17] Let F_{st}^{CDE} be the true assignment to treatment, as determined by the CDE. I estimate the following two-stage fuzzy regression discontinuity:

$$F_{st}^{CDE} = \theta F_{st} + f(X_{st}) + F_{st} \times h(X_{st}) + u_{st} \tag{1}$$

$$\%\Delta E_{s,t+1} = \beta \widehat{F_{st}^{CDE}} + j(X_{st}) + F_{st} \times k(X_{st}) + \epsilon_{st} \tag{2}$$

Equation (1) captures variation in the true assignment to failure that is driven by the running variable, the AYP proficiency gap. As in instrumental variables regression, the first stage includes exogenous covariates, which are a polynomial in the running variable that varies on either side of the treatment threshold. Equation (2) regresses the predicted failure status from the first-stage on the subsequent year percent change in enrollment, controlling for the same exogenous covariates. The parameter of interest, β, identifies the effect of just failing to meet the AYP achievement standard on enrollment in the subsequent year. Specifically, the failing a given year of AYP leads to a β percentage point change in enrollment growth the following year. I will also use the enter and exit variables as dependent variables to better understand enrollment mobility, while acknowledging that these measures are not as reliable as enrollment counts.

I examine AYP failure in several different samples to understand how students and schools responded to school choice independent of other policy effects. First, I estimate the causal effect of failing AYP in AY2001/02. This serves as a useful baseline, as failure at this stage does not result in any policy changes but may produce a school quality signal. Second, I limit the sample to

[17] Proficiency-based failure status corresponds to true failure status in 85% of schools in AY2001/02, 90% in AY2002/03 and 95% in AY2003/04.

schools that failed in AY2001/02 and estimate the effect of failure in AY2002/03, which triggered NCLB-mandated school choice, on subsequent enrollment growth in these schools.[18] The third sample captures the effect of school choice applied for two years. Leveraging the rule that schools had to pass AYP twice to have sanctions lifted, I restrict the sample to those schools that fail AYP in AY2001/02 and pass in AY2003/04. The marginally treated schools which failed in AY2002/03 would still be subject to the school choice sanction at the conclusion of AY2003/04, while the control group would never have had sanctions applied.[19] Finally, I compare those schools that failed AYP for two consecutive years across the threshold of failure in the third year of NCLB. Schools that failed for a third time would have added free tutoring and supplemental services on top of the school choice sanction, whereas schools that passed would only have school choice.[20] Investigating the effects of this additional sanction will shed light on whether responses vary depending on the nature of the accountability tool, and if so what kind of anticipatory effects may be present earlier on if agents accurately predict the sanction progression.

The four samples are used to estimate year-specific effects of NCLB policies as they were first applied to schools. They are summarized in the following table and will be referred to throughout by their numbering and treatment and control abbreviations. Histograms illustrating the distribution of the running variable in each sample are presented in Figure 1.

#	Year	Treatment & Control	Effect Estimated
1	01/02	Fail (F) vs. Pass (P)	Initial year failure (no sanctions)
2	02/03	F-F vs. F-P	School choice
3	02/03	F-F-P vs. F-P-P	School choice after 2 years
4	03/04	F-F-F vs. F-F-P	School choice + free tutoring vs. School choice only

Figure 2 provides a visual depiction of the first stage regression by plotting the average failure rate in 1 percentage point bins of the running variable. To the left of the cut-off, a few schools pass AYP despite failing based on proficiency as they are granted exemptions. To the right of the cut-off, a substantial fraction of schools fail AYP despite meeting the proficiency criteria, most

[18] Restricting to schools that failed in AY2001/02 is necessary since schools that passed could not be considered on the margin of receiving school choice in AY2002/03. In Section 6, I discuss alternative methods for estimating a fuller model of dynamic treatment effects.

[19] Restricting to schools that passed AYP in AY2003/04 prevents confounding effects with those from additional sanctions that are applied after a third year of failure. This approach also has the added benefit of using a sample of schools that is not under threat of additional sanctions, and thus is not confounded by possible anticipation effects. For clarity, I also report two-year effects unconditional on AY2003/04 AYP status in an appendix.

[20] It is worth noting that these services were to be prioritized for students from socio-economically disadvantaged backgrounds and low-performing subgroups. Supplemental services typically referred to the hiring of educational consultants to advise schools on how to improve achievement.

often because they fell below the 95% participation rate requirement.[21]

Table 2 reproduces the summary statistics from Panels A and B of Table 1 for schools within two percentage points of the running variable in each sample. Average school performance for samples 2 to 4 is generally lower since these samples select on prior school failure, as is reflected in Figure 1. These academically weaker schools are typically larger, have a greater share of Hispanic and economically disadvantaged students, and have higher student mobility and falling enrollments. Table 3 depicts the distribution of the lowest performing subgroup-test combinations for schools in all years and in each sample near the cut-off. It is most common for a school to perform lowest in its English learners subgroup writing the ELA test, and even more so in the samples where schools previously failed.

4.1 Identifying Assumptions

Given the samples selected for estimating the effects of school choice, it is important to be clear about the associated assumptions and why they should be expected to hold. There are two types of selection, depending on the sample: selection on prior failure histories (i.e. Sample 2: F-F vs. F-P and Sample 4: F-F-F vs. F-F-P) and selection on future failure histories (i.e. Sample 3: F-F-P vs. F-P-P). I will address each of these in turn following an introduction to the basic RD assumption.

Using the potential outcomes framework, where for simplicity $Y_{st} = \%\Delta Enrollment$ for school s from year t to $t+1$ and $D_{st} = F_{st}^{CDE}$, the treatment effect for Sample 1 is defined:

$$\beta = \frac{\lim_{g\to 0^-} \mathbb{E}[Y_{st}|X_{st}=x] - \lim_{x\to 0^+} \mathbb{E}[Y_{st}|X_{st}=x]}{\lim_{x\to 0^-} \mathbb{E}[D_{st}|X_{st}=x] - \lim_{x\to 0^+} \mathbb{E}[D_{st}|X_{st}=x]}$$

The parameter β identifies the local average treatment effect provided that the potential outcome for any treated, $Y_{st}(1)$, or control, $Y_{st}(0)$, school evolves continuously over the threshold of the

[21] It is worth commenting on the high fraction of schools failing to the right of the cut-off in Figure 2. First, this is not representative of the overall failure rate as I only plot points within 20 percentage points of the cut-off and each point does not represent the same number of observations. The especially high share of schools failing based on participation rates in the first year of the program (Sample 1) is consistent with the suddenness with which the policy was implemented. The act was signed into law in January 2002 and tests were written in the spring of that year.

running variable.[22] Mathematically:

$$\lim_{x \to 0^+} \mathbb{E}[Y_{st}(1)|X_{st} = g] = \lim_{x \to 0^-} \mathbb{E}[Y_{st}(1)|X_{st} = g]$$
$$\lim_{x \to 0^+} \mathbb{E}[Y_{st}(0)|X_{st} = g] = \lim_{x \to 0^-} \mathbb{E}[Y_{st}(0)|X_{st} = g]$$

I use the density manipulation test developed by Cattaneo et al. (2018) to show that this assumption holds for each sample. Figure 1 displays the resulting p-values from tests of the null hypotheses that there is no discontinuity in the density of the running variable at the cut-off. Although selection on prior AYP status alters the distribution of the running variable, the potential outcomes over the treatment threshold should remain continuous. That is, if a school in the selected sample failed regardless of which side of the cut-off they fell on their resulting enrollment growth should be a continuous function of the running variable. There is no clear reason why this should not continue to hold for observations with particular treatment histories.

The identifying assumptions for Sample 3 are somewhat more complicated:

$$\lim_{x \to 0^+} \mathbb{E}[Y_{s,t+1}(1)|X_{st} = x, F_{s,t-1} = 1, F_{s,t+1} = 0] = \lim_{x \to 0^-} \mathbb{E}[Y_{s,t+1}(1)|X_{st} = x, F_{s,t-1} = 1, F_{s,t+1} = 0]$$
$$\lim_{x \to 0^+} \mathbb{E}[Y_{s,t+1}(0)|X_{st} = x, F_{s,t-1} = 1, F_{s,t+1} = 0] = \lim_{x \to 0^-} \mathbb{E}[Y_{s,t+1}(0)|X_{st} = x, F_{s,t-1} = 1, F_{s,t+1} = 0]$$

That is, for schools which fail in $t = 1$ and pass in $t = 3$, enrollment growth from $t = 3$ to $t = 4$ does not jump discontinuously over the threshold of the running variable in $t = 2$. Essentially, schools to the right or left of the cut-off in year 2 would have to differ in a way that is not caused by year 2 failure but which affects enrollment growth after passing in year 3. If one believes the quasi-random assignment of schools across the threshold, this assumption should indeed hold, and is again supported by manipulation tests depicted in Figure 1, as well as robustness checks described in Section 6.

[22] The standard instrumental variables assumption of no defiers is likely satisfied here: schools who fail on other criteria when they meet proficiency rate standards should not also pass on exemptions if they were not to meet performance standards.

5 Results

5.1 Sample 1: Failure without sanctions

Beginning with the enrollment response to school failure in the first year of NCLB, Table 4 presents the RD estimates for Sample 1, where the first row represents the treatment effect on enrollment growth of just failing to make AYP based on the proficiency rate of the worst performing subgroup. Column 1 of Table 4 displays results using a linear function of the running variable, X_{st}, column 2 adds a cubic polynomial, column 3 reduces the bandwidth around the running variable to 2 percentage points[23], and column 4 presents nonparametric estimates from a local polynomial RD method developed by Titiunik (2014). The method implemented in column 4 is preferred, and columns 1 to 3 are included to demonstrate that the local polynomial method yields results that are consistent with parametric approaches. For a visual depiction of the result, Figure 3 plots averages of the dependent variable, percent change in enrollment, against the running variable, where each point contains several schools and every point represents approximately the same number of schools.[24] The figure also depicts a cubic polynomial in the running variable fitted on either side of the cut-off with 95% confidence intervals.

The coefficient found in the first row of column 4 says that just failing to meet test proficiency standards for NCLB caused enrollment growth to be 2.27 percentage points lower than in schools that just passed. This treatment does not involve any penalties for schools, and simply requires parents be notified that their students' school failed to make adequate progress. While students could in theory transfer schools or opt out of enrolling in a certain school at matriculation, failing AYP would not trigger the provision of transportation funds to reach alternative schools in their district. Thus, one can interpret this as the effect of a school quality signal to parents. This is consistent with Holbein's finding that student exit increases following AYP failure in general, though that paper does not examine how the effect varies based on failure history. The magnitude of the effect is surprising given the years-long history of standardized testing and reporting in California at this time, which should have provided ample resources for parents to learn about

[23] This is an arbitrary bandwidth. Coefficients are similar when restricting to 1 or 3 percentage points with a linear control function, and when using a mean squared error-optimal bandwidth selection method suggested by Titiunik (2014) with a quadratic or cubic control function.

[24] Note that this figure is representative of the estimate in column 2 of Table 4 but will not be identical as it ignores the first stage regression and only plots schools where proficiency-based failure and true failure corresponded. I also only display points that were within 20 percentage points of the cut-off for illustration.

school quality. Since this information shock has the potential to persist in later years of failure, it is important to keep this effect in mind when interpreting the results of failure-induced sanctions.

5.2 Sample 2: Failure triggering school choice

The results for Sample 2 with percent change in enrollment from 2002/03 to 2003/04 as the dependent variable are presented in Table 5. Figure 4 plots averages of the dependent variable where each point represents approximately the same number of schools, along with a cubic polynomial in the running variable fitted on either side of the cut-off. Although there is a small drop in enrollment growth for schools to the left of the cut-off, the data appear noisy. The small negative result of failure observed in Figure 4 is borne out in regression analysis. The estimates in the table indicate that the choice sanction had a statistically insignificant effect on enrollment growth in failing relative to passing schools, though the point estimate is consistently negative across specifications.

It would seem as though transfer take up was low under NCLB. Relative to the mean and standard deviation of enrollment growth for all schools in Sample 2, the effect, if it were statistically significant, is quite small: about -0.08σ. Replicating the table but controlling for baseline enrollment, the estimates are very similar. Overall, the evidence points to very little response to open enrollment under NCLB in California. This is consistent with anecdotal reports in several districts and general scepticism surrounding the policy and its enforcement.

One interpretation of this result in conjunction with that of Sample 1 is that parents respond to the initial informational aspect of school failure, and not the associated sanctions that come after additional failures. For this information story to be the valid, NCLB-mandated school choice would have to impose little change on the availability of transfers. Given the increase in funding for transportation of transfer students, one should expect that transfers at least became cheaper to parents post-NCLB. However, if school districts were reluctant to incorporate NCLB rules that differed from the prior transfer system, this lack of enforcement could explain the null effect of the policy change. I will explore whether there is suggestive evidence for this type of supply constraint in Section 7, looking at heterogeneous treatment effects.

As discussed previously, it is possible that expectations over future sanctions affect responses to ones already applied, meaning expectations about the second sanction, tutoring, could affect decisions made in response to the first, school choice. Further, while parents protested that they were not given adequate notice to transfer their children before the subsequent school year began, they would have at least had a second year to arrange these transfers, since it takes two years of a school passing for sanctions to be lifted. Investigating outcomes following failure in Samples 3 and 4 will shed light on the dynamic effects of multiple sanctions and whether additional time allowed for more choice-induced transfers.

5.3 Sample 3: Second year of school choice

Sample 3 includes schools that failed in first year and were thus eligible for school choice and passed in third year and were therefore not subject to escalating sanctions after the third year. By comparing these schools across the margin of failure in the second year of the program, the RD identifies the causal effect of being treated with school choice for two years but not being treated with additional sanctions, compared to not being treated with school choice at all, but nearly. It also potentially solves the issue of anticipatory effects of additional sanctions as I estimate outcomes *after* schools have passed and it is likely that parents expect them to continue to do so, alleviating the potential for future sanctions.

The results for Sample 3 are presented in Table 6, and Figure 5 depicts the results graphically. The effect of receiving school choice for a second year, without additional sanctions, relative to no choice at all was to decrease enrollment growth from AY 2003/04 to 2004/05 by 2.71 percentage points, or about a third of a standard deviation. The coefficient estimate is statistically significant at the 5% confidence level. Multiplying this coefficient by the average enrollment in schools near the cut-off in this sample (from Table 3) gives an enrollment decline of $0.027 * 693 \approx 19$ students. Estimating the regression in column 4 using enrollment change instead of percent change yields a statistically significant coefficient of -13.8 students.

The negative effect of failure on enrollment is larger in magnitude in the second year of transfer availability and when schools are no longer under threat of future sanction. It is worth noting that this is the effect of the enrollment change following a year in which both the treatment and control group *passed* AYP. It is possible that this effect simply captures the delayed response to transfer availability, which was not readily available the year before as intended. Alternatively, one might infer from this that the anticipatory effects of future sanctions were actually positive, biasing the year one estimate upwards. That is, parents may have kept their children in failing schools to take advantage of the potential third year sanction, free tutoring and supplemental education services.

Analyzing Sample 4 will provide insights into whether this was indeed the case.

5.4 Sample 4: Beyond school choice - tutoring sanction

If schools that failed for two consecutive years passed AYP in the third year, then parents who were hoping for free tutoring would be disappointed but still have the option to transfer. Thus, if it is indeed the case that the anticipatory effects of the third sanction were positive, it should follow that third year failure was associated with enrollment growth relative to passing. Testing this conjecture, Table 7 displays estimates of the enrollment response to the third consecutive year of failure for Sample 4 (schools which failed the first two years of NCLB).[25] Consistent with a theory of positive associations with the second sanction, the third year of consecutive failure was associated with greater enrollment growth relative to schools that failed for two years before passing. It is difficult to imagine an alternative explanation for why enrollment would grow by nearly 3 percentage points after schools just fail for a third time. This suggests that at least some parents are knowledgeable of details of the NCLB policy. Section 7 further examines this result by looking at which types of students were more responsive to the policy in this way.

5.5 District Case Study

Changes in school-level enrollment are informative insofar as competition for enrollment is intended as the rising tide that lifts all boats when families are provided more choice (Hoxby, 2003). The other key pathway for choice to improve student outcomes is access to more effective schools through transfers. To better understand whether students are able to take advantage of the NCLB transfer policy in this way, I use student-level data from a large urban district as a case study, which I will refer to as the 'District'. The District is includes a large number of schools which are more densely situated, and students are more likely to be free- or reduced-price lunch eligible, English learners, and perform worse on standardized tests. It enrolls a large share of students in California public primary schools, giving it outsize importance in the state's education landscape. Given large discrepancies in school performance and short distances between schools, one might expect a greater choice-induced demand response in this setting. On the other hand, if there is strong spatial correlation in school performance, the supply of high quality alternatives may be limited. I

[25] Note that both the treatment and control schools will still have access to the transfer option, given the two-year passing requirement for sanctions to be lifted.

use data on the number of seats available for open transfers at each school for the 2003/04 school year to investigate these supply channels.

I replicate my preferred specification from column 4 of Tables 4 to 7 replacing the dependent variable with an indicator variable measuring whether a student transfers to another school in the district next year. Standard errors are clustered at the school level. Results are presented in Table 8. As the first year of data is AY 2002/03, I am not able to estimate the first-year effect of AYP failure on transfers, and instead examine the second-year effect in column 1. The results indicate a positive but small and statistically insignificant effect of failing AYP in 2001/02 on transferring to a new school for the 2003/04 school year. While this suggests there was no delayed response to initial AYP failure, a less muted effect may have been detected outside the scope of the data in 2002/03.

The effects of NCLB-induced choice are presented in column 2 for one-year (Sample 2) effects and column 3 for two-year (Sample 3) effects. The local average treatment effect is a decrease of 2.6 percentage points in the likelihood of transferring schools as a result of open enrollment in the District. This result is counterintuitive, and contrasts with small negative effects on enrollment growth documented in the state as a whole.[26] The result could reflect students holding off on transfers in order to obtain the benefits of later sanctions. However, column 3 indicates that the effects remain negative, though statistically insignificant, a year later after schools pass and no additional sanctions are forthcoming. Column 4 presents estimates of the effect of marginally failing a third year of AYP and receiving escalating sanctions.[27] The negative coefficient estimate in column 4 is consistent with positive enrollment effects documented in the state, though it is small and statistically insignificant. Given muted reactions to these additional sanctions, or lack thereof, it seems unlikely that the reduction in transfers caused by second year AYP failure are driven primarily by demand side factors.

Several supply-side factors may limit the ability of open enrollment to provide meaningful transfer opportunities to students in failing schools. For students to gain access to better schools

[26] The effects on enrollment growth measured for the District alone are qualitatively similar to those in California as a whole, though smaller and not statistically significant, likely due to the smaller sample of schools. Coinciding enrollment declines with fewer out-transfers can be explained by fewer new students switching or matriculating into the school the following year (result not shown but available upon request).

[27] This is difficult to measure in the District as few schools that failed the first two years were able to get over the AYP threshold in the third year. With insufficient data for executing estimating the preferred nonparametric specification, I instead estimate a parametric regression function with a cubic polynomial in the running variable equivalent to the specification in the third column of the main results tables.

through transfers, there must be higher performing schools in the district, within feasible traveling distance, and which have the capacity to accept new students. In April 2003, 110 elementary and middle schools announced that they had a total of 3314 seats available to receive open enrollment transfers – enough for 1 in 100 eligible students. Only 1,491 (45%) of those seats were in schools that had passed AYP in 2002/03. In total, 2,666 students transferred to passing schools offering open enrollment seats, with fewer than half of these students coming from failing schools. With such scarcity in open enrollment spots, why would students in failing schools, who were to be prioritized, not take up the majority of these transfers? To address this puzzle, Figure 7 displays distributions of several measures of local schooling options for students in schools that passed and failed AYP in 2003. Though the total number of open seats within a 5km radius was not much higher at passing relative to failing schools, the number of nearby passing schools and open seats in those schools was much lower for the students who were meant to be prioritized for school choice. Thus, the well-intended provision of school choice under NCLB failed to overcome existing barriers to accessing high quality schools in the District.

6 Robustness

The results are robust to several tests of the identifying assumptions and sample restrictions, in addition to manipulation tests described in section 4. I conduct additional tests for suggestive evidence of manipulation of the running variable by checking for the presence of discontinuities in co-variates across the treatment threshold. This is accomplished by replacing the dependent variable in all regression specifications with one of the co-variates: total enrollment, a three-year pre-trend in school enrollment, a three-year pre-trend in district enrollment, share of the student body of various demographic or socio-economic subgroups, the number of schools in a district, and the distance to the nearest school. Graphical depictions of these results are included in Appendix Figure 91 to Figure 94. The only discontinuity uncovered by this exercise was in the share of African American students in Sample 1. Given the large number of co-variates tested, it is possible this result is spurious. Even so, the results are highly robust to the inclusion of any of these co-variates in regression specifications. A small exception is in the sample 3 regression where controlling for the school enrollment pre-trend increases standard errors enough to lose statistical significance.

6.1 Alternative Methods for Addressing Dynamic Treatment Effects

In order to capture the dynamic effects of failing one or more consecutive years of AYP, one could use the running variables from several years in one specification that employs a multidimensional RD (MDRD) design, as in Gilraine (2020). This could be done instead of analyzing a subset of schools with a particular failure history as in Sample 2.[28] Abstracting from the fuzzy design, one could estimate:

$$\begin{aligned} Y_{st} = \alpha + \tau_1 F_{st} + f(X_{st}, X_{s,t-1}) + \theta_1(f(X_{st}, X_{s,t-1}) * F_{st}) \\ + F_{s,t-1}[\tau_2 + \tau_3 F_{st} + \theta_2 f(X_{st}, X_{s,t-1}) + \theta_3(f(X_{st}, X_{s,t-1}) * F_{st})] + \epsilon_{st} \end{aligned}$$

where the estimate of interest, the effect of failing in both years and being treated with school choice relative to failing in only the first year, is:

$$\mathbb{E}[Y_{st}|F_{st} = 1, F_{s,t-1} = 1] - \mathbb{E}[Y_{st}|F_{st} = 0, F_{s,t-1} = 1] = \tau_1 + \tau_3 + (\theta_1 + \theta_3) f(X_{st}, X_{s,t-1})$$

Unfortunately, the multidimensional RD is highly demanding on the data and estimating this with a simple linear control function results in estimates with very large standard errors. Instead, my strategy is to subset the data to schools where $F_{s,t-1} = 1$ and estimate the regression and effect of interest:

$$\begin{aligned} Y_{st} = \gamma_1 + \beta_1 F_{st} + h(X_{st}) + \phi_1(h(X_{st}) * F_{st}) + \nu_{st} \quad for \quad F_{s,t-1} = 1 \\ \mathbb{E}[Y_{st}|F_{st} = 1, F_{s,t-1} = 1] - \mathbb{E}[Y_{st}|F_{st} = 0, F_{s,t-1} = 1] = \beta_1 + \phi_1(h(X_{st})) \end{aligned}$$

An important distinction between β_1 and $\tau_1 + \tau_3$ is that the former is not identified using a multidimensional control function, thus assuming that $X_{s,t-1}$ is not an important determinant of Y_{st}, either conditional on or interacting with X_{st}. Further, an approach recommended by prior papers implementing MDRD designs is to restrict to observations close to the cut-off of *both* running variables (Gilraine, 2020; Papay et al., 2014). This is an important consideration since schools near the cut-off in second year could have been far below the cut-off in first year, meaning that the local average treatment effects estimated for Samples 1 and 2 are identified off very different schools.

[28] While I focus on Sample 2 for exposition, the discussion of estimating dynamic treatment effects either by implementing MDRD or by selecting on prior treatment histories extends to Sample 4 (F-F-F vs. F-F-P).

Given the relevance of these techniques, I replicate my main results controlling for the prior year running variable and limiting observations based on various bandwidths of same. Furthermore, I display the results for a full set of treatment histories, both conditional and unconditional on prior failure status, to give the reader a full picture of the samples included and excluded from the main analysis. These results are presented in Appendix Table 101 and Table 102. Results in columns 1, 5 and 9 correspond to estimates associated with samples 1, 2 and 4, respectively.

The key results are robust to this bandwidth restriction and, for the most part, the tables depict that little is happening to enrollment following failure in the samples ignored by the main analysis. Though it appears as though second year failure following passing first year led to large decreases in enrollment growth (column 4), these effects disappear when limiting to schools that were near the cut-off the prior year. This is an intuitive result since the latter regression excludes schools those that saw large declines in performance from well above the cut-off to near it. The result in column 7 of Table 102 is somewhat concerning as it suggests that failing a second consecutive year caused enrollment growth instead of declines, despite triggering the open enrollment sanction for the first time. However, given there are so few schools from which this is identified, I suspect the result is not generalizeable.

7 Heterogeneous Treatment Effects

The main results prompt several questions. Who are the students that leave? Are they the disadvantaged students, for whom the program was intended? Is there evidence of cream-skimming, where schools will only take high performing, perhaps socio-economically advantaged students? Are there more students transferring where the costs of doing so are lower, in terms of distance between schools? Does dissimilarity between schools in a district predict greater transfers, as would be expected from the demand side, or fewer? These questions seek to explain the factors driving demand and supply of transfers under this policy. Answering them has broader implications for achieving the goals of open enrollment programs while minimizing their potential for unintended consequences.

To address the questions outlined above, this section presents estimates of heterogeneous treatment effects across student, school, and district characteristics. I reproduce the main results using enrollment of different subgroups of students as the dependent variable as well as for subsets of the data parsed by values of moderating variables. To identify causal heterogeneous treatment effects would require these moderating variables to be exogenous to enrollment growth, conditional on the running variable, X_{st}.[29] I discuss the plausibility of this assumption as I interpret the results. For several of the moderating variables I use this is unlikely to hold and findings should be taken as suggestive. Further, I do not adjust standard errors in this section for multiple testing as doing so in the non-parametric setting is not straightforward. Given the large number of estimates computed, adjusted standard errors are likely much larger, motivating added caution in making inferences from these results.

7.1 Heterogeneity across students

The NCLB Act stipulates that the resources allotted for implementing sanctions, such as transportation support for transfers and funding for tutoring services, should be prioritized towards socio-economically disadvantaged and academically struggling students. If enrollment changes caused by school failure are capturing student transfer through this program and not simply activist parents responding to signals about school quality, one should expect to see greater take up among socio-economically disadvantaged and lower performing students.

Examining the first dimension, Table 9 reproduces the main results for each sample using the standard enrollment percent change as the dependent variable in the first row, replacing this with enrollment percent change for socio-economically disadvantaged students as the dependent variable in the second row, and with the share of the school that are socio-economically disadvantaged in the third row, controlling for baseline shares to improve precision.[30] The results for socio-economic students differ from those for all students in several important ways. From column 1, it appears as though this group sees little enrollment response to the first year of failure, perhaps because they had fewer resources to transfer to other schools without additional help from the district. The point estimates for the initial response to school choice in Sample 2 are statistically insignificant, though in the opposite direction as one would expect. However, the large and statistically significant negative effect in column 3 suggests that with more time and after no more sanctions were forthcoming socio-economically disadvantaged students did in fact exit or cease enrolling in failing schools.

[29] It is also necessary to satisfy no bunching near the cut-off in regressions on these moderating variables, all of which passed this test as described in Section 6.

[30] It is possible that students are not consistently identified as socio-economically disadvantaged across years. While I would conjecture that the number changing status year to year is small, particularly since parents are unlikely to change their education levels, I do not have data to back this up.

Despite being prioritized for tutoring services, column 4 indicates that the enrollment boost for schools failing a third consecutive year was not coming from disadvantaged students. Replicating these results with a restricted sample of only Title I schools (not shown), the results in columns 1 to 3 are largely unchanged and the point estimate in column 4 is a statistically significant 3.3 percentage point increase in enrollment growth. This is suggestive that this result is driven by a desire to receive tutoring services, since it would only be offered in Title I schools.

Turning to the third row, it does not appear as though any of these results translated into significant changes in the share of students who were disadvantaged at a failing school, and this is true when restricting to Title I schools as well. Note that the enrollment figures in Table 9 are based on tested students, as the socio-economically disadvantaged category is not provided in data on total school enrollments. If instead I use enrollment of free- or reduced-price lunch eligible students[31] results are largely unchanged.

In Table 10, the data are transformed such that the unit of observation is the subgroup-school combination as opposed to the school. I include five subgroups of interest which are typically well-represented: African American, Asian, Hispanic, white and socio-economically disadvantaged. I do not include English learners as this group is largely redundant when combined with Hispanic and disadvantaged groups. The baseline results estimated at the school level are provided in the first row for reference. The replicated baseline results, controlling for subgroup, are in the second row. Estimated at the subgroup level, enrollment changes in Samples 3 and 4 are much larger in magnitude than for schools as a whole, and those for Sample 1 are somewhat smaller and lack statistical significance. This could be due to the particular selection of subgroups or the double-counting of students across the racial and economically disadvantaged subgroups.

The key point of the table is to compare these baseline results to the results across terciles of the ratio of a subgroup's proficiency rate to the lowest performing subgroup proficiency rate in their school. This is a measure of how well performing a student was relative to the students who were the intended beneficiaries of NCLB sanctions.[32] For subgroups who performed in the bottom tercile, which will include those who were the lowest performing, the enrollment effect for the second year of school choice is substantially larger in magnitude relative to the baseline, and for the tutoring sanction is somewhat smaller. On the other end, subgroups which performed much

[31] These measures, as well as Title I status, come from the National Center for Education Statistics database.

[32] Appendix Table 103 lists the values taken by the moderating variable in each tercile and sample, for this and other moderating variables used throughout Section 7.

better than the lowest achieving group appear to exit more than baseline after a school fails for the first time and less when a school is sanctioned with tutoring services. This last effect in column 4 is counter-intuitive since these students would not generally be targeted to receive tutoring. However, the finding disappears if I drop the economically disadvantaged subgroup, another prioritized group for tutoring funds.

Overall, there is suggestive evidence that more economically and academically advantaged students were the ones to exit after a school failed in the first year of NCLB. On the other hand, disadvantaged students responded more to the school choice aspect of the policy, as intended, but still only after the choice was available for a second year and no more sanctions were forthcoming.

7.2 Heterogeneity across schools and districts

Although differential take up of NCLB-induced transfers may be associated with differences in *school* characteristics, these tell us about decisions made at the *district* level since it is districts who have discretion over who gets to transfer where. There are several reasons why school districts might wish to prevent or discourage NCLB transfers. First, funding to transport transfer students was to be diverted from Title I funds, generating costs for districts who chose to obey the law. Though failing schools were required to use at least 5% of Title I funds for this purpose, there is little evidence that the federal Department of Education enforced this aspect of the law (EdSource, 2004). Districts may be more willing to allow students to transfer to nearby schools, though with my data it is not possible to disentangle this supply side response from parental demand for closer schools. Second, if schools in a district are highly polarized in their test score performance, administrators who care about the academic reputation of high achieving schools may be reluctant to accept low performing students into such schools.

To investigate these possible responses, I estimate heterogeneous treatment effects using the moderating variables: distance to the nearest other school in the district (Table 11), distance to the nearest passing school in the district (Table 12), ratio of own school performance to performance of the nearest passing school (Table 13), and a dissimilarity index in ELA and mathematics proficiency rates (Table 14). As in Table 10, all moderating variables are broken into terciles and values for each tercile are provided in Appendix Table 103.[33]

[33]The level of disaggregation was chosen to balance sample size and the ability to discern heterogeneous effects. Results across quartiles and medians are available upon request.

Table 11 reveals that the effects of AYP failure on enrollment changes in each sample are largest in magnitude and statistical significance in districts where schools are further away from one another.[34] This is counter-intuitive given the policy design, however the moderating variable could correlate to other aspects of high- and low-density school districts that predict stronger failure responses.

Limiting to districts where there was another passing school, Table 12 presents heterogeneous treatment effects for each sample across terciles of the distance to the nearest passing school in the district, which would be the nearest school to which students may have transferred under NCLB school choice. Trends are difficult to discern from these results given the small sample sizes after limiting to schools which had other passing schools in their district. Nonetheless, it appears as though being closer to other passing schools was associated with more enrollment declines in the second year of school choice (Sample 3).

To ascertain whether relative performance attracted students to nearby passing schools, I employ the academic performance index (API). This is a composite index of test scores compiled by CDE that was used in an antecedent accountability program and remained widely published in school report cards. Schools in the top tercile should most induce transfer demand as they are the worst performing relative to nearby alternatives. However, according to Table 13, middle tercile schools see the greatest enrollment changes, perhaps indicating that there is some push back from high performing schools or that they are overdemanded. Though I cannot disentangle these mechanisms, either one represents a limitation of NCLB-style open enrollment.

Examining if this pattern in relative school performance holds more widely, Table 14 estimates heterogeneous treatment effects across a district-level measure of school performance dissimilarity. The dissimilarity index dictates the share of students that would have to be moved across schools in a district in order for all schools to have the same proficiency rate. Similar to results in Table 13, the middle tercile sees the largest enrollment effects in Sample 3. Otherwise, the only statistically significant estimate is for schools in highly similar districts according to ELA proficiency, which have somewhat larger positive enrollment effects from tutoring services, though the differences across terciles appear small.

[34] In this and other school heterogeneity measures, schools that were the only one in their district are excluded. Effects for this group of schools were estimated separately but yielded very large standard errors as there were observations.

8 Conclusion

This study finds that the initial triggering of school choice under NCLB did not affect enrollment growth, though after students had the ability to transfer for a second year some did take up this opportunity. This late-onset take-up appears to mostly come from socio-economically disadvantaged students, indicating that the program had the intended effect of reducing the cost of transfers for students who had the fewest resources to do so on their own. On the other hand, the opposite group of students left schools after they failed the first year of NCLB: economically and academically advantaged students declined enrollment in failing schools without receiving any funding for transfers from the district.

A less central, yet relevant, finding of this research is that parents responded to the NCLB sanction of tutoring and supplemental services with an increase in enrollment. This was likely attributable to a decline in out-transfers, though exits are imprecisely estimated in the data and cannot explain the entire enrollment effect. The fact that parents were savvy enough to adjust their school choice decisions based on the potential for tutoring suggests that NCLB-mandated notifications letters were successful at promoting salience of the policy details. In regard to the delay in transfer take-up, it may simply be the case that the letters were sent too late in the fall of 2003 to allow time for the bureaucratic process of transferring.

Although the open enrollment program appears to eventually take effect as intended, the estimate of the number of students to transfer as a result of the policy is not substantial: between 10 and 20 out-transfers for an average school population of about 700 students.[35] It is difficult to say whether this would exert sufficient pressure on the school to increase their productivity. It does seem unlikely that the loss of this many peers would produce noticeable externalities on students who remain in failing schools, alleviating concerns about schools becoming worse off due to a loss in high achieving students. However, the lack of stronger effects for schools which had much better nearby passing schools relative to those that did not suggests that this small take-up could be coming from supply constraints. A similar pattern is implied by results of heterogeneous treatment effects across measures of within-district dissimilarity in school performance. It is possible that districts prevented students from transferring by simply denying them, or there may have been a lack of space and funds to expand overdemanded schools, as administrators claimed. More data

[35]Similar numbers are estimated for students who exit after first year failure.

are needed to shed light on these competing narratives.

Without the ability to track individual students across schools and over time, it is difficult to come to conclusions regarding choice-induced mobility of students and the mechanisms influencing these mobility decisions. Nonetheless, the quasi-randomly identified RD estimates present compelling evidence that students and parents did respond to NCLB sanctions in ways that could be expected. However, results from heterogeneous treatment effect estimation do not follow patterns that would be expected from student demand, potentially suggesting that supply side constraints did influence the ability of students to transfer under NCLB-induced open enrollment. This may point to a broader implementation flaw that explains why the policy did not generate economically large changes in student enrollment. If it is indeed the case that NCLB-induced choice did little to improve test scores in California, as Ahn and Vigdor (2014) find to be the case in North Carolina, barriers to transfer take up could represent a major impediment to achieving these gains. Understanding the link between these outcomes is vital to designing successful open enrollment policies in the future.

References

Ahn, T., & Vigdor, J. (2014). *The impact of no child left behind's accountability sanctions on school performance: Regression discontinuity evidence from north carolina* (Working Paper No. 20511). National Bureau of Economic Research. https://doi.org/10.3386/w20511

Avery, C., & Pathak, P. A. (2021). The distributional consequences of public school choice. *American Economic Review*, *111*(1), 129–52. https://doi.org/10.1257/aer.20151147

Betts, J. R., & Loveless, T. (2005). School choice, equity, and efficiency. *Getting choice right: Ensuring equity and efficiency in education policy* (pp. 1–13). Brookings Institution Press.

Betts, J. R., Rice, L. A., Zau, A. C., Tang, Y. E., & Koedel, C. (2006). *Does school choice work? effects on student integration and achievement* (tech. rep.). Public Policy Institute of California.

Bruhn, J. (2020). *The consequences of sorting for understanding school quality* [Manuscript].

Cattaneo, M. D., Jansson, M., & Ma, X. (2018). Manipulation testing based on density discontinuity. *The Stata Journal*, *18*(1), 234–261. https://doi.org/10.1177/1536867X1801800115

Chakrabarti, R. (2014). Incentives and responses under no child left behind: Credible threats and the role of competition. *Journal of Public Economics*, *110*, 124–146. https://doi.org/https://doi.org/10.1016/j.jpubeco.2013.08.005

Dee, T. S., & Jacob, B. (2011). The impact of no child left behind on student achievement. *Journal of Policy Analysis and Management*, *30*(3), 418–446.

Deming, D. J., Hastings, J. S., Kane, T. J., & Staiger, D. O. (2014). School choice, school quality, and postsecondary attainment. *American Economic Review*, *104*(3), 991–1013.

EdSource. (2004). *No child left behind in california? the impact of the federal nclb act so far* (tech. rep.) [https://edsource.org/wp-content/publications/NCLB_1-04.pdf].

Gilraine, M. (2018). *School accountability and the dynamics of skill formation.*

Gilraine, M. (2020). A method for disentangling multiple treatments from a regression discontinuity design. *Journal of Labor Economics*, *38*(4), 1267–1311.

Grady, S., Bielick, S., & Aud, S. (2010). *Trends in the use of school choice: 1993 to 2007* (tech. rep.) [(NCES 2010-004). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education]. National Center for Education Statistics.

Hastings, J. S., Kane, T. J., & Staiger, D. O. (2006). *Preferences and Heterogeneous Treatment Effects in a Public School Choice Lottery* (NBER Working Papers No. 12145). National Bureau of Economic Research, Inc. https://ideas.repec.org/p/nbr/nberwo/12145.html

Hastings, J. S., & Weinstein, J. M. (2008). Information, school choice, and academic achievement: Evidence from two experiments. *The Quarterly Journal of Economics*, *123*(4), 1373–1414.

Hemelt, S. W. (2011). Performance effects of failure to make adequate yearly progress (ayp): Evidence from a regression discontinuity framework. *Economics of Education Review*, *30*(4), 702–723.

Hoff, D. F. (2006). Groups press for enforcement on nclb choice option. *Education Week*, *26*(4), 28–30. https://www.edweek.org/ew/articles/2006/09/%5C%5C20/04choice.h26.html

Holbein, J. (2016). Left behind? citizen responsiveness to government performance information. *The American Political Science Review*, *110*(2), 353–368.

Hoxby, C. M. (2003). School Choice and School Productivity. Could School Choice Be a Tide that Lifts All Boats? *The Economics of School Choice* (pp. 287–342). National Bureau of Economic Research, Inc.

Krieg, J. M. (2008). Are Students Left Behind? The Distributional Effects of the No Child Left Behind Act. *Education Finance and Policy*, *3*(2), 250–281.

Krieg, J. M. (2011). Which Students Are Left behind? The Racial Impacts of the No Child Left Behind Act. *Economics of Education Review*, *30*(4), 654–664.

Lovenheim, M. F., & Walsh, P. (2018). Does choice increase information? evidence from online school search behavior. *Economics of Education Review*, *62*, 91–103. https://doi.org/https://doi.org/10.1016/j.econedurev.2017.11.002

Macartney, H., McMillan, R., & Petronijevic, U. (2015). *Incentive design in education: An empirical analysis* (Working Paper No. 21835). National Bureau of Economic Research.

Neal, D., & Schanzenbach, D. W. (2010). Left Behind by Design: Proficiency Counts and Test-Based Accountability. *The Review of Economics and Statistics*, *92*(2), 263–283.

Papay, J. P., Murnane, R. J., & Willett, J. B. (2014). High-school exit examinations and the schooling decisions of teenagers: Evidence from regression-discontinuity approaches. *Journal of Research on Educational Effectiveness*, *7*(1), 1–27.

Polikoff, M. S., & Wrabel, S. L. (2013). When is 100% not 100%? The Use of Safe Harbor to Make Adequate Yearly Progress. *Education Finance and Policy*, *8*(2), 251–270.

Reback, R., Rockoff, J., & Schwartz, H. L. (2014). Under pressure: Job security, resource allocation, and productivity in schools under no child left behind. *American Economic Journal: Economic Policy*, *6*(3), 207–41.

Titiunik, R. (2014). Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, *14*(4), 909–946(38).

Wong, M., Cook, T. D., & Steiner, P. M. (2009). *No child left behind: An interim evaluation of its effects on learning using two interrupted time series each with its own non-equivalent comparison series* (Working Paper WP-09-11). Institute for Policy Research, Northwestern University.

Figures

Figure 1: Histograms of the Running Variable, X_{st} by Sample

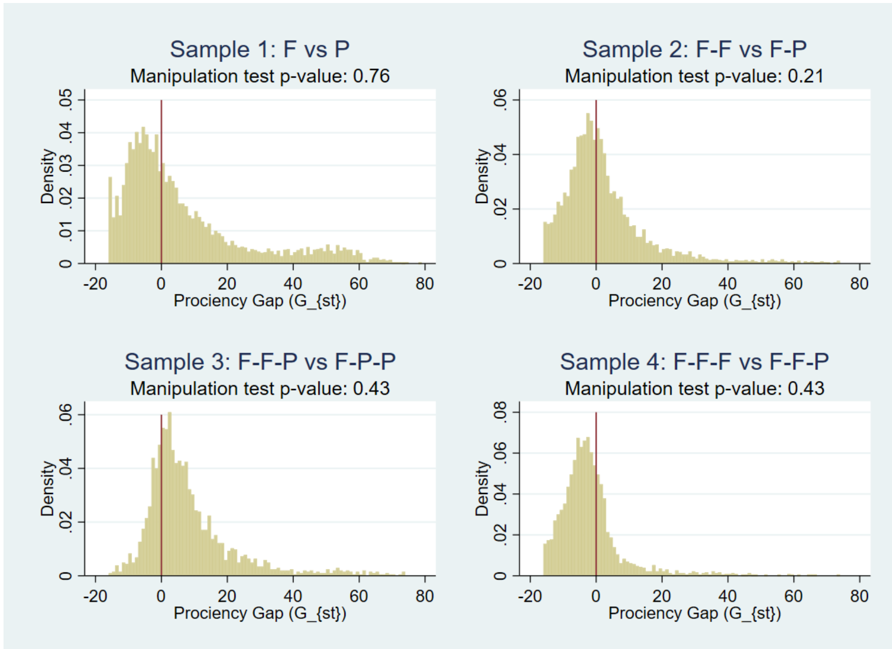
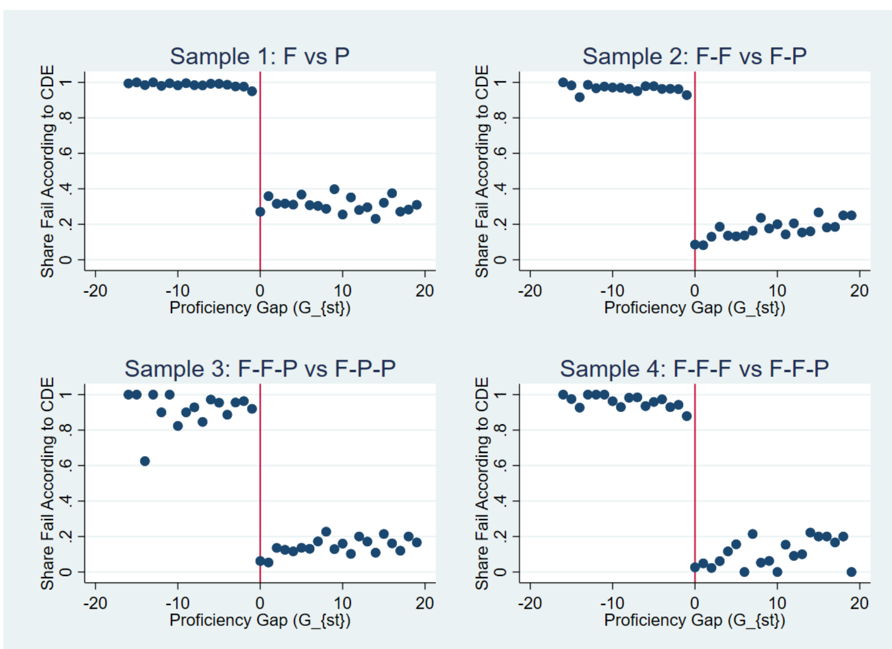
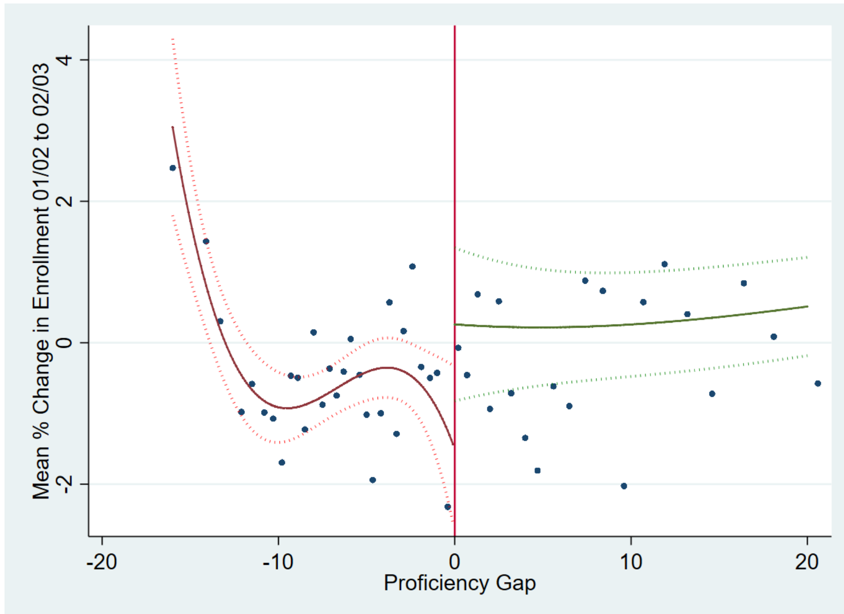


Figure 2: First Stage Illustration by Sample



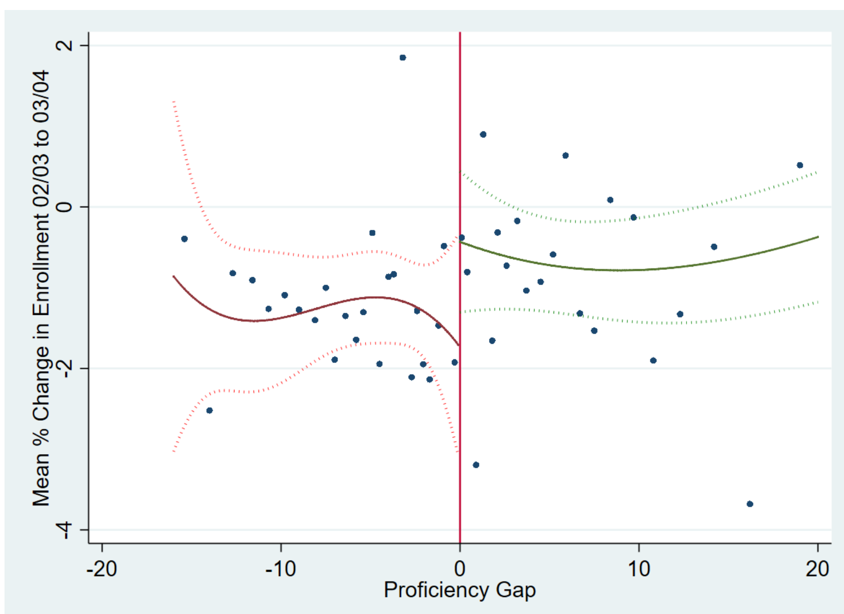
Notes: Schools are binned into 1 percentage point intervals of the running variable. The failure rate for each bin is plotted against the running variable, the proficiency rate gap. Schools to the left of the threshold $X_{st} = 0$ failed the proficiency requirement of NCLB while schools to the right passed.

Figure 3: Sample 1 - Year 1 Failure Effect on Enrollment Growth (no sanctinos)



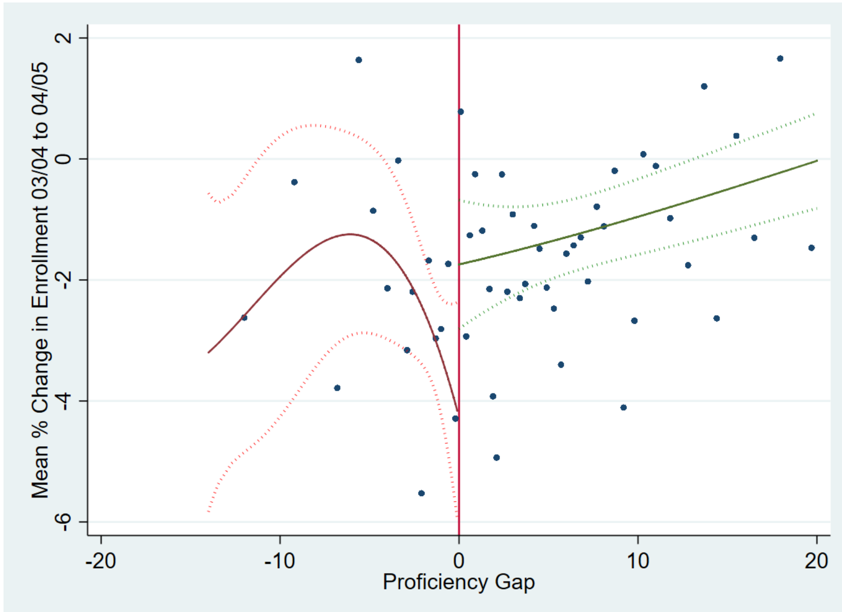
Notes: Figure includes schools with a proficiency gap that was between -20 and 20 percentage points. The figure plots a third order polynomial in the running variable fitted on either side of the cut-off with 95% confidence intervals. Points plot the average enrollment growth of schools that fall in an interval of the running variable, where intervals are defined such that points have approximately the same number of schools.

Figure 4: Sample 2 - School Choice Effect on Enrollment Growth (1st Year)



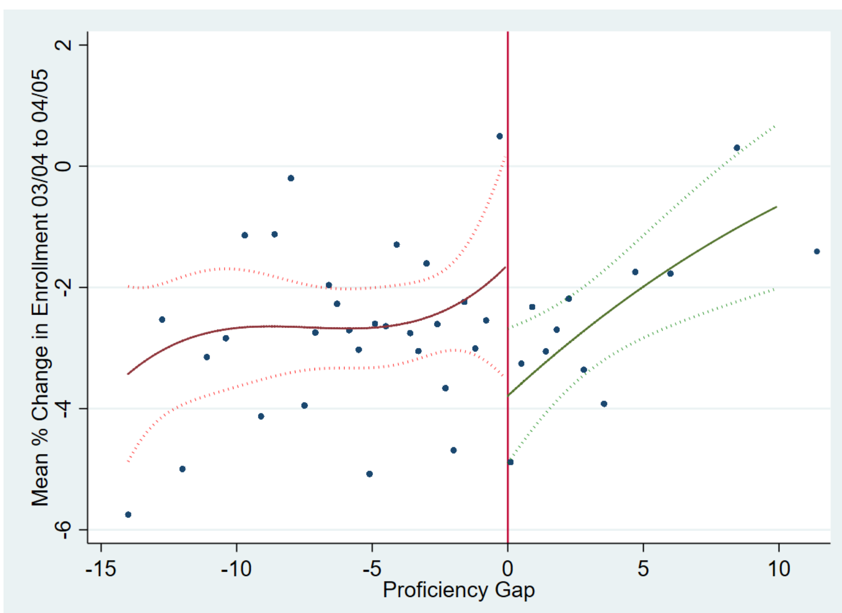
Notes: Figure includes schools which failed and year 1, and for which the proficiency gap was between -20 and 20 percentage points. The figure plots a third order polynomial in the running variable, proficiency gap, fitted on either side of the cut-off with 95% confidence intervals. Points plot the average enrollment growth of schools that fall in an interval of the running variable, where intervals are defined such that points have approximately the same number of schools.

Figure 5: Sample 3 - School Choice Effect on Enrollment Growth (2nd Year)



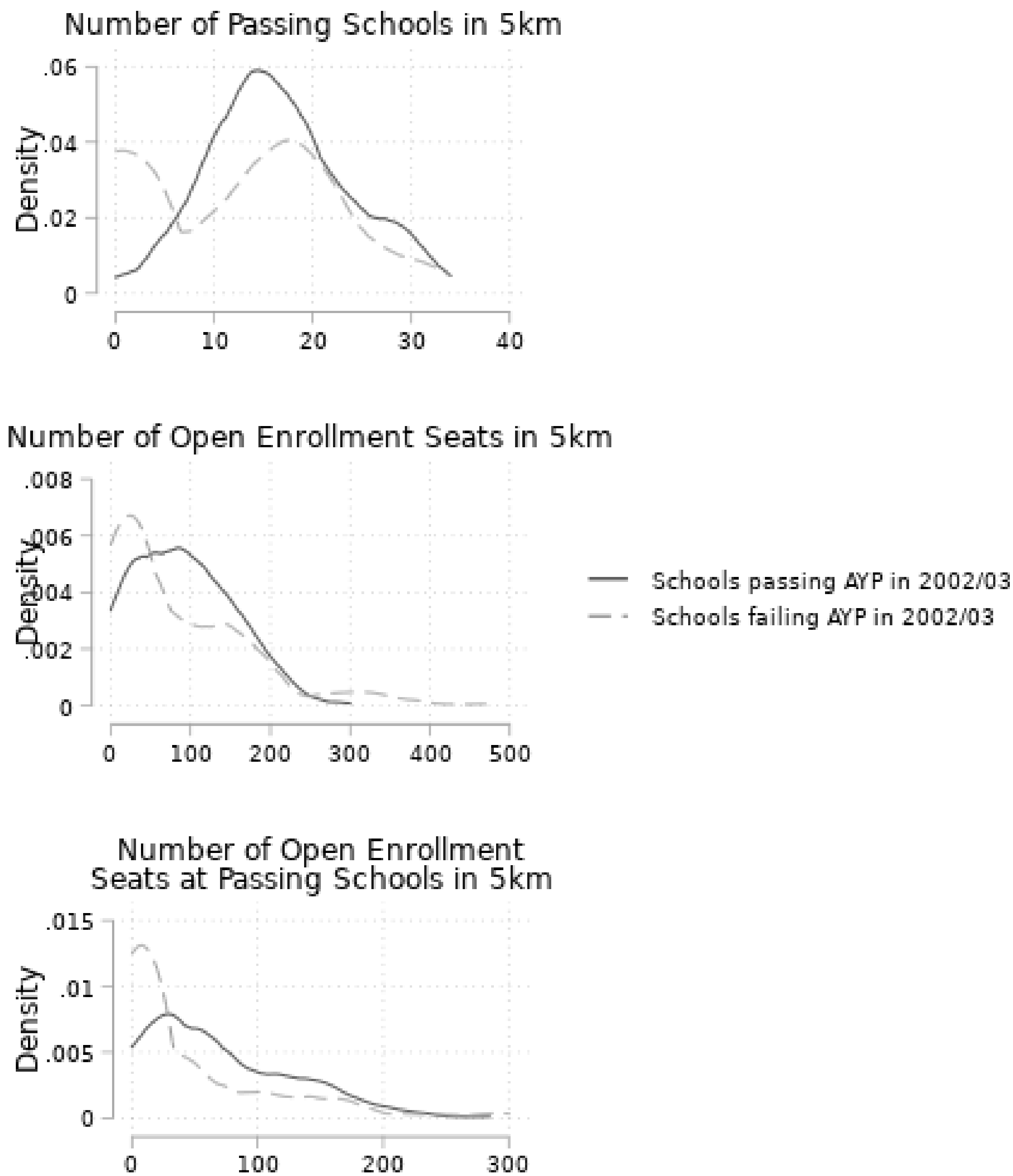
Notes: Figure includes schools which failed in year 1 and passed in year 3, and for which the proficiency gap was between -20 and 20 percentage points. The figure plots a third order polynomial in the running variable fitted on either side of the cut-off with 95% confidence intervals. Points plot the average enrollment growth of schools that fall in an interval of the running variable, where intervals are defined such that points have approximately the same number of schools.

Figure 6: Sample 4 - Tutoring and School Choice Effect Relative to School Choice Alone



Notes: Figure includes schools which failed and Years 1 and 2, for which the proficiency gap was between -20 and 20 percentage points. The figure plots a third order polynomial in the running variable, proficiency gap, fitted on either side of the cut-off with 95% confidence intervals. Points plot the average enrollment growth of schools that fall in an interval of the running variable, where intervals are defined such that points have approximately the same number of schools.

Figure 7: Distribution of Transfer Options in Case Study District, by 2002/03 AYP Status



Notes: Each figure includes two kernel density plots of the relevant measure: one for the sample of schools that passed AYP in 2002/03 and one for the sample of schools that failed. All schools failed AYP in 2001/02, and thus would have been subject to school choice if they failed in 2002/03. All variables are measured in 2002/03.

Tables

Table 1: Summary Statistics (AY 2001/02 to 2005/06)

Panel A: California Schools	Mean	(St. Dev.)
Total School Enrollment	660	(363)
Test Enrollment	512	(347)
Share African American	0.08	(0.12)
Share Asian	0.08	(0.12)
Share Hispanic	0.43	(0.29)
Share White	0.36	(0.28)
Share Other	0.05	(0.07)
Share Socio-Economically Disadvantaged	0.52	(0.31)
Year-to-year Enrollment Percent Change	-0.6	(9.8)
Year-to-year Enrollment Absolute Change	-7.7	(62.0)
Panel B: California Districts		
Number of schools in District	7	(21)
Share of Districts with Single School	0.31	(0.46)
Distance (km) to Nearest School in District	1.6	(9.7)
Dissimilarity Index in English Proficiency	0.10	(0.09)
Dissimilarity Index in Math Proficiency	0.13	(0.11)
Panel C: Los Angeles Unified		
Total School Enrollment	996	(659)
Test Enrollment	764	(643)
Share African American	0.13	(0.19)
Share Asian	0.04	(0.08)
Share Hispanic	0.68	(0.28)
Share White	0.12	(0.18)
Share Other	0.03	(0.05)
Share Socio-Economically Disadvantaged	0.75	(0.25)
Year-to-year Enrollment Percent Change	-1.5	(9.4)
Year-to-year Enrollment Absolute Change	-23.8	(83.6)
Intra-District Transfer Rate	0.07	(0.26)
Distance (km) to Nearest School in District	0.9	(0.5)
Dissimilarity Index in English Proficiency	0.31	(0.02)
Dissimilarity Index in Math Proficiency	0.33	(0.01)

Table 2: Summary Statistics by Sample Near the Cut-Off

	Sample 1		Sample 2		Sample 3		Sample 4	
Total observations	6,344		3,968		2,050		2,251	
Observations within 2pp of cut-off	800		787		423		478	
Panel A: School Characteristics	Mean	(St. Dev.)	Mean	(St. Dev.)	Mean	(St. Dev.)	Mean	(St. Dev.)
Total School Enrollment	649	(295)	704	(312)	693	(292)	734	(316)
Test Enrollment	480	(252)	527	(279)	513	(252)	568	(308)
Share African American	0.08	(0.11)	0.09	(0.13)	0.08	(0.12)	0.07	(0.10)
Share Asian	0.07	(0.10)	0.05	(0.08)	0.05	(0.09)	0.06	(0.10)
Share Hispanic	0.49	(0.26)	0.60	(0.26)	0.59	(0.27)	0.64	(0.25)
Share White	0.31	(0.25)	0.22	(0.22)	0.23	(0.22)	0.18	(0.19)
Share Other	0.05	(0.07)	0.04	(0.06)	0.04	(0.07)	0.04	(0.06)
Share Socio-Economically Disadvantaged	0.60	(0.24)	0.69	(0.23)	0.69	(0.22)	0.72	(0.22)
Year-to-year Enrollment Percent Change	-0.5	(7.8)	-1.3	(8.1)	-1.2	(8.1)	-2.7	(7.9)
Year-to-year Enrollment Absolute Change	-5.1	(50.5)	-9.6	(59.5)	-8.2	(56.1)	-24.7	(70.8)
Panel B: Estimated Student Mobility								
Estimated Number of New Students Enter	77	(60)	94	(89)	95	(97)	92	(70)
Number Enter / Total Enrollment	0.1	(0.1)	0.1	(0.1)	0.1	(0.1)	0.1	(0.1)
Estimated Number of Students Exit	83	(65)	106	(90)	106	(97)	105	(79)
Number Exit / Total Enrollment	0.1	(0.1)	0.2	(0.1)	0.2	(0.1)	0.1	(0.1)
Panel C: District Characteristics								
Number of schools in District	15	(33)	16	(35)	19	(40)	18	(40)
Share of Districts with Single School	0.07	(0.26)	0.06	(0.23)	0.05	(0.21)	0.06	(0.24)
Distance (km) to Nearest School in District	1.5	(6.3)	1.4	(6.4)	1.5	(8.6)	1.1	(2.2)
Dissimilarity Index in English Proficiency	0.15	(0.10)	0.15	(0.10)	0.17	(0.10)	0.16	(0.10)
Dissimilarity Index in Math Proficiency	0.17	(0.09)	0.19	(0.09)	0.20	(0.08)	0.19	(0.09)

Table 3: Lowest Performing Subgroup-Test Distribution Overall and by Sample Near the Cut-off

Lowest Performing Subgroup-Test	All Schools (02/03 to 05/06)	Sample 1	Sample 2	Sample 3	Sample 4
	Share of schools				
African American - ELA	0.03	0.04	0.04	0.03	0.03
African American - Math	0.04	0.05	0.05	0.06	0.03
Disability - ELA	0.04	0.03	0.04	0.03	0.04
Disability - Math	0.07	0.02	0.03	0.03	0.04
English Learner - ELA	0.33	0.40	0.58	0.57	0.62
English Learner - Math	0.04	0.05	0.04	0.04	0.06
Hispanic - ELA	0.09	0.12	0.08	0.11	0.07
Hispanic - Math	0.05	0.09	0.05	0.05	0.04
Socio-econ. Disadv. - ELA	0.12	0.09	0.03	0.03	0.03
Socio-econ. Disadv. - Math	0.06	0.09	0.04	0.03	0.02
Other	0.13	0.01	0.01	0.01	0.01
Total Number of Schools	26,374	804	793	424	488

Table 4: Sample 1 - Year 1 Failure Effect on Enrollment Growth (no sanctions)

	(1)	(2)	(3)	(4)
Failed AYP ($F^{CDE}_{s,01/02}$)	-1.25**	-1.76*	-3.35**	-2.27**
	(.536)	(.999)	(1.64)	(1.14)
AYP Proficiency Gap ($G_{s,01/02}$)	.0176**	-.0183	.212	
	(.008)	(.058)	(.618)	
$F_{s,01/02} \times G_{s,01/02}$	-.136***	-.673**	-1.49	
	(.033)	(.312)	(.951)	
$G_{s,01/02}{}^2$		.002		
		(.002)		
$G_{s,01/02}{}^3$		-.00002		
		(.00002)		
$F_{s,01/02} \times G_{s,01/02}{}^2$		-.128***		
		(.044)		
$F_{s,02/03} \times G_{s,02/03}{}^3$		-.006***		
		(.001)		
Constant	.021	.26	.856	
	(.347)	(.553)	(1.08)	
Observations	6,344	6,344	800	6,344
R-squared	.0003	.001	.003	

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Columns (1) and (2): parametric RD with 1st and 3rd order polynomial, respectively.
Column (3): limits the bandwidth around the cut-off to 2pp on either side.
Column (4): local polynomial RD estimate.

Notes: The dependent variable is the percent change in total enrolment from AY 01/02 to 02/03. Results are from the second stage of a two-stage least squares regression where the excluded first stage instrument is a dummy for whether schools failed to meet NCLB test proficiency standards, $F_{s,01/02}$. The coefficient of interest is on $F^{CDE}_{s,01/02}$, which indicates true AYP failure. AYP proficiency gap is the distance between the mandated proficiency rate and the proficiency rate of a school's lowest performing subgroup. Sample 1 includes all elementary schools with sufficient enrollment to report test outcomes.

Table 5: Sample 2: School Choice Effect on Enrollment Growth (1st Year)

	(1)	(2)	(3)	(4)
Failed AYP ($F^{CDE}_{s,02/03}$)	-0.197	-1.31	-0.875	-0.806
	(.506)	(.939)	(1.47)	(.869)
AYP Proficiency Gap ($G_{s,02/03}$)	0.043***	-0.083	-0.586	
	(.013)	(.078)	(.615)	
$F_{s,02/03} \times G_{s,02/03}$	-0.048	-0.22	0.868	
	(.0505)	(.422)	(.964)	
$G_{s,02/03}{}^2$		.005		
		(.003)		
$G_{s,02/03}{}^3$		-.00005		
		(.00003)		
$F_{s,02/03} \times G_{s,02/03}{}^2$		-.050		
		(.062)		
$F_{s,02/03} \times G_{s,02/03}{}^3$		-.001		
		(.003)		
Constant	-1.07***	-.432	-.444	
	(.303)	(.445)	(.783)	
Observations	3,968	3,968	787	3,968
R-squared	.003	.001	.002	

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Columns (1) and (2): parametric RD with 1st and 3rd order polynomial, respectively.
Column (3): limits the bandwidth around the cut-off to 2pp on either side.
Column (4): local polynomial RD estimate.

Notes: The dependent variable is the percent change in total enrolment from AY 02/03 to 03/04. Results are from the second stage of a two-stage least squares regression where the excluded first stage instrument is a dummy for whether schools failed to meet NCLB test proficiency standards, $F_{s,02/03}$. The coefficient of interest is on $F^{CDE}_{s,02/03}$, which indicates true AYP failure. AYP proficiency gap is the distance between the mandated proficiency rate and the proficiency rate of a school's lowest performing subgroup. Sample 2 includes all elementary schools with sufficient enrollment to report test outcomes and which failed AYP in AY 01/02.

Table 6: Sample 3: School Choice Effect on Enrollment Growth (2nd Year)

	(1)	(2)	(3)	(4)
Failed AYP ($F^{CDE}_{s,02/03}$)	-0.779	-2.53**	-4.21**	-2.71**
	(.683)	(1.21)	(2)	(1.17)
AYP Proficiency Gap ($G_{s,02/03}$)	0.078***	0.068	-1.27	
	(.016)	(.089)	(1.01)	
$F_{s,02/03} \times G_{s,02/03}$	-0.211**	-1.2	0.156	
	(.104)	(.791)	(1.33)	
$G_{s,02/03}{}^{2}$		.001		
		(.004)		
$G_{s,02/03}{}^{3}$		-.00002		
		(.00004)		
$F_{s,02/03} \times G_{s,02/03}{}^{2}$		-.127		
		(.135)		
$F_{s,02/03} \times G_{s,02/03}{}^{3}$		-.004		
		(.006)		
Constant	-1.95***	-1.74***	.003	
	(.328)	(.542)	(1.4)	
Observations	2,007	2,007	414	2,007
R-squared	.013	.005	-.022	

Robust standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Columns (1) and (2): parametric RD with 1st and 3rd order polynomial, respectively.
Column (3): limits the bandwidth around the cut-off to 2pp on either side.
Column (4): local polynomial RD estimate.

Notes: The dependent variable is the percent change in total enrolment from AY 03/04 to 04/05. Results are from the second stage of a two-stage least squares regression where the excluded first stage instrument is a dummy for whether schools failed to meet NCLB test proficiency standards, $F_{s,02/03}$. The coefficient of interest is on $F^{CDE}_{s,02/03}$, which indicates true AYP failure. AYP proficiency gap is the distance between the mandated proficiency rate and the proficiency rate of a school's lowest performing subgroup. Sample 3 includes all elementary schools with sufficient enrollment to report test outcomes and which failed AYP in AY 01/02 but passed in 03/04.

Table 7: Sample 4: Tutoring & School Choice Effect Relative to Choice Alone

	(1)	(2)	(3)	(4)
Failed AYP ($F^{CDE}_{s,03/04}$)	0.589	2.16*	5.67***	2.94**
	(.623)	(1.19)	(1.95)	(1.34)
AYP Proficiency Gap ($G_{s,03/04}$)	0.127***	0.411**	0.59	
	(.029)	(.164)	(.723)	
$F_{s,03/04} \times G_{s,03/04}$	-0.048	0.050	2.67**	
	(.069)	(.566)	(1.36)	
$G_{s,03/04}{}^2$		-.011		
		(.007)		
$G_{s,03/04}{}^3$		.00008		
		(.00008)		
$F_{s,03/04} \times G_{s,03/04}{}^2$		.076		
		(.084)		
$F_{s,03/04} \times G_{s,03/04}{}^3$		.003		
		(.004)		
Constant	-2.79***	-3.79***	-3.9***	
	(.402)	(.567)	(.895)	
Observations	2,251	2,251	478	2,251
R-squared	.012	.011	-.011	

Robust standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Columns (1) and (2): parametric RD with 1st and 3rd order polynomial, respectively.
Column (3): limits the bandwidth around the cut-off to 2pp on either side.
Column (4): local polynomial RD estimate.

Notes: The dependent variable is the percent change in total enrolment from AY 03/04 to 04/05. Results are from the second stage of a two-stage least squares regression where the excluded first stage instrument is a dummy for whether schools failed to meet NCLB test proficiency standards, $F_{s,03/04}$. The coefficient of interest is on $F^{CDE}_{s,03/04}$, which indicates true AYP failure. AYP proficiency gap is the distance between the mandated proficiency rate and the proficiency rate of a school's lowest performing subgroup. Sample 3 includes all elementary schools with sufficient enrollment to report test outcomes and which failed AYP in both AY 01/02 and 02/03.

Table 8: Local Polynomial RD Estimates for Each Sample Effect on Student Transfer

	(1) Sample 1: F vs. P	(2) Sample 2: F-F vs. F-P	(3) Sample 3: F-F-P vs. F-P-P	(4) Sample 4: F-F-F vs. F-F-P
Effect on Out-Transfer	0.003	-0.026*	-0.019	-0.029
	(0.010)	(0.012)	(0.018)	(0.025)
N	539,985	451,966	242,802	304,832
School clusters	523	397	187	215

Notes: Columns (1) to (3) replicate the non-parametric regressions used to estimate column (4) in Tables 4 to 6, employing the relevant samples from the Los Angeles Unified School District. Column (4) replicates the regression in Table 7 using the parametric method in column (2) of that table. The dependent variable is an indicator for whether a student transfers to a different school in the district the next year. Standard errors are clustered at the school-level.

Table 9: Effects for Socio-Economically Disadvantaged Students

Outcome:	(1) Sample 1: F vs. P	(2) Sample 2: F-F vs. F-P	(3) Sample 3: F-F-P vs. F-P-P	(4) Sample 4: F-F-F vs. F-F-P
Enrollment % Change (All students)	-2.273**	-0.806	-2.712**	2.940**
	(1.140)	(0.869)	(1.171)	(1.340)
Enrollment % Change - Socio-econ. Disadv.	-0.848	1.498	-5.328***	1.943
	(1.779)	(1.450)	(1.948)	(1.987)
Share of School is Socio-econ. Disadv.	0.017	0.023	0.040	0.002
	(0.030)	(0.026)	(0.032)	(0.041)

Notes: Each column replicates column (4) from Tables 4 to 7 for the relevant sample. The first row displays the results from the previous tables, the second row replaces the dependent variable in these regressions with the percent change in enrollment of tested students identified as socio-economically disadvantaged, and the third row with the share of total tested enrollment that was socio-economically disadvantaged.

Table 10: Subgroup-level Effects by Terciles of Proficiency Relative to Lowest Performing Group

Moderating variable: Ratio of own-subgroup proficiency to lowest performing subgroup proficiency

		(1) Sample 1: F vs. P	(2) Sample 2: F-F vs. F-P	(3) Sample 3: F-F-P vs. F-P-P	(4) Sample 4: F-F-F vs. F-F-P
Baseline Results	Coef.	-2.273**	-0.806	-2.712**	2.940**
	(St. err.)	(1.140)	(0.869)	(1.171)	(1.340)
	N	6,344	3,698	2,007	2,251
Baseline Results Estimated at Subgroup Level		-2.066	-0.980	-7.683**	9.118***
		(1.290)	(0.802)	(3.701)	(2.041)
		19,324	12,459	9,567	7,037
Bottom tercile		-1.623	0.742	-31.75**	5.084**
		(1.656)	(1.665)	(13.67)	(2.406)
		5,109	3,446	3,388	1,948
Middle tercile		-3.794*	0.117	-4.529	5.912***
		(2.109)	(2.362)	(7.397)	(2.171)
		6,129	4,174	3,587	2,394
Top tercile		-4.562*	-2.541	-5.067	16.02***
		(2.755)	(1.941)	(5.457)	(5.799)
		6,428	4,433	2,115	2,544

Notes: Each column replicates column (4) from Tables 4 to 7 for the relevant sample. The first row displays the results from the previous tables. The second row re-structures the data as one observation per subgroup-school and re-estimates the results, controlling for subgroup. Lastly, the subgroup-school sample is subset into terciles based on the ratio of a subgroup's proficiency rate to the lowest performing subgroup's proficiency rate in their school. Values for each tercile are given in Appendix Table A3. The results are re-estimated separately for each tercile in the final three rows.

Table 11: Heterogeneous Treatment Effects by Terciles of Distance to Nearest School

Moderating variable: Distance (in km) to nearest school in district

		(1) Sample 1: F vs. P	(2) Sample 2: F-F vs. F-P	(3) Sample 3: F-F-P vs. F-P-P	(4) Sample 4: F-F-F vs. F-F-P
Baseline Results	Coef.	-2.273**	-0.806	-2.712**	2.940**
	(St. err.)	(1.140)	(0.869)	(1.171)	(1.340)
	N	6,344	3,698	2,007	2,251
Bottom tercile		0.817	-0.933	-4.584*	3.917
		(2.199)	(1.647)	(2.704)	(2.855)
		1,980	1,242	539	713
Middle tercile		-1.405	-1.893	2.424	1.628
		(1.307)	(1.553)	(1.635)	(1.998)
		2,114	1,324	670	743
Top tercile		-4.869**	-1.114	-5.071**	5.441**
		(2.335)	(1.882)	(2.207)	(2.535)
		2,037	1,301	738	737

Notes: Each column replicates column (4) from Tables 4 to 7 for the relevant sample. The first row displays the results from the previous tables. The remaining three rows subset the data into terciles according to the distance to the nearest school within the same district and re-estimate the results separately for each susbet. Values for each tercile are given in Appendix Table A3.

Table 12: Heterogeneous Treatment Effects by Terciles of Distance to Nearest Passing School

Moderating variable: Distance (in km) to nearest **passing** school in district					
		(1) Sample 1: F vs. P	(2) Sample 2: F-F vs. F-P	(3) Sample 3: F-F-P vs. F-P-P	(4) Sample 4: F-F-F vs. F-F-P
Baseline Results	Coef.	-2.273**	-0.806	-2.712**	2.940**
	(St. err.)	(1.140)	(0.869)	(1.171)	(1.340)
	N	6,344	3,698	2,007	2,251
Bottom tercile		-3.688	-0.623	-4.721*	2.826
		(2.436)	(1.664)	(2.608)	(1.918)
		1,309	997	511	570
Middle tercile		-4.551**	-1.642	1.549	2.676
		(1.938)	(1.730)	(3.085)	(2.794)
		1,340	1,032	610	587
Top tercile		0.113	-0.494	-2.301	3.000
		(1.486)	(1.397)	(2.235)	(2.890)
		1,284	1,014	516	577

Notes: Each column replicates column (4) from Tables 4 to 7 for the relevant sample. The first row displays the results from the previous tables. The remaining three rows subset the data into terciles according to the distance to the nearest *passing* school within the same district and re-estimate the results separately for each susbet. Values for each tercile are given in Appendix Table A3.

Table 13: Results by Terciles of Performance Relative to Nearest Passing School

Moderating variable: Ratio of nearest passing school's academic performance index (API) relative to own-school API

		(1) Sample 1: F vs. P	(2) Sample 2: F-F vs. F-P	(3) Sample 3: F-F-P vs. F-P-P	(4) Sample 4: F-F-F vs. F-F-P
Baseline Results	Coef.	-2.273**	-0.806	-2.712**	2.940**
	(St. err.)	(1.140)	(0.869)	(1.171)	(1.340)
	N	6,344	3,698	2,007	2,251
Bottom tercile		-1.619	-1.654	-0.0864	4.607**
		(2.568)	(2.278)	(3.133)	(2.135)
		1,314	870	675	502
Middle tercile		-3.771**	0.254	-4.674**	2.498
		(1.626)	(1.619)	(1.912)	(2.954)
		1,026	1,049	603	569
Top tercile		-3.502	0.760	-0.211	11.09*
		(2.180)	(2.072)	(3.255)	(6.249)
		1,344	1,028	314	549

Notes: Each column replicates column (4) from Tables 4 to 7 for the relevant sample. The first row displays the results from the previous tables. The remaining three rows subset the data into terciles according to the ratio of the nearest passing school's Average Performance Index (API) to a school's own API and re-estimate the results separately for each susbet. API is a composite index of school test scores. Values for each tercile are given in Appendix Table A3.

Table 14: Results by Terciles of District Dissimilarity in ELA and Mathematics Proficiency

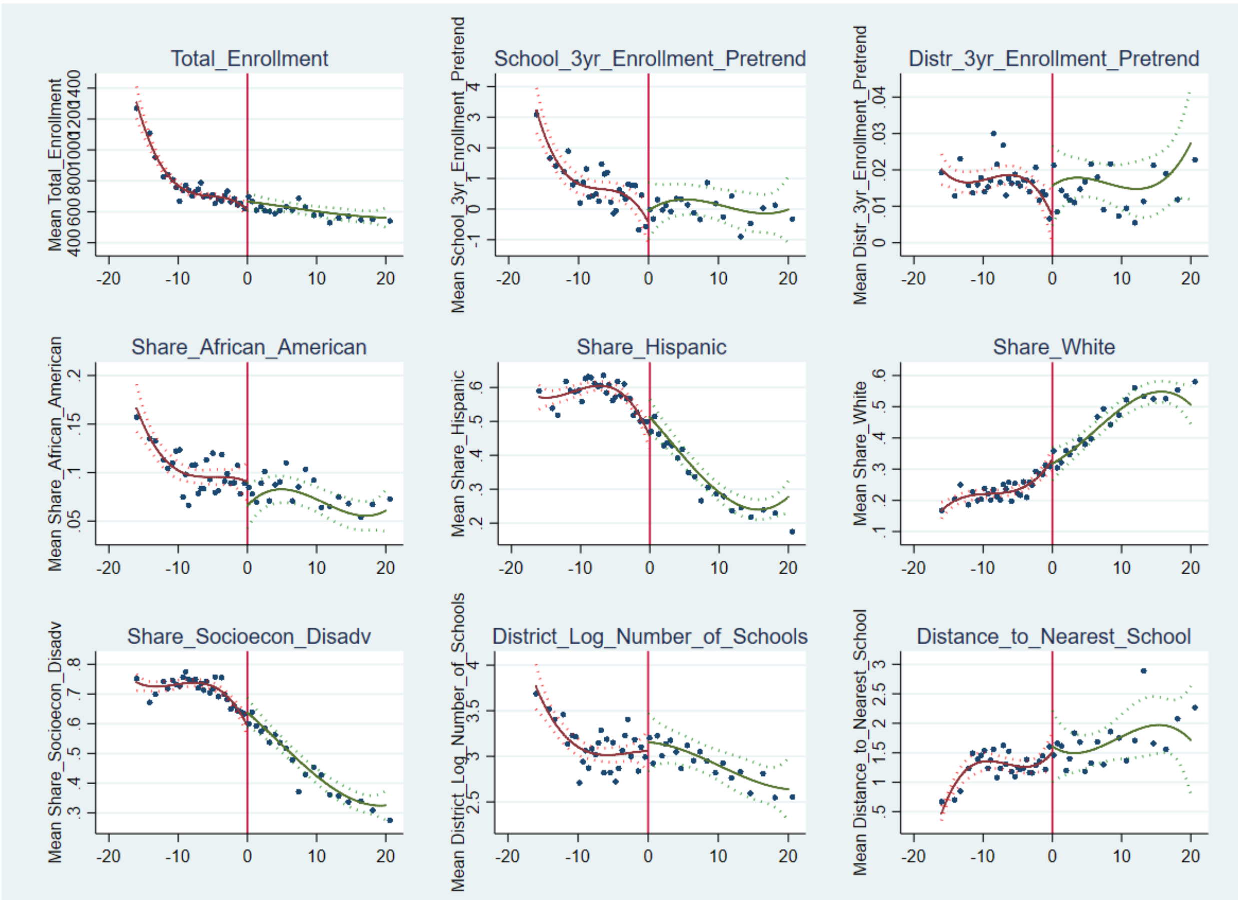
Moderating variable: District Dissimilarity Index in ELA and Mathematics Profociency

			(1) Sample 1: F vs. P	(2) Sample 2: F-F vs. F-P	(3) Sample 3: F-F-P vs. F-P-P	(4) Sample 4: F-F-F vs. F-F-P
Baseline Results		Coef.	-2.273**	-0.806	-2.712**	2.940**
		(St. err.)	(1.140)	(0.869)	(1.171)	(1.340)
		N	6,344	3,698	2,007	2,251
ELA	Bottom tercile		-1.288	-0.190	-2.581	3.801**
			(2.916)	(1.539)	(1.913)	(1.808)
			2,046	1,292	692	739
	Middle tercile		-2.697	-0.0196	-4.136*	1.839
			(2.131)	(1.386)	(2.230)	(2.121)
			2,059	1,295	657	723
	Top tercile		-1.463	-1.719	-0.787	2.918
			(1.449)	(1.400)	(1.837)	(3.015)
			2,027	1,283	600	732
Math	Bottom tercile		-2.950	0.366	-2.880	2.913
			(2.409)	(1.362)	(2.477)	(1.871)
			2,048	1,317	743	735
	Middle tercile		-0.570	-2.255	-5.317**	2.204
			(1.715)	(1.989)	(2.332)	(1.627)
			2,045	1,272	624	736
	Top tercile		-2.427	-1.510	-0.681	5.115
			(1.544)	(1.727)	(2.432)	(4.505)
			2,039	1,281	582	723

Notes: Each column replicates column (4) from Tables 4 to 7 for the relevant sample. The first row displays the results from the previous tables. The remaining three rows subset the data into terciles according to the distance to the a measure of the within-district academic achievement dissimilarity across schools. This measures the share of students who would have to be moved across schools in order for all schools in the district to have the same proficiency rate in ELA or mathematics. Values for each tercile are given in Appendix Table A3.

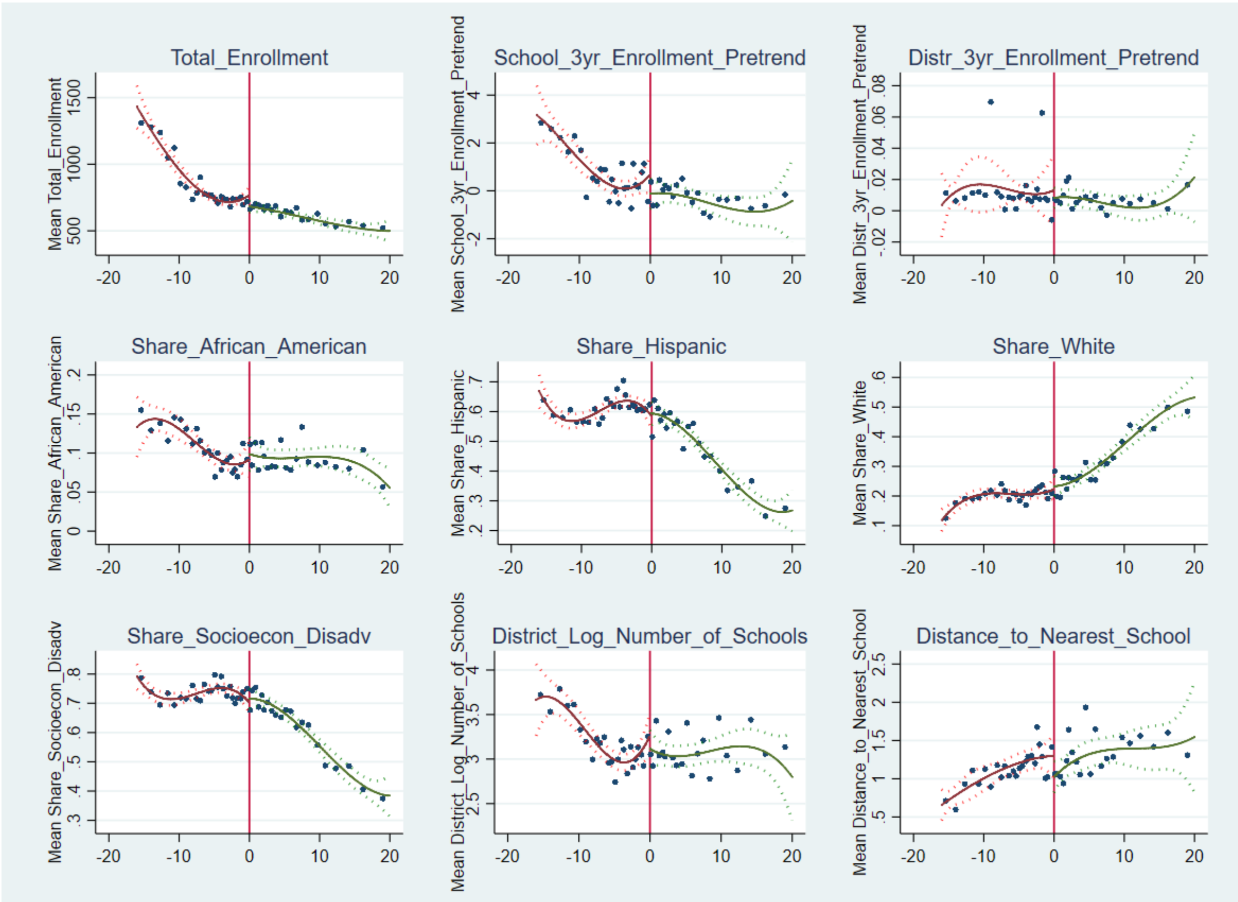
9 Appendix Figures

Figure 91: Sample 1 - Regression Discontinuity on Covariates



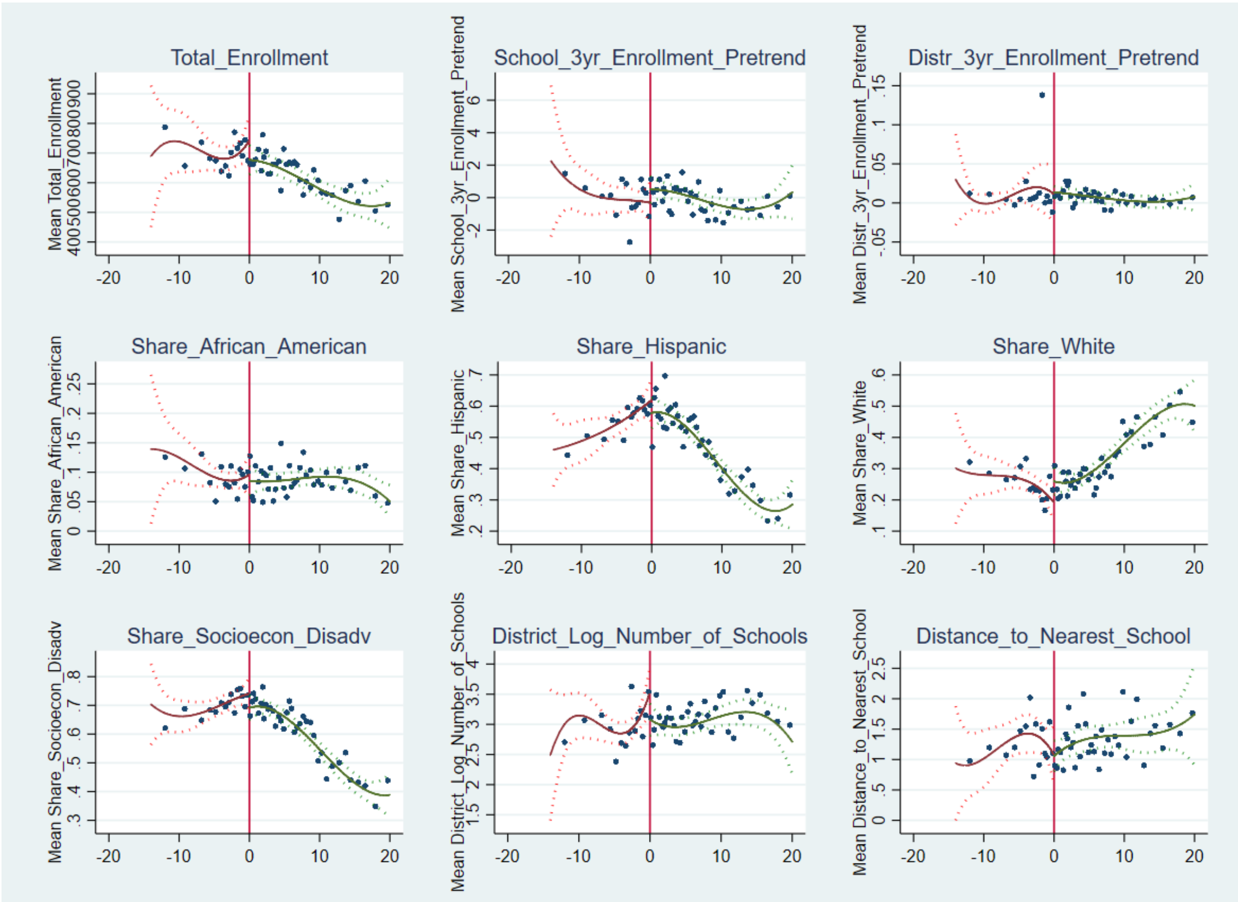
Notes: Each graph replicates Figure 3 but replacing the dependent variable with a covariate to ensure there are no discontinuities at the cut-off.

Figure 92: Sample 2 - Regression Discontinuity on Covariates



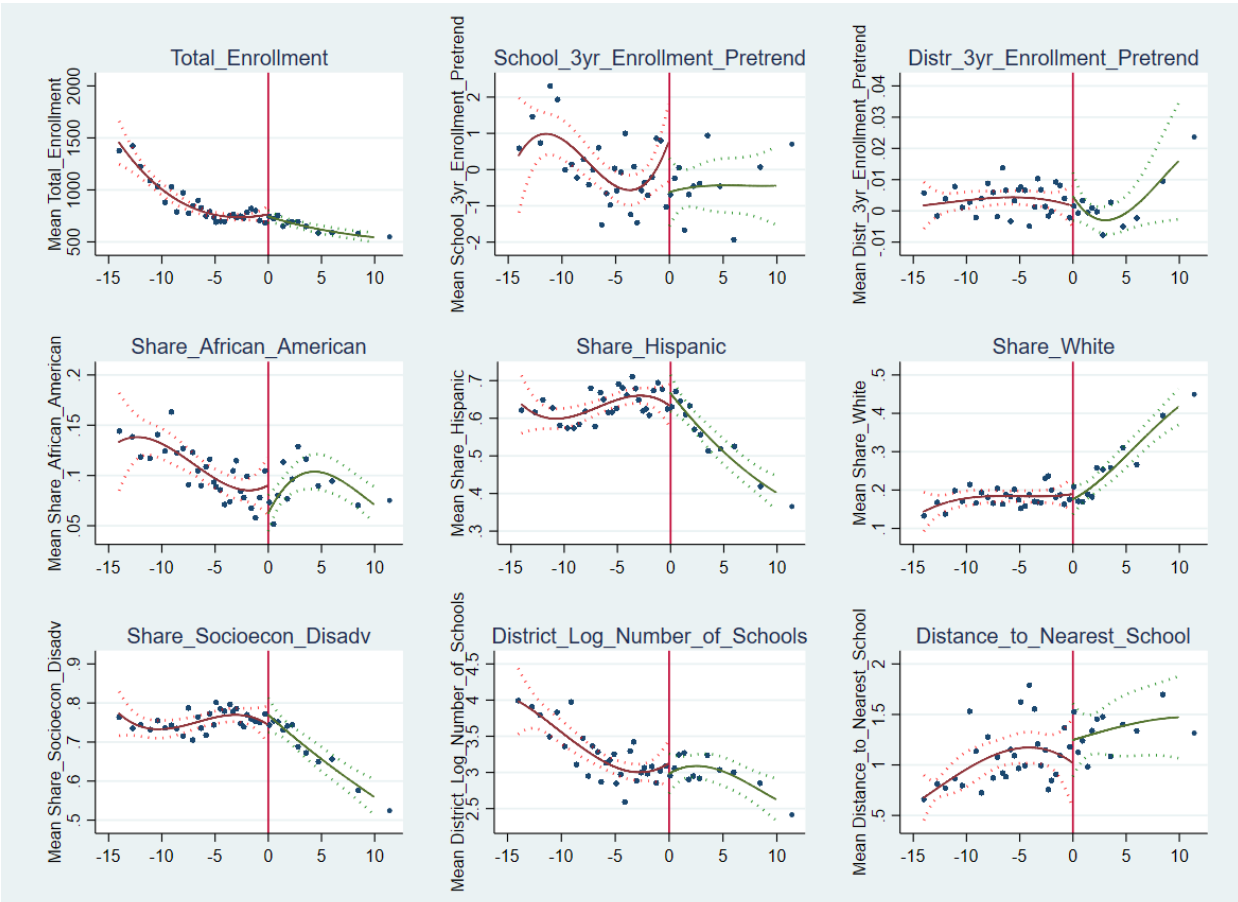
Notes: Each graph replicates Figure 4 but replacing the dependent variable with a covariate to ensure there are no discontinuities at the cut-off.

Figure 93: Sample 3 - Regression Discontinuity on Covariates



Notes: Each graph replicates Figure 5 but replacing the dependent variable with a covariate to ensure there are no discontinuities at the cut-off.

Figure 94: Sample 4 - Regression Discontinuity on Covariates



Notes: Each graph replicates Figure 6 but replacing the dependent variable with a covariate to ensure there are no discontinuities at the cut-off.

10 Appendix Tables

Table 101: Results for All Treatment Histories

	Unconditional on Prior Treatment Histories		
	(1)	(2)	(3)
	Year 1	Year 2	Year 3
RD Estimate	-2.273**	-1.547	0.853
	(1.140)	(0.993)	(0.957)
Observations	6,344	6,405	6,485

Conditional on Prior Treatment Histories					
(4)	(5)	(6)	(7)	(8)	(9)
Year 2 - Passed Y1	Year 2 - Failed Y1	Year 3 - PP	Year 3 - PF	Year 3 - FP	Year 3 - FF
-8.340*	-0.806	0.366	1.012	-0.777	2.940**
(4.473)	(0.869)	(3.185)	(2.354)	(1.511)	(1.340)
2,437	3,968	2,227	177	1,633	2,251

Table 102: Results limiting bandwidth and controlling for prior year running variable

	Unconditional on Prior Treatment Histories		
	(1)	(2)	(3)
	Year 1	Year 2	Year 3
RD Estimate	-2.273**	-3.814	0.911
	(1.140)	(2.458)	(1.177)
Observations	6,344	1,199	2,013

Conditional on Prior Treatment Histories					
(4)	(5)	(6)	(7)	(8)	(9)
Year 2 - Passed Y1	Year 2 - Failed Y1	Year 3 - PP	Year 3 - PF	Year 3 - FP	Year 3 - FF
-2.368	-3.533	-1.436	4.667**	-1.453	3.400*
(4.214)	(3.059)	(3.880)	(2.298)	(1.475)	(1.901)
392	807	193	64	718	1,014

Table 103: Values of Moderating Variables across Terciles for Each Year

		Bottom tercile		Middle tercile		Top tercile	
		min	max	min	max	min	max
Distance to nearest other school in district (km)	Year 1 Sample	0	0.67	0.67	1.14	1.14	93.61
	Year 2 Sample	0	0.61	0.61	1.05	1.05	71.66
	Year 3 Sample	0	0.55	0.56	0.99	1	72.37
Distance to nearest passing school in district (km)	Year 1 Sample	0	1.05	1.05	2.25	2.25	93.61
	Year 2 Sample	0	0.87	0.87	1.54	1.55	71.66
	Year 3 Sample	0	0.80	0.80	1.45	1.45	72.37
Ratio of own-school performance index (API) to that of nearest passing school	Year 1 Sample	0.72	1	1.00	1.08	1.08	3.09
	Year 2 Sample	0.72	1	1	1.09	1.09	2.12
	Year 3 Sample	0.61	1.04	1.04	1.12	1.12	2.75
Dissimilarity index in math proficiency across district schools	Year 1 Sample	0	0.17	0.17	0.27	0.27	0.76
	Year 2 Sample	0	0.20	0.20	0.27	0.27	0.48
	Year 3 Sample	0	0.21	0.21	0.28	0.28	0.66
Dissimilarity index in English proficiency across district schools	Year 1 Sample	0	0.15	0.15	0.26	0.26	0.68
	Year 2 Sample	0	0.16	0.16	0.28	0.28	0.47
	Year 3 Sample	0	0.16	0.16	0.27	0.28	0.73

Notes: Terciles are defined based on a sample of schools which failed all the previous years of NCLB. Thus, the Year 1 sample includes all schools, the Year 2 sample includes schools which failed in Year 1, and the Year 3 sample includes schools which failed in both Years 1 and 2. In regards to the samples defined for analysis, "Year 2 Sample" terciles are used for both Samples 2 and 3 while "Year 3 Sample" is used for Sample 4.