Choice and Consequence:
Assessing Mismatch at Chicago Exam Schools

Joshua D. Angrist Parag A. Pathak Román A. Zárate*

July 2019

Abstract

The educational mismatch hypothesis asserts that students are hurt by affirmative action policies that place them in selective schools for which they wouldn't otherwise qualify. We evaluate mismatch in Chicago's selective public exam schools, which admit students using neighborhood-based diversity criteria as well as test scores. Regression discontinuity estimates for applicants favored by affirmative action indeed show no gains in reading and negative effects of exam school attendance on math scores. These results hold for more selective schools and for applicants most likely to benefit from affirmative-action, a pattern suggestive of mismatch. However, exam school effects in Chicago are explained by schools attended by applicants who are *not* offered an exam school seat. Specifically, mismatch arises because exam school admission diverts many applicants from high-performing Noble Network charter schools, where they would have done well. Consistent with these findings, exam schools reduce Math scores for applicants applying from charter schools in another large urban district. Exam school applicants' previous achievement, race, and other characteristics that are sometimes said to mediate student-school matching play no role in this story.

*We are grateful to Miikka Rokkanen for his contributions to this project. Our thanks to the Chicago Public Schools, the Noble Network, and an anonymous large urban school district for graciously sharing data, to Clemence Idoux and Ignacio Rodriguez for excellent research assistance, and to MIT SEII program managers Annice Correia, Eryn Heying, and Anna Vallee for invaluable administrative support. We thank Will Dobbie, Glenn Ellison, Amy Finkelstein, Michael Greenstone, Peter Hull, Chris Walters and seminar participants at Chicago, MIT, and Princeton for helpful input. Financial support from the Laura and John Arnold Foundation is gratefully acknowledged. Pathak also thanks the National Science Foundation and the W.T. Grant Foundation for research support. Angrist's daughter teaches at a charter school.

1 Introduction

The educational mismatch hypothesis predicts poor outcomes for the beneficiaries of diversity preferences if those benefitting are ill-prepared for the sort of education provided by selective schools. Proposed by Sander (2004) as a cause of racial gaps in bar exam passage rates, the mismatch hypothesis has inspired a large and expanding literature on the consequences of affirmative action in college admissions (see, e.g., Sander and Taylor (2012) and Arcidiacono and Lovenheim (2016)). The debate here is more than academic: mismatch arguments were prominent in the 2016 *Fisher v. University of Texas* Supreme Court case, in which the court upheld the constitutionality of affirmative action policy at the University of Texas.[1]

Public discussion of mismatch has focused on higher education, but the same issues arise for admission to selective public high schools, often called exam schools. Like selective colleges, exam schools offer high-achieving students the opportunity to attend public schools with similarly bright classmates, experienced teachers, additional courses and clubs, and perhaps other resources beyond those available to students at non-selective public schools. Because public exam schools select their students using entrance exams and course grades, criteria that disadvantage minority applicants, affirmative action (AA) is seen by many as necessary for exam school diversity.

In the 1970s and 1980s, exam schools in Boston, Chicago, and San Francisco guaranteed exam school seats for minorities after these districts came under court control as a consequence of desegregation efforts. US courts have since treated the issues raised by affirmative action in exam school admissions and higher education similarly. The landmark 1978 Bakke decision outlawed racial quotas in college admissions. Since the 1990s, exam school admissions have likewise moved to de-emphasize race.[2] But the question of whether minority exam school applicants are indeed well-served by affirmative action policies remains.

This paper assesses the case for mismatch arising from affirmative action at exam schools in the mostly nonwhite Chicago Public School (CPS) district. Affirmative action in Chicago lowers admissions cutoffs for applicants from lower-income neighborhoods and raises them elsewhere. We use exam school cutoffs to construct regression discontinuity (RD) estimates of exam school effects that compare applicants who are just above and below neighborhood-specific cutoffs. These results

[1] During the court proceedings, Justice Antonin Scalia stated: "There are those who contend that it does not benefit African-Americans to get them into the University of Texas, where they do not do well, as opposed to having them go to a less-advanced school, a slower-track school where they do well."

[2] The U.S. Supreme Court ruled against Seattle and Louisville's race-based public school admissions systems in 2007.

show statistically significant negative effects of exam school attendance on math scores.[3]

Applicants from the most disadvantaged neighborhoods face markedly lower cutoffs. Like the population affected by affirmative action, these low-tier applicants enter exam schools with scores well below those of most admitted students. But the extent of this sort of "preparedness mismatch" varies across school selectivity and with individual student background. Chicago's system for centralized exam school assignment allows us to construct instrumental variables estimates around a wide range of cutoffs, for applicants from different tiers. These results show similar effects of enrollment at more and less selective schools and for applicants more and less likely to benefit from affirmative action, a pattern of findings inconsistent with the mismatch hypothesis.

The fact that high-tier applicants who clear their cutoff have baseline achievement well above the median at the school where they're offered a seat. This observation leads to a direct test of Sander (2004)'s formulation of mismatch, which argues that the gap between applicant credentials and those of the median student in his or her school determine match quality.[4] Estimates from models allowing effects to differ according to whether applicants have scores above or below those of the median student provide evidence on this claim. Although not definitive, this direct test of mismatch suggests forces other than mismatch lie behind negative exam school effects.

Mixed evidence for "preparedness mismatch" motivates our exploration of alternative explanations of CPS exam school effects. Many of Chicago's exam school applicants who are not admitted end up in another non-traditional sector. The most important alternatives are charter and magnet schools, both of which use lotteries to select their students. This allows us to combine the RD design generated by exam school admissions with lottery variation in a three-sector model. This model isolates non-traditional sector effects relative to a traditional (mostly neighborhood school) counterfactual. Multi-sector estimates reveal that negative exam school effects are driven by diversion away from the Noble Network of charter schools. Noble is a high-performing, "No-Excuses"-style charter management organization (CMO). Because Noble enrollment boosts achievement sharply, the diversion away from Noble induced by an exam school offer reduces test scores.[5]

[3] Barrow et al. (2016) report RD estimates of the effects of Chicago exam school offers. We comment on these results below. Other related studies of American exam schools include Abdulkadiroğlu et al. (2014) and Dobbie and Fryer (2014). These studies mostly find no effect of exam school attendance on achievement or college attendance.

[4] Sander (2004) writes: "If there is a very large disparity at a school between the entering credentials of the median student and the credentials of students receiving large preferences, then the credentials gap will hurt those the preferences are intended to help. A large number of those receiving large preferences will struggle academically, receive low grades, and actually learn less in some important sense than they would have at another school where their credentials were closer to the school median."

[5] Other examples of this sort of sectoral or program-operator substitution appear in studies by Heckman et al. (2000), Kline and Walters (2016), Angrist et al. (2013), and Chabrier et al. (2016). Kirkeboen et al. (2016) explore

Finally, we report briefly on exam school effects for a sample of applicants enrolled in charter middle schools in a second large urban district. This district has three exam schools and a high-performing charter sector that mostly embraces No-Excuses-style pedagogy. Exam school applicants enrolled in a charter school at the time of application typically remain in the charter sector if they fail to win an exam school seat. Consistent with our Chicago findings, exam school offers reduce achievement for this district's charter-originating exam school applicants.

The rest of the paper is organized as follows. Section 2 discusses institutional background related to Chicago's exam school sector. Section 3 describes our data and presents descriptive statistics. Section 4 reports exam school enrollment effects identified by RD tie-breaking in the CPS centralized match. The methodology here is an application of Abdulkadiroğlu et al. (2017b). Section 5 isolates the leading non-traditional exam school alternatives—magnet and charter schools—and shows how charter school effects determine exam school effects. Section 6 sketches a formal model of mismatch and shows that a simple additive model of school effects fits the CPS data remarkably well. Results for a second large urban district are discussed in Section 7. Section 8 concludes.

2 Institutional Background

2.1 Exam Schools and Affirmative Action in Chicago

CPS is the third largest American school district, with more than 600 schools and roughly 400,000 students. As in many urban districts, most CPS students are black or Hispanic and from low-income families. CPS high school students attend neighborhood schools by default, but can choose other schools and programs. Choice options include exam schools (known locally as selective enrollment schools), magnet schools, charter schools, military academies, and career academies.

In the years covered by this study, CPS operated nine exam schools.[6] Exam schools offer their students a curriculum emphasizing honors and Advanced Placement (AP) courses. Not surprisingly, exam school students are higher-achieving than most of their public school peers. Chicago's most selective exam schools, Northside and Payton, are frequently listed as among the best U.S. public high schools. In the 2016 U.S. News & World Report ranking, for example, Northside was ranked 39, while Payton was 41st.

substitution patterns and their consequences across college majors.

[6]These are Brooks College Prep High School, Jones College Prep High School, King College Prep High School, Lane Tech High School, Lindblom Math and Science Academy, Northside College Prep High School, Payton College Prep High School, Westinghouse College Prep High School, and Whitney M. Young Magnet High School. South Shore International High School opened in 2013 and Hancock College Prep High School opened in 2015.

Selective enrollment applicants apply by ranking the exam schools to which they wish to be admitted. Until the 2010–2011 admissions cycle, applicants could rank up to four schools. Applicants have since been able to rank up to six schools. The selective school admissions formula assigns equal weight to entrance test results, a standardized test taken in middle school, and to letter grades earned in 7th grade. These criteria generate a *composite score* running from 0 to 900 that schools use to rank applicants. The most selective schools admit students with composite scores above about 800; the least selective admit students with scores as low as 650, typically around the 66th quantile in the applicant composite score distribution.

Chicago's school choice system grew out of a 1980 desegregation consent decree in which the district agreed to promote school integration and to increase access for black and Hispanic students by offering diverse school choice options. Initially, the choice system considered race. But in 2009, in response to a federal court decision vacating the original consent decree, CPS adopted an affirmative action plan for selective schools based on the characteristics of the census tracts where applicants live. This system remains in place.

Chicago's system of residential preferences has been seen as a model for race-neutral admissions at selective schools (see, e.g., Kahlenberg (2014)). Federal guidelines on acceptable alternatives to promote diversity at selective K-12 schools also resemble Chicago's plan (see OCR (2011)).[7] The Chicago system assigns each census tract a score equal to the sum of its percentile rank on five dimensions: median family income; a measure of adult educational attainment; home ownership rates; and the prevalence of single-parent households and non-native English speakers.[8] These scores are then used to assign tracts to one of four tiers, so that roughly a quarter of CPS students live in each tier. Barrow et al. (2016) reproduce a map of Chicago's census tracts by tier. Higher-income tracts, in tier 4, are on the city boundary, while the most disadvantaged, in tier 1, are concentrated in the southern and western parts of the city center.

Since 2009, the CPS exam school admissions process has used a version of the deferred acceptance (DA) algorithm, now used for school assignment at CPS's other schools and around the country. The exam school implementation of DA reflects tier-based affirmative action. Specifically, the CPS exam-school assignment mechanism allocates a fixed share of seats, called merit seats, using composite scores only, while remaining seats are split equally across tiers. In 2011, for example,

[7] Ellison and Pathak (2016) quantify the allocative efficiency of race-neutral affirmative action plans and compare Chicago's new plan with alternatives.

[8] Since 2010, the index has included measures of local-area school performance.

70% of seats were assigned within tiers and the rest were merit seats.[9]

DA for school assignment considers school priorities over students (such as tier) as well as applicant preferences. As noted by Abdulkadiroğlu et al. (2017b), however, the CPS system can be represented as DA without priorities by dividing each school into 5 sub-schools, one containing merit seats and the other four containing equal numbers of tier-reserved seats. Applicants who rank a school are first considered for merit seats and then for the seats reserved for their tier. DA without priorities is called serial dictatorship (SD). An important feature of SD (and DA) is that admission decisions are determined solely by a set of school-specific cutoffs (Abdulkadiroğlu et al., 2017a). CPS's version of DA therefore produces distinct cutoffs for each school and tier.[10]

Figure 1 reports the Fall 2011 cutoffs for each school and tier. At Northside and Payton, the most selective schools, students from tier 4 neighborhoods needed almost 900 points (the maximum score) to get in. By contrast, cutoffs for students from tier 1 neighborhoods are about 100 points lower. These differences are represented by the dots in the figure. The figure also shows that cutoffs for these schools in a system without affirmative action would be near the tier 3 cutoff of 860. Tier 1 and 2 applicants with scores between the cutoff for their tier and 860 are admitted because of affirmative action.

The median composite score for admitted students, also plotted in Figure 1, provides a point of comparison for tier-specific cutoffs. At Northside, for example, the median for those admitted exceeds the cutoffs for tiers 1-3. Admitted applicants from these lower-SES tiers are therefore especially likely to have scored below the school median, a gap that becomes a possible source of mismatch. Payton cutoffs show a similar pattern.

Lower down the hierarchy of exam school selectivity, at Lane, Young, and Jones, gaps between the tier 4 and tier 1 cutoffs are also about 100 points. Cutoffs at Brooks show an interesting reversal: tier 3 applicants must clear a higher bar than applicants from (higher-SES) tier 4. This reflects differences in demand for these schools across tiers. Similarly, at the three least selective exam schools (King, Lindblom, and Westinghouse), affirmative action considerations generate a cutoff

[9] Dur et al. (2016) detail the CPS implementation of DA further.

[10] Chicago implements tier reservations by running DA after splitting each school into merit and tier seats. Applicants are first considered for merit seats at schools they rank; if rejected at school s, they apply to school s seats for their tier. Dur, Pathak, and Sonmez (2016) show that this implies that tier cutoffs are lower than cutoffs for merit seats, so cutoffs for merit seats can be ignored in our description of assignments. Provided every school is over-subscribed, this DA implementation is the same as SD for a market without priorities, where applicants from tier t rank tier-t seats at s immediately below merit seats at s, as if these seats were at a different school. This can be seen by noting that when schools are over-subscribed, an applicant from tier t is never assigned to seats reserved for other tiers. As a result, applicants can be processed in order of their composite score as in SD and need only be considered at merit seats and the tier seats allocated to their tier. Chicago isn't quite SD because not all exam schools are always over-subscribed for each tier. But SD replicates 99.7 % of assignments in our sample.

for tier 2 and 3 applicants higher than for tier 4. Typically, however, applicants from lower tiers must clear a lower cutoff for exam school admission. If students from disadvantaged neighborhoods are less prepared for the rigors of an exam school education, this reduction in admissions standards might produce mismatch.

2.2 Non-Traditional Exam School Alternatives

The two most important non-traditional alternatives for Chicago's exam school applicants (as measured by enrollment) are magnet schools and charter schools. Magnet schools enroll applicants district-wide and offer arts programs, agricultural science courses, and International Baccalaureate programs. In the period covered by our sample, CPS had four magnet high schools: Chicago High School for Agricultural Sciences, Clark Academic Prep, Curie Metropolitan, and Von Steuben Metropolitan. Cullen et al. (2006) use lotteries to estimate the achievement consequences of attendance at a Chicago magnet school in 2000, uncovering little evidence of a magnet impact.

Magnet schools admit students through school-specific lotteries, though low-tier and nearby applicants have a higher chance of winning a seat. Applicants must earn a minimum qualifying score on grade 7 achievement tests. Siblings of enrolled students are subject to a separate sibling lottery. Students who live within 2.5 miles of a school participate in a proximity lottery that allocates up to 40% of non-sibling seats. Remaining capacity is divided equally into four tier-specific lotteries. Seats are offered to applicants with the highest lottery numbers for their tier. Unlike exam schools, magnet offers are decentralized, so applicants may get more than one.

Charter schools are autonomous publicly-funded schools that operate in a framework known as a charter. The first Illinois charter school opened in 1996. Illinois charters are typically approved by local school districts, though some are granted by the state. Charters usually run for 5 years and must be certified by the Illinois State Board of Education. Current law restricts Chicago to 75 charters, but allows expansions by charter schools established before 2003.

The Noble Network of Charter Schools is one of Chicago's most visible charter providers, enrolling 40% of Chicago's 9th grade charter students. Founded in 1999, the Network started with a single campus, Noble Street College Prep, but has expanded rapidly since 2006. In the years covered by our study, Noble had nine additional campuses.[11] With the exception of Gary Comer College Prep, which includes a middle school, Noble runs high schools spanning grades 9–12.

[11] These are Chicago Bulls College Prep, Gary Comer College Prep, Golder College Prep, Johnson College Prep, Muchin College Prep, Pritzker College Prep, Rauner College Prep, Rowe-Clark Math & Science Academy, and UIC College Prep. Since 2014, Noble Network has added six schools.

Noble Network pedagogy is similar to that of other "No Excuses" charter schools, emphasizing extended instruction time, discipline and comportment, and data-driven feedback and teacher training. Noble Network schools attract attention inside and outside Chicago. The network was awarded an expansion grant from the U.S. Department of Education in 2015, one of only 12 Charter Management Organizations to be so recognized. Using data and admissions lotteries for applicants to the flagship campus for 2003-5, Davis and Heller (2019) find large positive effects of Noble attendance on college attendance. In other work, we've seen promising achievement results for No Excuses charters in Boston (Abdulkadiroğlu et al., 2011; Angrist et al., 2013), Denver (Abdulkadiroğlu et al., 2017a), New Orleans (Abdulkadiroğlu et al., 2016), and New York City (Dobbie and Fryer, 2013). Dobbie and Fryer (2015) also report encouraging results for effects of No Excuses charter attendance on teen pregnancy and involvement in the criminal justice system.[12]

Any Chicago resident student who completes 8th grade can apply to as many Noble campuses as they like, and may receive multiple Noble offers. Seats at oversubscribed campuses are assigned by (school-specific) lotteries. Students who not receiving a lottery offer go onto a randomly-ordered waitlist. Some Noble campuses are typically under-subscribed. Applicants with siblings enrolled at a Noble school are automatically admitted there.

3 Data and Descriptive Statistics

Data provided by CPS data contain information on enrollment, test scores, exam school admission, and magnet school admission. We also obtained application and admissions records from the Noble Network. Our CPS files cover the 2009–2010 to 2014–2015 school years. The files include information on students' school and grade, as well as demographic information such as gender, race, free/reduced price lunch status, and special education status.

Our analysis looks at school attendance effects on PLAN and ACT scores for exam school applicants enrolled in CPS at the time of application. PLAN is a preliminary ACT test, typically taken in 10th grade. ACT is a college readiness assessment taken as part of the CPS Prairie State Achievement exams (PSAE), usually administered in 11th grade, and required of all high school students until recently. Eighth grade Illinois State Achievement Test (ISAT) scores in math and reading provide a baseline achievement control. Scores are standardized to be mean zero and have

[12]Our tabulations of school climate data from Illinois State Board of Education report cards (available ar https://www.illinoisreportcard.com) show that other Chicago charters have fewer students taking advanced courses, higher student absenteeism and chronic truancy rates, and are less likely to share features of the No Excuses model.

unit standard deviation in the CPS population each year. We use the first score available for students who repeat an exam. The restriction to applicants enrolled in CPS ensures that exam school students are compared to students attending other within-district choice options and reduces loss to follow-up. The Data Appendix details sample construction further.

The exam and magnet school admissions files cover three application cohorts, from 2009–2010 to 2011–2012, the beginning of the tier-based affirmative action period. These files record applicants' ranking of schools, their residential tier, and school offered. Each applicant's composite score appears in the applicant file; it's these that determine the cutoffs discussed above. For each magnet applicant, we also observe magnet school lottery numbers, and whether an applicant had sibling priority, or participated in the proximity or tier lottery. Noble Network admissions files cover Noble Network applicants from 2009–2010 to 2011-2012. These files record campuses applied to, Noble lottery numbers (one for each school), and lottery outcomes (offer or waitlist). Noble files also identify applicants admitted via sibling priority.

Most CPS students are nonwhite, and most are poor enough to qualify for a subsidized lunch. These and other comparisons of student characteristics by school sector appear in Table 1. Exam school students are less likely to be non-white or poor, and have higher baseline scores, than other CPS students. Magnet students are more likely to be Hispanic and less likely to be black than the CPS population; they also have baseline scores above those of other CPS students, but their baseline scores are well below those of incoming exam school students. Charter students are more likely to be black and less likely to be Hispanic than other CPS students. They also have lower baseline scores than the typical CPS 9th grader. Noble enrolls a lower proportion of black students than other charter schools, but a higher proportion of Hispanics. Noble students also have baseline scores above those of other charter students, but below the CPS average.

4 The Case for Mismatch

4.1 RD Estimates

We're interested in the achievement consequences of exam school attendance for affirmative action beneficiaries and other exam school applicants. As detailed in Abdulkadiroğlu et al. (2017b), the CPS exam school assignment mechanism generates a regression-discontinuity (RD) research design that identifies exam school effects. In combination with instrumental variables derived from magnet and charter school lotteries, we can use this variation to estimate a variety of school sector effects.

As in other RD designs, exam-school RD compares applicants across a set of cutoffs. The CPS assignment mechanism can be modeled as a serial dictatorship that generates cutoffs $\tau_s(t)$, for each school s and tier t. Cutoffs are tier-specific because applicants compete for seats reserved for their tier, only after merit seats are assigned. Following Abdulkadiroğlu et al. (2017a), let $\theta_i = (\succ_i, t_i)$ denote applicant i's type, where $\succ_i$ is the applicants ranking of schools and t_i is her tier. The composite score used for admissions is denoted by r_i; this is the RD running variable.

Let S_θ denote the set of schools ranked by applicants of type θ. In serial dictatorship, the probability an applicant is offered a seat at *any* of these schools is the probability the applicant has a composite score that clears the most forgiving (that is, the lowest) of the these schools' cutoffs. Formally, the *qualifying cutoff* faced by applicant i, denoted $q(\theta_i)$, and the school that identifies it, $s^*(\theta_i)$, are defined as:

$$q(\theta_i) = \min_{s \in S_{\theta_i}} \{\tau_s(t_i)\},$$

$$s_i^* = s^*(\theta_i) = \arg\min_{s \in S_{\theta_i}} \{\tau_s(t_i)\},$$

where t_i fixes applicant tier.[13] As in other nonparametric applications of RD research designs, we look at applicants near the relevant qualifying cutoff. A dummy variable indicating any exam school offer to applicants of type θ_i in this window can be coded as

$$D_i = 1[r_i \geq q(\theta_i)].$$

Conditional on suitable running variable controls, this any-offer dummy is an instrument for any-exam-school enrollment, C_i, measured in years.

Benchmark just-identified fuzzy RD estimates come from the following first- and second-stage equations,

$$C_i = \gamma D_i + \alpha_1(s_i^*, t_i) + X_i'\delta_1 + h(r_i) + \nu_i \tag{1}$$

$$Y_i = \beta C_i + \alpha_2(s_i^*, t_i) + X_i'\delta_2 + g(r_i) + \varepsilon_i, \tag{2}$$

where s_i^* and t_i are the school and tier determining $q(\theta_i)$, with corresponding fixed effects α_1 and

[13] Abdulkadiroğlu et al. (2017b) adopt the convention that a lower admissions score is more favorable, in which case $q(\theta_i)$ is the maximum cutoff for schools in S_θ.

α_2. The running variable control function in the first stage is

$$h(r_i) = \kappa_0(s_i^*, t_i)r_i + \kappa_1(s_i^*, t_i)\max\{0, r_i - q(\theta_i)\}, \tag{3}$$

with an analogous parametrization, denoted $g(r_i)$, included in the second stage. Coefficients κ_0 and κ_1 also vary freely with s_i^* and t_i. X_i includes baseline covariates (race, gender, lunch status, and baseline math and reading scores).

The running variable control function described by equation (3) implements a local linear estimation strategy with slope changes at each school-tier cutoff. Bandwidths around these cutoffs are computed using the Imbens and Kalyanaraman (2012) procedure. Because the probability of an exam school offer varies across schools, years, and tiers, we compute cutoff-specific bandwidths using the sample that faces risk around each cutoff. Bandwidths are also computed separately for each outcome variable.[14] Equations (1) and (2) include dummies indicating the school and tier generating $q(\theta_i)$. The parameters in these equations also vary by application year, which ranges from 2009 to 2011. The estimation sample contains the set of students with running variable values in the bandwidths around the relevant qualifying cutoff.

An applicant who clears the admissions cutoff at the most-forgiving exam school for her tier spends more time at an exam school than an applicant who scores just below this cutoff. This can be seen in the first stage estimates reported in Panel A of Table 2, which show that cutoff-clearing marginal applicants spend roughly an additional 0.6 years at an exam school prior to taking the PLAN test in grade 10 and an additional 0.9 years prior to taking the ACT in grade 11. Second stage estimates of equation (2), reported in Panel B in the table, are negative; those for math are also statistically significant.

School-Specific Offer Instruments

We also use dummies for *individual* school offers as instruments for C_i. In a serial dictatorship, applicant i from tier t_i obtains an offer at school s by clearing $\tau_s(t_i)$ and failing to get an offer from a higher ranked school. Individual school offer dummies, D_{is}, can therefore be written

$$D_{is} = \begin{cases} 1[r_i \geq \tau_s(t_i)] & \text{if } D_{i\tilde{s}} = 0 \text{ for all } \tilde{s} \in S_\theta \text{ where } \tilde{s} \succ_i s, \\ 0 & \text{otherwise.} \end{cases}$$

[14] Estimates using variations on this bandwidth selection scheme are similar.

Because applicants receive a single offer, we have

$$D_i = \sum_s D_{is},$$

where s indexes exam schools. The full set of individual school offers can be used as instruments for C_i in an over-identified fuzzy RD two-stage least squares (2SLS) setup. The first-stage for school-specific offers likely differs by school, so we expect the resulting over-identified 2SLS estimates to be more precise than those using a single "any-offer" instrument.

An important piece of the school-specific offer strategy is control for differential risk of an offer by applicant type. In a match governed by serial dictatorship, applicants to school s face two sources of variation in offers. First, applicants must qualify for a seat at s by clearing the school-s cutoff. Second, qualified applicants to s may be seated elsewhere by qualifying at a higher-ranked choice. Abdulkadiroğlu et al. (2017b) use this fact to show that in an local-to-cutoffs empirical strategy of the sort widely used to exploit RD-type variation, the probability of receiving an offer at a CPS exam school is either 0, 1, or 0.5; this is the local-to-cutoffs SD propensity score for an offer at s. The theoretical appendix sketches this result.

This understanding of assignment risk under SD motivates a 2SLS empirical strategy that uses school-specific offers as instruments for exam school attendance, controlling for the propensity score for each. The resulting 2SLS procedure estimates β in the system

$$C_i = \sum_s \gamma_s D_{is} + \sum_s \eta_s(p_{is}) + X_i'\delta_1 + h(r_i) + \nu_i \tag{4}$$

$$Y_i = \beta C_i + \sum_s \lambda_s(p_{is}) + X_i'\delta_0 + g(r_i) + \varepsilon_i. \tag{5}$$

The first stage here, equation (4), includes a full set of D_{is} dummies indicating offers at exam school s; p_{is} is the relevant propensity score with effect η_s and λ_s. X_i includes baseline covariates (race, gender, lunch status, and baseline math and reading scores). The first stage also includes running variable controls of the form

$$h(r_i) = \kappa_0(t_i)r_i + \sum_s \kappa_{1s}(t_i)\max\{0, r_i - \tau_s(t_i)\}, \tag{6}$$

with $g(r_i)$ defined similarly in (5).[15] This model is estimated in a sample of applicants close to

[15] Coefficients κ_0 and κ_1 vary freely with tier t_i and year.

at least one school-specific cutoff. As detailed in the theoretical appendix, applicants to school s face non-degenerate risk at s if they're in the school-s bandwidth or the bandwidth around the lowest (easiest-to-clear) cutoff among the cutoffs for schools they prefer to s. We use the same cutoff-specific bandwidths as for the qualifying-cutoff RD.

As can be seen in Panel C of Table 2, over-identified estimates using school-specific offers are slightly more precise but otherwise similar to the just-identified estimates reported in Panel B. Although the precision gain from multiple instruments seems unimpressive, the availability of multiple instruments is useful for the analysis that follows.[16]

It's also noteworthy that the overidentification test statistics associated with equation (4) and (5) generate a decisive rejection for one math outcome and marginal rejection for another. Rejection implies that the underlying instruments used one-at-a-time produce statistically distinguishable IV estimates. Mismatch is a potential source of effect heterogeneity in this context: we might expect exam school effects to be lower or more negative when identified by offers to elite schools than when identified by offers from less selective schools.

4.2 Exam School Effects By School Type and Tier

Figure 1 shows that affirmative action beneficiaries who were offered seats at more selective exam schools are likely to have composite scores further from those of the median student at these schools. Mismatch may therefore be especially important at elite schools. To investigate differential effects by school selectivity, we distinguish enrollment effects at the four most selective schools (Northside, Payton, Young and Jones) from effects at other schools. Multiple school offers can be used to distinguish exam school treatment effects at different types of schools. To that end, let C_{1i} denote years of elite exam enrollment, while C_{2i} counts years enrolled at non-elite exam schools. A two-

[16]The appendix reports results from an exploration of threats to a causal interpretation of the estimates in Table 2. Follow-up scores are available for around 83% of the sample. Appendix Table A1 shows that exam school offers are largely unrelated to the probability of observing an outcome test score. Attrition is therefore unlikely to induce selection bias (for the theory behind this claim, see, e.g., Angrist (1997)). Table A2 presents encouraging evidence of covariate balance by offer status.

endogenous-variable model identifies separate causal effects of elite and non-elite enrollment:

$$
\begin{aligned}
C_{1i} &= \sum_s \gamma_{1s} D_{is} + \sum_s \eta_{1s}(p_{is}) + X_i'\delta_1 + h_1(r_i) + \nu_{1i}, \\
C_{2i} &= \sum_s \gamma_{2s} D_{is} + \sum_s \eta_{2s}(p_{is}) + X_i'\delta_2 + h_2(r_i) + \nu_{2i}, \\
Y_i &= \beta_1 C_{1i} + \beta_2 C_{2i} + \sum_s \lambda_s(p_{is}) + X_i'\delta_0 + g(r_i) + \varepsilon_i,
\end{aligned} \tag{7}
$$

where the first-stage effects vary for the two types of exam schools. The parameter β_1 is the causal effect of exposure to an elite school environment and β_2 is a non-elite exam school effect. The running variable controls are as in equation (6).

Estimates of elite and non-elite exam school effects using school-specific offer instruments, reported in Panel A of Table 3, are broadly similar to those from the single-effect model. In this case, however, we also see a substantial negative and statistically significant decline in PLAN reading scores due to elite enrollment, as well negative effects for math. The estimates in Table 3 are not consistent with the mismatch hypothesis: only one out of three negative effects are more pronounced at elite schools. The p-values in the table also show the elite/non-elite gap in estimated effects on PLAN reading to be marginally significant.

Another take on the mismatch story looks a different groups of *applicants*, distinguishing those most likely to benefit from affirmative action from others. We identify AA-affected applicants by running the CPS match with no tier reservations, comparing the sample of applicants seated at an exam school in this scenario to the sample seated in the (actual) match with tier reservations. The results of this comparison, detailed in appendix Table A3, show that 65% of beneficiaries reside in tier 1 and a further 33% are from tier 2.

Exam school effects by tier are computed by replacing the distinction between years of enrollment at elite and non-elite schools in equation (7) with endogenous variables distinguishing effects of exam school exposure for applicants in tiers 1 and 2 from enrollment effects on applicants in tiers 3 and 4. The instruments in this case are a full set of school-specific offers interacted with tier (for a total of 36 instruments; as before, the model controls for tier main effects). Here too, the estimates are inconsistent with mismatch. The lower panel of Table 3 shows larger and more negative estimates for low-tier applicants only for ACT math (though no contrast in estimated effects by tier reported in column 3 is significantly different from zero). These findings are also broadly consistent with the statistically significant negative reduced form effects of CPS exam school offers

on letter grades and selective college enrollment for low-tier applicants reported in Barrow et al. (2016).

Appendix Table A3 compares characteristics of the AA-affected group with those of exam school compliers. The AA-affected are almost all low-tier residents. But AA-affected and exam-offer compliers are about equally likely to be minority, low income, and to have baseline scores above those of all exam school applicants (this score gap exceeds that for compliers). Figures not reported in the table show that, among applicants for whom AA yields an additional offer, a little over half are exam-offer compliers. At the same time, among low-tier exam-offer compliers, about a third get an improved offer by virtue of AA. For tier 1 compliers, the proportion AA-affected hits 55%. Results for exam offer compliers therefore seem likely to indicative of effects on applicants affected by AA cutoff changes.

Finally, we look at exam treatment effects interacted with applicant position relative to the median, an analysis motivated by Sander (2004). These results are generated by replacing the two endogenous variables in equation (7) with years of exam enrollment and exam years times an indicator for above-median baseline scores (the above-median dummy is switched on when an applicant scores above the median at the exam school where they enroll). In this specification, large positive interaction terms support mismatch, especially when large enough to suggest achievement gains for high baseline achievers. The model with interaction terms is identified using school-specific offers and school-specific offers interacted with tier dummies as instruments.

As can be seen in Table 4, enrollment effects from the interacted model are close to those estimated without interaction terms. For example, these are estimated to be -0.099 for PLAN math and -0.059 for ACT math (ignoring interactions, that is, for below median applicants). Estimates adding tier-offer interactions to the instrument list appear in Panel B of Table 4. Other than a modest precision increase in Panel B, the estimates with and without tier interactions in the instrument list are similar. Estimates of the above-median interaction term are mostly positive, but marginally significant only for ACT math and only in Panel B (an estimate of 0.036). Moreover, these positive interactions terms are too small to imply a positive exam school effect for high-baseline applicants. On balance, therefore, the results in this table offer little support for preparedness-mediated mismatch suggested by Sander (2004).

5 Diagnosing Diversion

5.1 Counterfactual Destinies

The results in Table 3 seem inconsistent with the mismatch hypothesis. And the tests of Sander-style mismatch in Table 4 fail to uncover evidence of strong preparedness interactions or positive effects for high achievers. It seems likely, therefore, that negative exam school effects are generated by forces other than mismatch. We show here that negative exam school effects reflect *diversion*: exam school offers divert many applicants away from high-performing high schools in the Noble Network. Students who enroll at Noble instead of an exam school do better as a result, so exam school offers reduce achievement. The first step in this argument establishes Noble schools as a leading counterfactual "destiny" for offer-affected applicants denied an exam school seat.

The population of interest in this context consists of offer-compliers, that is, those who enroll in an exam school when offered a seat but not otherwise. To formalize this, let $W_i \in \{exam, traditional, charter, magnet, other\}$ denote the sector in which applicant i enrolls in the year in which she takes the ACT test (the last category is divided into smaller non-traditional sectors). Define *potential* sector enrollment variables W_{1i} and W_{0i}, indexed by exam offers D_i as

$$W_i = W_{0i} + (W_{1i} - W_{0i})D_i.$$

Potential exam school enrollment under alternative offer scenarios can now be written $C_{di} = 1\{W_{di} = exam\};\ d = 0, 1$. We're interested in $E[1\{W_{0i} = j\}|C_{1i} > C_{0i}]$. This is the counterfactual sector distribution for exam-offer compilers in the scenario where they're not offered an exam school seat. As in Abdulkadiroğlu et al. (2016), we estimate these counterfactual destinies by applying results in Abadie (2002).[17]

As can be seen in the first bar plotted in Figure 2, the most common counterfactual destiny is a traditional CPS school. But the charter sector, followed by magnet schools, is the most important destiny outside the traditional sector. Roughly a quarter of exam school offer compliers enroll in a charter school when not offered an exam school seat. The outlined portion of the charter segment shows that Noble schools account for over half of the charter sector counterfactual. Since applicants from tier 1 and tier 2 are disproportionately likely to benefit from AA, we also tabulate destinies

[17] Our procedure amounts to 2SLS estimation of a version of (2) with $(1 - D_i)1\{W_i = j\}$ as the outcome variable and $(1 - D_i)$ as the endogenous variable. Some applicants are non-compliant because they fail to take offers; In the language of Angrist et al. (1996), these are never-takers. Other applicants (always-takers) find an exam school seat even when not offered one through the official match.

for applicants from low tiers; these are appear in the second bar. The destiny distribution for these applicants shows charter and Noble shares similar to those for the sample as a whole.[18]

Motivated by the large number of exam school applicants who end up in these two non-traditional sectors when not offered an exam school seat, we compute destinies for the subsample of exam school applicants who also apply to magnet school or Noble schools. Even for the 26% of exam applicants who apply to a magnet school, charter schools remain the most important alternative after traditional public schools. And for the 13% of exam school offer compliers who also apply to the Noble Network, a Noble school is the dominant counterfactual destiny.

A graphical exam school RD hints at the importance of diversion. Figure 3 plots enrollment and ACT math scores against the exam school admissions composite centered at applicants' qualifying cutoffs (positive values indicate qualification for admission), after regression-adjusting for applicant risk-sets. In the sample of applicants who apply to exam and magnet schools, exam offers boost exam enrollment while decreasing magnet exposure sharply. This is apparent in the left side of Panel A in the figure. At the same time, the right-hand side of this picture shows a small downward shift in ACT math scores. Among applicants to both exam and Noble schools, sector substitution is even more pronounced: the drop in Noble exposure at the qualifying cutoff is almost as large as the increase in exam school exposure. The right-hand side of this figure also shows a marked decrease in test scores at the cutoff.

5.2 Multi-sector Models

We evaluate the hypothesis that exam school effects are explained by sector substitution using a multi-sector model with three endogenous variables. Let C_i, M_i, and N_i denote years of exam, magnet, and Noble enrollment, respectively. The magnet and Noble sectors both admit students via decentralized single-school lotteries. As in other decentralized lotteries, applicants may apply to more than one school. Since the likelihood of getting an offer is higher for applicants who apply to multiple schools, the multi-sector model controls for the set of schools applied to (risk set controls are included with other baseline covariates).

For each magnet applicant, we code an offer instrument, D_{im}, indicating offers at over-subscribed schools (defined as those where some students received an offer, while others didn't), in risk strata determined by residential proximity, tier, sibling status, and cohort. Focusing on over-subscribed

[18] "Outside CPS" in Figure 2 refers to enrollment in the test year. Some of those missing when not offered a seat return, while some of those offered a seat exit later, leading to the balanced attrition rates by offer status seen in Appendix Table A1.

schools yields a stronger first stage. Likewise, for each Noble applicant, we code an offer instrument, D_{in}, for any applicant who is offered a seat at an oversubscribed campus in risk strata determined by sibling status and cohort.[19]

A multi-sector model identified by three offer instruments can be described by this four-equation setup:

$$\begin{aligned}
C_i &= \sum_s \gamma_{cs} D_{is} + \gamma_{cm} D_{im} + \gamma_{cn} D_{in} + \sum_s \eta_{cs}(p_{is}) + X_i'\delta_{1c} + h_c(r_i) + \nu_{ci} \\
M_i &= \sum_s \gamma_{ms} D_{is} + \gamma_{mm} D_{im} + \gamma_{mn} D_{in} + \sum_s \eta_{ms}(p_{is}) + X_i'\delta_{1m} + h_m(r_i) + \nu_{mi} \\
N_i &= \sum_s \gamma_{ns} D_{is} + \gamma_{nm} D_{im} + \gamma_{nn} D_{in} + \sum_s \eta_{ns}(p_{is}) + X_i'\delta_{1n} + h_n(r_i) + \nu_{ni} \\
Y_i &= \beta_c C_i + \beta_m M_i + \beta_n N_i + \sum_s \lambda_s(p_{is}) + X_i'\delta_0 + g(r_i) + \varepsilon_i,
\end{aligned} \tag{8}$$

where γ's are three first-stage coefficients and β's are three sector effects. As in (7), this model includes exam-school offer propensity score and running variable controls, and a vector of covariates, X_i, that includes dummies for magnet and Noble risk sets. This model is estimated in the sample of exam school applicants who also apply to either a magnet or Noble school.

The multi-sector first stages reported in Table 5 show that the offer of a seat in each sector boosts enrollment in the sector offered, while reducing enrollment elsewhere. For instance, an exam sector offer increases exam enrollment by about 0.94 years before taking the ACT (this differs from the estimate in Table 2 because the sample here includes exam school applicants that also applied either to magnet or Noble campuses and because the model measures effects relative to students not offered a seat in any of these three sectors. Magnet and Noble offers increase own-sector enrollment by about 0.70 years.

The estimates in the first column of Table 5 show that magnet and Noble offers reduce exam school attendance about equally. At the same time, exam offers pull applicants away from Noble schools a little more than from magnet schools (compare -0.20 and -0.13 in columns 2 and 3 of the first row of the table). Exam and Noble offers reduce magnet attendance similarly (compare -0.13 and -0.11 in column 2). Exam and magnet offers reduce Noble enrollment, but exam offers pull students away from Noble much more strongly than do offers from magnets (compare -0.200 and -0.067 in column 3).

[19] Appendix Tables A1 and A2 report statistics related to attrition and balance for the magnet and Noble research designs.

As a benchmark, Table 6 reports single-sector estimates of exam school effects analogous to those reported in Table 2, computed in the sample used to estimate the multi-sector model. These appear in odd numbered columns in the table. The math effects here are somewhat more negative than those reported in Table 2. For example, the estimated effect of exam school years on ACT math, reported in column 5 is -0.120 in Panel A (compared with -0.058 in Table 2). The over-identified estimates reported in Panel B of Table 6 again show a modest increase in precision. The modest increases in negative magnitudes seen here are likely attributable to larger cross-sector substitution in the sample including only exam school applicants who also applied to a magnet or Noble school.

2SLS estimates of multiple sector effects, reported in even-numbered columns in Table 6, suggest that negative effects of exam school exposure on math scores and PLAN reading are explained by the combination of large gains from Noble enrollment and exam-induced substitution away from this sector. In column 2, for example, the Noble estimate ranges from 0.37 to 0.43, while the associated exam school estimates are small and not significantly different from zero. This pattern repeats in columns 5 and 6, with large positive effects of Noble exposure and a commensurate decline in exam school effects, to the point where the latter are not significantly different from zero in the multi-sector model.

Estimates from the just-identified version of this model, reported in Panel A, suggest magnet schools may also boost PLAN math. But these estimates are considerably smaller than the corresponding Noble effects and are not significantly different from zero in Panel B. Estimates in columns 4 and 8 suggest that Noble attendance boosts reading as well as math. Single-sector estimates of exam school effects on PLAN reading effects in column 3 are also negative (equal to -0.136 and statistically significant in Panel A and -0.064 in Panel B). Though not always significantly different from zero, these estimates can be explained by the combination of individual sector effects in column 4 and substitution patterns like those in Table 5.

Consistent with the estimated effects on reading in Tables 2, 3, and 4, the benchmark estimates in column 7 for ACT reading show no evidence of an exam school effect. Estimates of the multivariate model in column 8 show positive Noble effects, but these are not large enough for sector small enough for sector substitution to generate a negative estimate in column 7, when Noble and magnet enrollment are ignored.

6 Mismatch vs. Diversion

6.1 Theoretical Framework

The estimates in Table 6 suggest exam school effects are driven primarily by the quality of schools attended by applicants not offered an exam school seat. This section considers the possibility that exam-Noble substitution accounts for *all* effect heterogeneity observed in our data, including variation across subgroups defined by tier and baseline achievement. This strong and simple null hypothesis is measured against an alternative that allows for possible match effects.

Our exploration of treatment effect heterogeneity is motivated by the distinction between school quality and student-specific match effects in the Ellison and Pathak (2016) model of human capital production. This model postulates a production function that depends on student background (ability or preparedness) and a match component that decreases in the distance between preparedness and curriculum difficulty. In particular, suppose that the expected outcome for student i of preparedness level a_i is described by:

$$V(a_i, q_s, \varphi_s) = f(a_i, q_s) - \zeta(a_i - \varphi_s)^2, \tag{9}$$

where q_s is school quality, φ_s is curriculum characterized by the median preparedness of the students school s serves, and ζ parameterizes the relative importance of student-curriculum matching. Let

$$a_i = x_i'\psi,$$

where x_i is a fully saturated description of discrete applicant characteristics.

This production function yields a testable distinction between school quality effects common to all students and a set of interaction terms generated by mismatch. To see this, suppose that the non-match-related component of human capital production is

$$f(a_i, q_s) = q_s + a_i = q_s + x_i'\psi. \tag{10}$$

The mismatch term in (9) can now be written

$$\zeta(x_i'\psi - \varphi_s)^2 = x_i'\delta_2 + x_i'\delta_1\varphi_s + \varrho_s, \tag{11}$$

where $x_i'\delta_2 = \zeta\psi' x_i x_i'\psi$, $\delta_1 = -2\zeta\psi$, and $\varrho_s = \zeta\varphi_s^2$.

Combining terms in (10) and (11) produces a linear-in-parameters specification,

$$\begin{aligned} y_{is} &= (q_s + \varrho_s) + x_i'[\psi + \delta_2] + \varphi_s x_i'\delta_1 \\ &= \mu_s + x_i'\pi + \varphi_s x_i'\delta_1, \end{aligned} \tag{12}$$

where $\mu_s = q_s + \varrho_s$ and $\pi = \psi + \delta_2$. Equation (12) casts mismatch as a source of interactions between school selectivity (φ_s) and applicant characteristics related to ability or preparedness. School quality effects, by contrast, are additive, contributing to a school-specific intercept.[20]

We test for mismatch using an empirical model that omits interaction terms but allows for differing first-stage effects of exam school offers on Noble enrollment. Specifically, we ask whether subgroup variation in the *Noble* first stage explains heterogeneous *exam school* offer effects. This test is constructed via 2SLS estimation of versions of

$$Y_i = \sum_s C_{is}\mu_s + X_i'\pi + \sum_s C_{is}X_i'\delta_s + \varepsilon_i, \tag{13}$$

where C_{is} indicates enrollment at school s with quality effect μ_s. Coefficient δ_s parameterizes student-school interactions. Equation (13) is identified by using D_{is} and $D_{is}X_i$ as instruments for C_{is} and $C_{is}X_i$, respectively.

Suppose that matching matters little, so school quality effects are the same for all students, and captured by μ_s. When $\delta_s = 0$, the reduced form becomes:

$$E[Y_i|X_i, D_{is}, D_{in}, D_{im}] = \sum_s E[C_{is}|X_i, D_{is}, D_{in}, D_{im}]\mu_s + X_i'\pi. \tag{14}$$

Suppose further that only Noble attendance matters, a sector with common school effect μ_n. Let Z_i be any instrument that changes Noble enrollment. Equation (14) then simplifies to:

$$E[Y_i|X_i, Z_i] = \mu_0 + E[N_i|X_i, Z_i]\mu_n + X_i'\pi. \tag{15}$$

Finally, let $\rho_Z(X)$ and $\phi_Z(X)$ denote the reduced form and first-stage for an X-specific Noble effect

[20] The school quality component of the production function described by (10) need not be additive. Charter schools, for example, might boost achievement more for low-ability students than for others. On the other hand, the absence of substantial school/student interactions of any sort weighs against mismatch.

identified by instrument Z. Differencing by Z_i, we obtain:

$$\rho_Z(X_i) = (E[N_i|X_i, Z_i = 1] - E[N_i|X_i, Z_i = 0])\mu_n = \phi_Z(X)\mu_n. \tag{16}$$

Equation (15) implies that the reduced form effect of instrument Z_i for *each* X_i and *any* Z_i is driven by the corresponding covariate-specific Noble first stage generated by Z. In other words, the covariate-specific reduced form effect of *any* offer should be proportional to the corresponding covariate-specific first stage effect. The slope of the line linking offer-status differences outcomes to the corresponding differences in average years enrolled at Noble is the causal effect of interest, μ_n.

6.2 Covariate-Specific Estimates Explained

The proportionality hypothesis embodied in equation (16) is tested in two ways. First, we use both exam and Noble offers as instruments for a Noble effect, asking whether the resulting IV estimates are similar. As can be seen in Panel A of Table 7, use of Noble offers to instrument Noble exposure generates large precisely estimated effects on math ranging from about 0.34 to 0.41. Effects on PLAN reading are also large at about 0.35, while the ACT reading effect is a marginally significant 0.148. The sample here is the same as used for the multi-sector estimates reported in Tables 5 and 6.

Estimates using *exam school* offers to instrument *Noble enrollment*, reported (with standard errors) in the second pair of rows in Table 7, are remarkably similar those using Noble offers. Compare, for example, 0.33 with 0.41 for PLAN math and 0.46 with 0.34 for ACT math. Not surprisingly, the estimates using exam offers to instrument Noble enrollment are less precise than those using Noble instruments, but both sets of estimates show large and at least marginally significant effects on the same three outcomes for which exam school effects are most pronounced in Tables 2-4. The overidentification test statistic associated with a 2SLS procedure that uses both sorts of offers as instruments yields a formal test of offer-IV equality.[21] As can be seen in the third pair of rows in Table 7, 2SLS estimates generated by this procedure fall in between those using either offer type alone. The p-values arising from overidentification tests associated with these estimates offer no hint of divergence between estimates using Noble and exam school offers as instruments.

A second exploration of the proportionality hypothesis embodied in (16) uses offer instruments

[21] Angrist and Pischke (2009) discusses this interpretation of overidentification tests.

interacted with covariates. This identification strategy tests the equality of IV estimates generated for different covariate subgroups. Specifically, we instrument Noble enrollment using D_{in} and a set of 9 offer-covariate interactions for lunch status (interactions for free and reduced-price), race (interactions for black and Hispanic), gender, baseline test scores (an interaction for low), and tier dummies (interactions for Tiers 1-3). With 9 interactions and an offer main effect in the instrument list, the overidentification test has 9 degrees of freedom, and tests whether estimates using different covariate cells are equal.

The covariate interaction strategy built around Noble offers generates estimates much like those using a single Noble offer dummy. Compare, for example, the estimate of approximately 0.35σ for PLAN math in the first pair of rows in Panel B of Table 7 with 0.41 at the top of the table. Estimates for the remaining three outcomes at the top of Panels A and B are about as close. With one exception, the overidentification test statistics associated with these estimates are far showing evidence of effect heterogeneity.

Panel B of Table 7 also reports estimates interacting *exam offers* with the same covariate subgroups. Because the F statistics here are low enough to raise concerns about finite-sample bias, this part of the table shows LIML estimates rather than 2SLS. Although the LIML estimates of Noble effects on PLAN scores are less precise than the corresponding just-identified estimates using exam offers, the LIML estimates in Panel B are remarkably to those generated by a single exam offer dummy.[22] Here, the over-identification p-values are mostly lower than the p-values generated by interacting Noble offers with covariates. This partly reflects the relative imprecision of IV estimates using exam offers interacted with covariates to instrument Noble enrollment.

Finally, the bottom of Table 7 reports the results of combining both offer types and the full set of covariate interactions in a single 2SLS procedure using 20 instruments. Not surprisingly, the resulting estimates are in between the corresponding exam-offer and Noble-offer estimates including covariate interactions as instruments. The over-identification test statistic associated with these estimates has 19 degrees of freedom. None of the test statistics generated by this specification offer evidence against the proportionality hypothesis. That is, the underlying IV estimates are consistent with constant covariate-specific IV estimates that are also similar across offer types.

This remarkable homogeneity is visible in Figure 4, which offers a graphical representation of the model with 20 instruments. Specifically, the figure plots covariate-specific reduced form estimates

[22] Angrist and Pischke (2009) discuss the approximate unbiasedness of LIML when instruments are weak. The precision of 2SLS increases as the number of instruments increase. Reduced precision in moving from the just-identified to over-identified model is a consequence of the switch to LIML.

for ACT math against the corresponding first stage estimates for Noble exposure, constructed using both exam and Noble offer instruments. Red numbers in the figure label results using the Noble offer instrument, while black numbers label those using exam offers. Each reduced form and first stage estimate in the figure is fit one-at-a-time for a variety of (overlapping) covariate cells. For example, points labeled with "1" show first stage and reduced form estimates for applicants who qualify for a free lunch. Solid dots in the figure plot results using any-offer instruments across all covariate cells.[23]

Consistent with the substitution patterns in Table 5, the Noble lottery instrument generates large positive first stages for Noble enrollment and a corresponding set of positive reduced form estimates. At the same time, first-stage effects of exam offers on Noble enrollment are mostly negative, as are the corresponding reduced form estimates. Importantly, however, the points from both research designs fall roughly along a straight line with a slope equal to about 0.36.[24] This "visual IV" estimate indeed appears to rationalize the collection of reduced form and first stage estimates plotted in the figure, whether driven by Noble or exam school offers.

7 Exam Schools Elsewhere

A brief examination of exam school effects in a second (anonymous) large urban district (LUD) generates results remarkably consistent with those from Chicago. Like Chicago, most LUD students are nonwhite. LUD exam schools also admit students using a centralized assignment DA-based scheme similar to that used in the CPS match. And LUD has a robust charter sector, with many schools following the same sort of "No Excuses" practices characteristic of the Noble network. Our analysis of LUD looks at exam school effects for applicants who originate in charter middle schools, and are therefore likely to have character enrollment destinies when not offered an exam school seat.

The analysis sample for this investigation includes applicants who applied for an exam school seat from 2003 to 2015. Outcome variables come from statewide math and English assessments, with scores standardized to the district mean and standard deviation. Our analysis focuses on students enrolled in a charter school at the time they applied for an exam school seat. Two-thirds

[23] All estimates in the figure control for covariate-group main effects.

[24] This estimate comes from a weighted fit, with weights given by the reciprocal of the homoskedastic sampling variance of the reduced form estimates. Note that reduced form and first stage estimates in mutually exclusive covariate splits, such as for men and women, average to the any-offer effects plotted in the figure. The slope reported in the text therefore omits the any-offer estimates.

of the exam-school offer compliers in this group end up in a charter high school when not offered an exam school seat. The research design for LUD controls for non-random variation in exam school offer risk in a 2SLS setup analogous to that described by equation (5) for CPS.

2SLS estimates of exam school effects for the sample of charter-originating LUD applicants show that LUD exam enrollment reduces math scores by around 0.2 standard deviations per year of enrollment. This can be seen in Table 8, which reports results from tests taken at the end of the first school year after the application year and from a larger sample that includes additional assessments. The negative 2SLS estimates for math reported in Table 8 are significantly different from zero, mirroring those from Chicago.

8 Conclusions

Diversity concerns have long taken center stage in public discussion of exam school access (see, e.g., Boston Latin Needs More Diversity, 2016 and Confronting Segregation in NYC Schools, 2017). The mismatch hypothesis is a touchstone in this debate. Yet our findings show that as far as achievement goes, exam school value-added is driven more by alternative schooling options than by measures of applicant preparedness. More generally, we find little in the way of match effects of any kind. Remarkably, a simple additive model with constant charter, exam, and magnet effects rationalizes the pattern of achievement gaps for applicants who have applied to more than one sector.

The test score effects reported here are echoed in studies of CPS exam school effects on other outcomes. In particular, Barrow et al. (2016) show that tier 1 applicants to CPS exam schools are substantially less likely to attend a selective college when they cross an exam school admissions cutoff. At the same time, Davis and Heller (2019) report positive college effects for Noble's flagship campus in 2003-05. It therefore seems likely that Noble enrollment explains negative exam school effects on college selectivity as well as on test scores.

Chicago is not alone among large urban districts in featuring important exam and charter sectors. It's noteworthy, therefore, that data from another large urban district shows evidence of the same sort of consequential sector substitution among exam school applicants originating in charter schools. These results suggest that policies focused solely on selective school diversity or access, rather than school quality in the form of causal value-added, may do little to increase learning for disadvantaged groups.

References

Abadie, A. (2002). Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models. *Journal of the American Statistical Association*, 97(457):284–292.

Abdulkadiroğlu, A., Angrist, J., Dynarski, S., Kane, T., and Pathak, P. (2011). Accountability and Flexibility in Public Schools: Evidence from Boston's Charters And Pilots. *Quarterly Journal of Economics*, 126(2):699–748.

Abdulkadiroğlu, A., Angrist, J. D., Hull, P. D., and Pathak, P. A. (2016). Charters Without Lotteries: Testing Takeovers in New Orleans and Boston. *American Economic Review*, 106(7):1878–1920.

Abdulkadiroğlu, A., Angrist, J. D., Narita, Y., and Pathak, P. A. (2017a). Research Design meets Market Design: Using Centralized Assignment for Impact Evaluation. *Econometrica*, 85(5):1373–1432.

Abdulkadiroğlu, A., Angrist, J. D., Narita, Y., and Pathak, P. A. (2019). Breaking Ties: Regression Discontinuity Designs Meets Market Design. IZA Discussion Paper 12205.

Abdulkadiroğlu, A., Angrist, J. D., Narita, Y., Pathak, P. A., and Zarate, R. A. (2017b). Regression Discontinuity in Serial Dictatorship: Achievement Effects at Chicago's Exam Schools. *American Economic Review: Papers and Proceedings*, 107(5).

Abdulkadiroğlu, A., Angrist, J. D., and Pathak, P. A. (2014). The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica*, 82(1):137–196.

Angrist, J. (1997). Conditional independence in sample selection models. *Economics Letters*, 54(2):103–112.

Angrist, J., Pathak, P., and Walters, C. (2013). Explaining Charter School Effectiveness. *American Economic Journal: Applied Economics*, 5(4):1–27.

Angrist, J. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics*. Princeton University Press, Princeton.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455.

Arcidiacono, P. and Lovenheim, M. (2016). Affirmative Action and the Quality-Fit Tradeoff. *Journal of Economic Literature*, 54(1):3–51.

Barrow, L., Sartain, L., and de la Torre, M. (2016). The Role of Selective High Schools in Equalizing Educational Outcomes: Heterogeneous Effects by Neighborhood Socioeconomic Status. Federal Reserve Bank of Chicago, WP 2016-17.

Boston Globe (2016). Boston latin needs more diversity. Boston Globe editorial, September 28.

Chabrier, J., Cohodes, S., and Oreopoulous, P. (2016). What Can We Learn from Charter School Lotteries? *Journal of Economic Perspectives*, 30(3):57–84.

Cullen, J. B., Jacob, B. A., and Levitt, S. (2006). The Effect of School Choice on Participants: Evidence from Randomized Lotteries. *Econometrica*, 74(5):1191–1230.

Davis, M. and Heller, B. (2019). No Excuses Charter Schools and College Enrollment: New Evidence from a High School Network in Chicago. *Education Finance and Policy*, 14(3):414–440.

Dobbie, W. and Fryer, R. (2013). Getting beneath the Veil of Effective Schools: Evidence from New York City. *American Economic Journal: Applied Economics*, 5(4):28–60.

Dobbie, W. and Fryer, R. (2014). The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools. *American Economic Journal: Applied Economics*, 6(3):58–75.

Dobbie, W. and Fryer, R. (2015). The Medium-Term Effects of High-Achieving Charter Schools on Non-Test Score Outcomes. *Journal of Political Economy*, 123(5):985–1037.

Dur, U., Pathak, P. A., and Sönmez, T. (2016). Explicit vs. Statistical Preferential Treatment in Affirmative Action: Theory and Evidence from Chicago's Exam Schools. NBER Working Paper 22109.

Ellison, G. and Pathak, P. A. (2016). The Efficiency of Race-Neutral Alternatives to Race-Based Affirmative Action: Evidence from Chicago's Exam Schools. NBER Working Paper 22589.

Heckman, J., Hohmann, N., Smith, J., and Khoo, M. (2000). Substitution and Drop Bias in Social Experiments: A Study of an Influential Social Expeirment. *Quarterly Journal of Economics*, 115:651–694.

Imbens, G. and Kalyanaraman, K. (2012). Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *Review of Economic Studies*, 79(3):933–959.

Kahlenberg, R. (2014). Elite, Separate, and Unequal: New York City's Top Public Schools Need Diversity. NY Times, June 22.

Kirkeboen, L., Leuven, E., and Mogstad, M. (2016). Field of Study, Earnings, and Self-Selection. *Quarterly Journal of Economics*, 131(3):1057–1111.

Kline, P. and Walters, C. (2016). Evaluating Public Programs with Close Substitutes: The Case of Head Start. *Quarterly Journal of Economics*, 131:1795–1848.

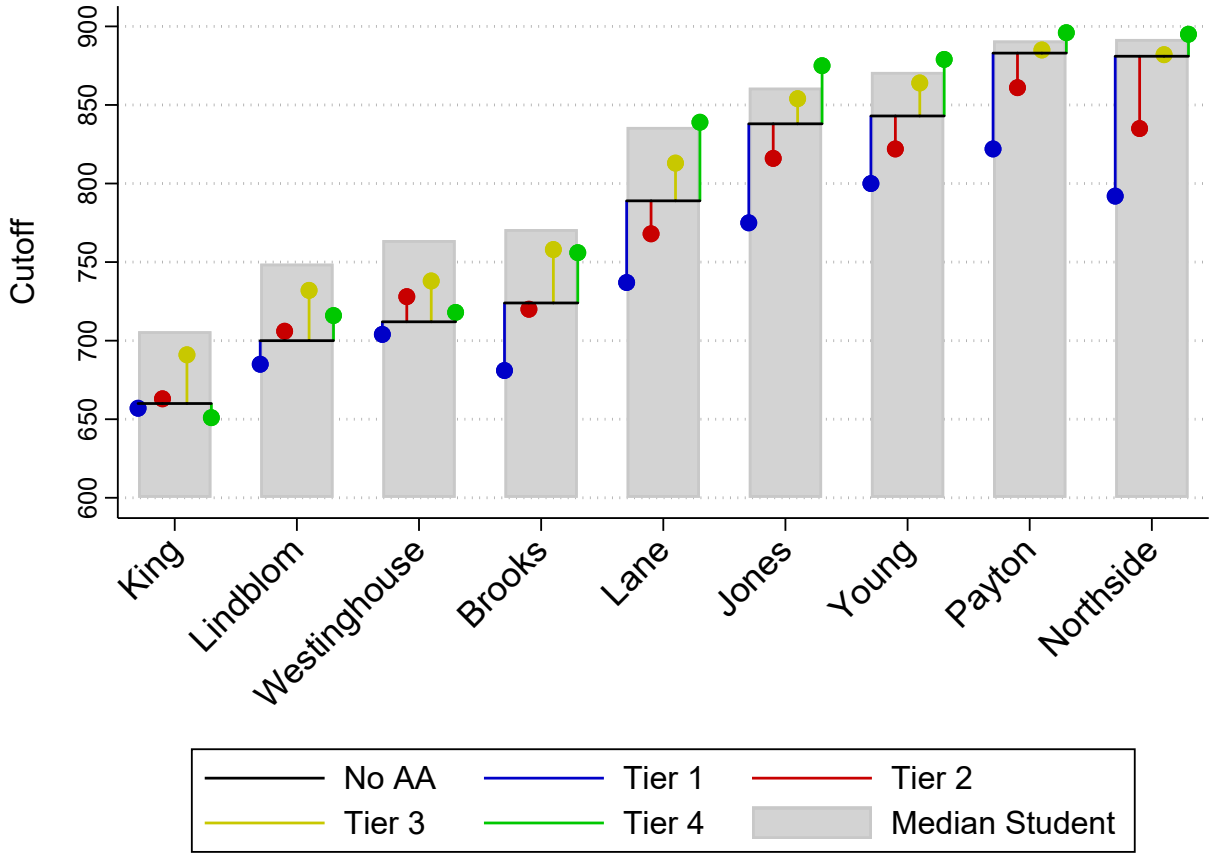
New York Times (2017). Confronting Segregation in New York City Schools. New York Times, May 15.

OCR (2011). Guidance on the Voluntary Use of Race to Achieve Diversity and Avoid Racial Isolation in Elementary and Secondary Schools. Memorandum of U.S. Department of Justice and U.S. Department of Education, Office of Civil Rights, Available at: http://www2.ed.gov/about/offices/list/ocr/docs/guidance-ese-201111.pdf and Last accessed: April 15, 2017.

Sander, R. H. (2004). A Systematic Analysis of Affirmative Action in American Law Schools. *Stanford Law Review*, 57(2):367–483.

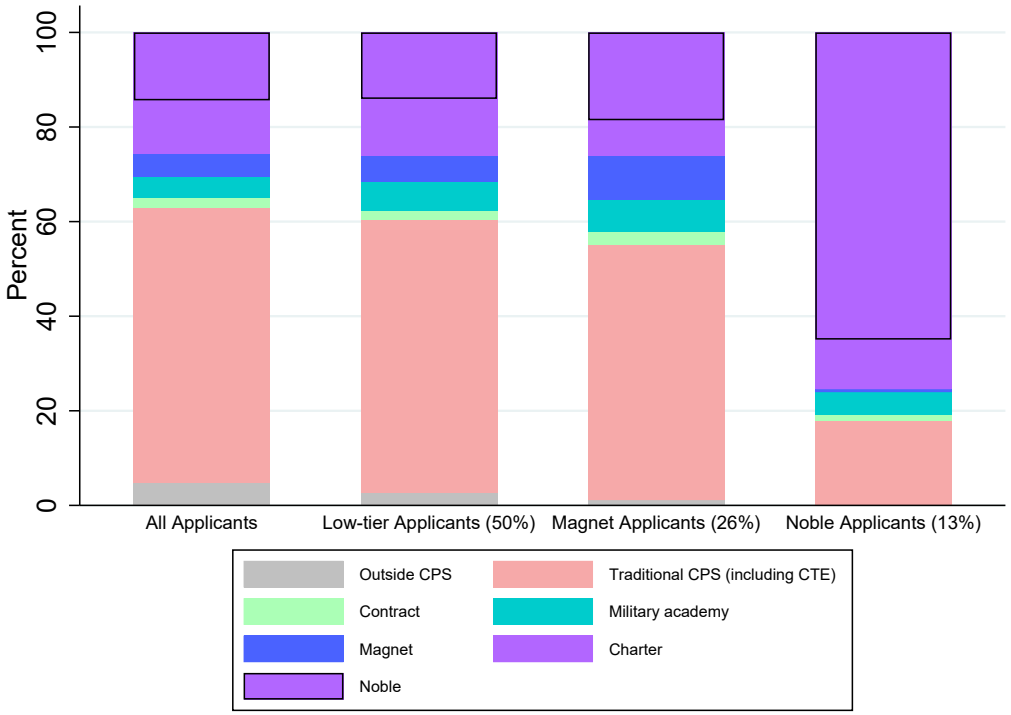
Sander, R. H. and Taylor, S. (2012). *Mismatch: How Affirmative Action Hurts Students It's Intended to Help, and Why Universities Won't Admit it.* Basic Books.

Figure 1: Chicago Admission Cutoffs by Tier (2011)



Note: This figure shows the admission cutoffs for each tier at each exam school in Chicago in 2011, along with the cutoff that would have obtained without tier-based AA. The grey bars indicate the median admissions score among those offered a seat.

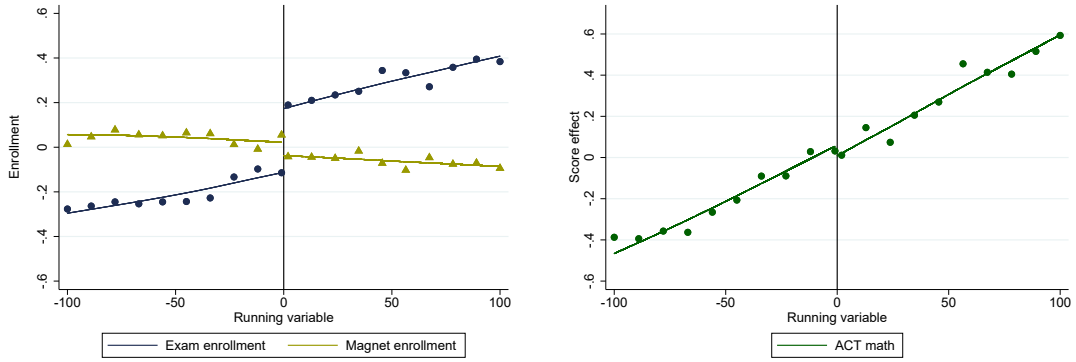
Figure 2: Enrollment Destinies



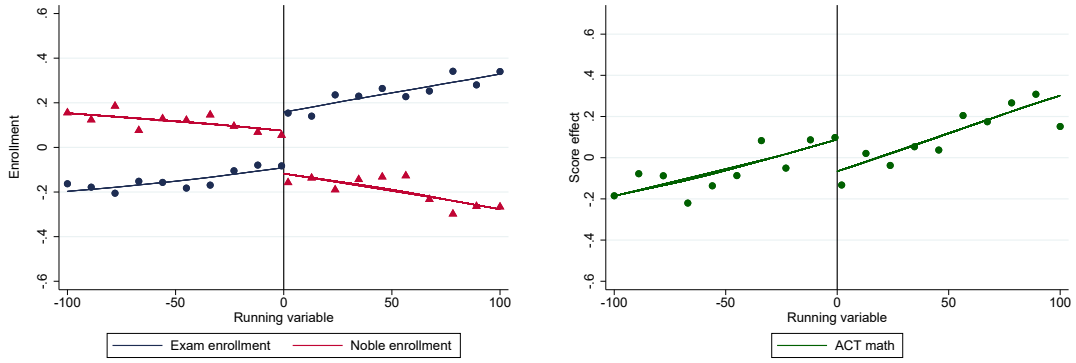
Note: This figure shows the enrollment destinies of exam school compliers when not offered an exam school seat. Enrollment compliers are applicants who attend an exam school when offered a seat but not otherwise. The 1st bar plots non-exam destinies for all rejected exam school applicants, and the 2nd bar plots for low-tier applicants. The 3rd bar plots non-exam destinies for rejected exam school applicants who also applied to a magnet school. The 4th bar plots non-exam destinies for rejected exam school applicants who also applied to a Noble school. Destinies are estimated as in Abdulkadiroğlu et al. (2014). Enrollment rates are measured in the fall of the year of the ACT math test. A student who does not take ACT math is counted as Outside CPS.

Figure 3: Effects at Qualifying Cutoffs

Panel A: Magnet and Exam School Applicants

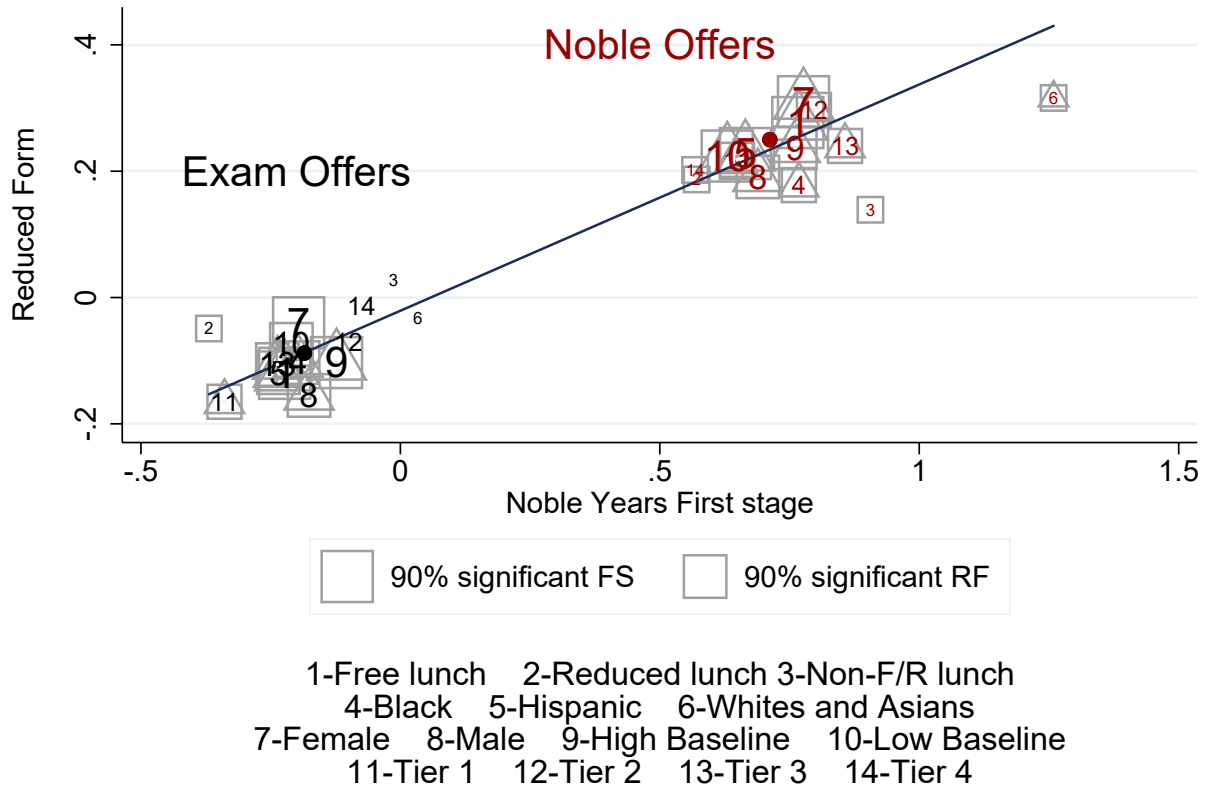


Panel B: Noble and Exam School Applicants



Note: The left side of this figure plots enrollment at exam, Noble, and magnet schools near qualifying (exam-school) cutoffs against the exam school running variable. Applicants who clear their qualifying cutoff are offered an exam school seat. The right side plots ACT math against tie-breaker over the same range. All variables are plotted after partialing out risk sets. The x-axis in both columns is centered at the qualifying cutoff for each applicant. Plotted points are averages in 10-point windows; lines in the plots are estimated conditional mean functions smoothed using local linear regression (LLR).

Figure 4: Covariate VIV for the Effects of Noble Enrollment on ACT Math



Note: This is a visual instrumental variables (VIV) figure that plots reduced form effects of exam school and Noble offers on ACT math against first stage effects of exam school and Noble offers on Noble years, separately for a set of 14 covariate-defined groups. Exam offer effects are plotted in black; Noble offer effects are plotted in red. Covariate-specific estimates are computed one at a time. Estimates for all applicants are plotted with dots, these are omitted from the VIV calculation. The slope of the solid line through these estimates is 0.362, a VIV estimate of the effects of Noble enrollment, using offer-covariate interactions as instruments.

Table 1. CPS 9th Graders by Sector

	Traditional schools (1)	Exam schools (2)	Magnet schools (3)	Charter schools (4)	Noble Network (5)
Black	0.444	0.332	0.295	0.578	0.437
Hispanic	0.428	0.304	0.553	0.370	0.520
White	0.075	0.233	0.092	0.022	0.016
Asian	0.033	0.086	0.043	0.013	0.012
Female	0.489	0.578	0.489	0.480	0.510
Free lunch	0.753	0.448	0.725	0.788	0.763
Reduced price lunch	0.084	0.114	0.100	0.098	0.128
Special education	0.166	0.091	0.139	0.178	0.157
Tier 1	0.267	0.166	0.176	0.337	0.349
Tier 2	0.284	0.204	0.315	0.301	0.299
Tier 3	0.264	0.274	0.300	0.248	0.241
Tier 4	0.186	0.355	0.209	0.114	0.111
Baseline math	-0.033	1.445	0.078	-0.289	-0.149
Baseline reading	-0.034	1.242	0.087	-0.232	-0.080
N	89,542	9,020	5,477	17,828	7,146

Notes: This table reports average characteristics of 9th grade CPS students attending traditional, exam, magnet, charter, and Noble Network schools between 2010 and 2012 (applying for seats in 2009-11). Students are assigned the tier for the census tract of their residence, as recorded in CPS enrollment files. Baseline scores are standardized to the CPS test-taking population.

Table 2. Fuzzy RD (2SLS) Estimates of Exam School Enrollment Effects

	PLAN		ACT	
	Math	Reading	Math	Reading
	(1)	(2)	(3)	(4)
	Panel A: First stage for any exam offer			
Clears qualifying cutoff	0.645***	0.613***	0.906***	0.952***
	(0.025)	(0.025)	(0.036)	(0.037)
	Panel B: 2SLS using exam offer			
Exam years	-0.087***	-0.053	-0.058***	-0.003
	(0.032)	(0.038)	(0.022)	(0.023)
N	16,451	16,855	17,826	16,889
	Panel C: 2SLS using school-specific offers			
Exam years	-0.091***	-0.059	-0.056***	0.005
	(0.031)	(0.037)	(0.021)	(0.023)
Overid (DF=8) p-value	0.02	0.08	0.00	0.42
F	77.0	74.5	73.0	73.1
N	17,528	17,708	18,805	18,162

Notes: This table reports 2SLS estimates of exam school exposure effects for applicants applying 2009-11. The endogenous variable is years of enrollment between application and test date. The first stage for any exam school offer is reported in Panel A; the corresponding 2SLS estimates appear in Panel B. These models are identified by variation around the qualifying cutoff. The estimates in Panel C use a full set of school-specific offers as instruments, controlling for offer risk. All models include the set of baseline covariates listed in Table A2. Models in Panels A and B also include school-by-tier-by year fixed effects and school-by-tier-by year running variable controls. Models in Panel C also include year-by-tier fixed effects and year-by-tier running variable controls as described in the text. The table also shows the first stage F statistic for over-identified models. Bandwidths use the formulas in Imbens and Kalyanaraman (2012). Robust standard errors are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3. Exam Effects by School Type and Tier

	PLAN		ACT	
	Math	Reading	Math	Reading
	(1)	(2)	(3)	(4)
	Panel A: Effects by school type			
Elite exam years enrolled	-0.086**	-0.107**	-0.043*	0.006
	(0.038)	(0.045)	(0.023)	(0.026)
Non-elite exam years enrolled	-0.092***	-0.050	-0.061***	0.006
	(0.031)	(0.037)	(0.021)	(0.023)
p-value (elite=non-elite)	0.79	0.05	0.23	0.99
F (elite exam)	206.3	198.9	214.7	213.8
F (non-elite exam)	91.6	87.5	85.2	86.9
N	17,529	17,709	18,805	18,162
	Panel B: Effects by tier			
Estimates for low tier applicants	-0.073*	-0.063	-0.074**	0.008
	(0.042)	(0.052)	(0.030)	(0.031)
Estimates for high tier applicants	-0.094**	-0.055	-0.024	-0.007
	(0.042)	(0.049)	(0.028)	(0.031)
p-value (low tier = high tier)	0.72	0.91	0.22	0.72
F (low tier)	10.8	10.2	10.1	10.4
F (high tier)	11.7	11.7	11.2	11.4
N	17,528	17,708	18,805	18,162

Notes: This table reports 2SLS estimates of the effects of exam school exposure for applicants applying 2009-11 using school-specific offers as instruments for any exam school enrollment. Panel A splits enrollment effects by school type, distinguishing between effects of the number of years enrolled in an elite exam school (Northside, Payton, Young, and Jones) and effects of the number of years enrolled in a non-elite exam school (King, Lindblom, Westinghouse, Brooks, and Lane). Panel B splits applicants by tier, distinguishing effects of years enrolled for applicants in tiers 1-2 from effects of years enrolled for applicants in tiers 3-4. Estimates in this panel add instruments by interacting school-specific offers with tier dummies. All models include the set of baseline covariates listed in Table A2, year-by-tier fixed effects, and year-by-tier running variable controls, as described in the text. The table also reports Angrist and Pischke first-stage F statistics for models with multiple endogenous variables and the p-value for a test of subgroup effect equality. Robust standard errors appear in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4. Exam Effects for Applicants Above and Below the Median

	PLAN		ACT	
	Math	Reading	Math	Reading
	(1)	(2)	(3)	(4)
Panel A: 2SLS using school-specific offers				
Exam years	-0.099***	-0.062*	-0.059***	0.007
	(0.032)	(0.038)	(0.021)	(0.024)
Exam years x above median	0.024	0.044	0.011	-0.003
	(0.039)	(0.045)	(0.023)	(0.025)
p-value (exam years+interaction=0)	0.10	0.75	0.12	0.89
F (exam years)	87.6	85.5	82.2	82.7
F (exam years x above median)	86.0	72.4	87.2	89.0
N	17,462	17,646	18,740	18,099
Panel B: 2SLS using school-specific offers interacted with tier dummies				
Exam years	-0.092***	-0.063*	-0.054***	0.003
	(0.030)	(0.036)	(0.020)	(0.022)
Exam years x above median	0.028	0.041	0.036**	-0.003
	(0.028)	(0.033)	(0.016)	(0.019)
p-value (exam years+interaction=0)	0.10	0.62	0.48	1.00
F (exam years)	22.0	21.3	20.9	21.1
F (exam years x above median)	37.5	35.4	36.1	38.4
N	17,462	17,646	18,740	18,099

Notes: This table reports 2SLS estimates of the effects of exam school exposure in models with an interaction term for applicants whose baseline scores are above those of the median student enrolled in the exam school attended. Estimates in Panel A use school-specific offer dummies as instruments. Estimates in Panel B use school-specific offers interacted with tier dummies. All models include the set of baseline covariates listed in Table A2, year-by-tier fixed effects, and year-by-tier running variable controls, as described in the text. The table also reports p-values for tests of whether effects for applicants above the median are zero. The table also reports Angrist and Pischke first-stage F statistics for models with multiple endogenous variables. Robust standard errors appear in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5. First Stage Estimates for Multi-Sector Models

	Exam Years	Magnet Years	Noble Years
	(1)	(2)	(3)
Clears qualifying cutoff	0.938***	-0.131***	-0.200***
	(0.058)	(0.037)	(0.042)
Magnet lottery offer	-0.197***	0.673***	-0.067*
	(0.053)	(0.070)	(0.036)
Noble lottery offer	-0.178***	-0.110***	0.724***
	(0.055)	(0.032)	(0.077)
N	6,954	6,954	6,954

Notes: This table reports first stage estimates for three enrollment variables, exam years, magnet years, and noble years, estimated in the ACT math sample. The instruments are dummies indicating any exam offer (determined by clearing the qualifying cutoff), lottery offers at over-subscribed magnet schools, and lottery offers at over-subscribed Noble schools. First stage equations include running variable controls and risk set dummies for exam, magnet, and Noble offers, as described in the text, and the covariates listed in Table A2. All models include school-by-year-by-tier running variable controls. Robust standard errors appear in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6. 2SLS Estimates of Sector Effects

	PLAN				ACT			
	Math		Reading		Math		Reading	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Any exam offer (clears qualifying cutoff)								
Exam years	-0.080	0.079	-0.136**	-0.040	-0.120***	0.014	-0.020	0.028
	(0.051)	(0.063)	(0.060)	(0.078)	(0.037)	(0.042)	(0.036)	(0.046)
Magnet years		0.171*		0.075		0.076		-0.025
		(0.088)		(0.103)		(0.059)		(0.061)
Noble years		0.433***		0.345***		0.373***		0.125*
		(0.087)		(0.103)		(0.056)		(0.067)
p-value (exam = magnet)		0.23		0.20		0.24		0.35
p-value (exam = Noble)		0.00		0.00		0.00		0.09
N	6,372	6,372	6,516	6,516	6,954	6,954	6,582	6,582
Panel B: School-specific offers								
Exam years	-0.167***	-0.049	-0.064	0.004	-0.108***	-0.019	0.006	0.044
	(0.048)	(0.064)	(0.054)	(0.077)	(0.032)	(0.040)	(0.033)	(0.045)
Magnet years		0.119		0.036		0.032		-0.017
		(0.084)		(0.102)		(0.055)		(0.062)
Noble years		0.369***		0.372***		0.338***		0.169***
		(0.088)		(0.104)		(0.054)		(0.062)
p-value (exam = magnet)		0.02		0.69		0.26		0.25
p-value (exam = Noble)		0.00		0.00		0.00		0.02
F (exam)	33.1	20.2	33.2	20.8	33.3	22.6	34.3	22.9
F (magnet)		9.2		9.4		10.6		10.5
F (Noble)		11.3		10.6		11.6		11.0
N	6,694	6,694	6,728	6,728	7,171	7,171	6,967	6,967

Notes: This table reports 2SLS estimates of the effects of exam, magnet, and Noble years enrolled, estimated in multi-sector models. These estimates were computed using the sample of 2009-11 exam school applicants who applied to exam schools and a Noble or magnet school, as in Table 5. The instruments used to contruct the estimates in Panel A are described in the note to Table 5. The instrument list for Panel B replaces the any-exam-offer dummy with a set of school-specific offer dummies, adding appropriate controls for offer risk. The table also reports Angrist and Pischke first-stage F statistics for models with multiple endogenous variables and the p-value for tests of subgroup-effect equality. All models include the set of baseline covariates listed in Table A2 and dummies for risk sets for Noble and magnet applicants. Models in Panel A also include school-by-tier-by-year fixed effects and school-by-tier-by-year running variable controls. Models in Panel B also include year-by-tier fixed effects and year-by-tier specific running variable controls as described in the text. Robust standard errors appear in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7. Noble Effects Identified by Offer Instruments and Covariate Interactions

	PLAN		ACT	
	Math	Reading	Math	Reading
	(1)	(2)	(3)	(4)
Panel A: Offer instruments without covariates				
Noble lottery offer				
Noble years	0.412***	0.350***	0.344***	0.148**
	(0.071)	(0.089)	(0.047)	(0.057)
Exam offer (clears qualifying cutoff)				
Noble years	0.333*	0.500**	0.455***	0.118
	(0.179)	(0.209)	(0.113)	(0.128)
Both offers				
Noble years	0.386***	0.355***	0.361***	0.115**
	(0.070)	(0.084)	(0.045)	(0.054)
First stage F	57.47	58.36	58.79	52.69
Overid (DF=1) p-value	0.759	0.261	0.422	0.898
N	6,550	6,699	7,131	6,759
Panel B: Offer instruments with covariate interactions				
Noble lottery offer				
Noble years	0.352***	0.311***	0.309***	0.109**
	(0.074)	(0.083)	(0.048)	(0.055)
First stage F	11.52	11.44	12.09	9.67
Overid (DF=9) p-value	0.002	0.446	0.702	0.498
Exam offer (LIML)				
Noble years	0.379	0.451*	0.394***	0.160
	(0.239)	(0.246)	(0.094)	(0.111)
First stage F	4.04	4.35	4.89	4.57
Overid (DF=9) p-value	0.168	0.219	0.926	0.951
Both offers				
Noble years	0.358***	0.368***	0.350***	0.126**
	(0.063)	(0.075)	(0.040)	(0.048)
First stage F	8.36	8.38	8.55	7.42
Overid (DF=19) p-value	0.128	0.606	0.961	0.905
N	6,550	6,699	7,131	6,759

Notes: This table reports alternative IV estimates of the effects of Noble exposure for 2009-11 applicants. The sample includes applicants who applied to exam schools and either to magnet or Noble schools. In Panel A, the instrument list includes a dummy for the type of offer indicated in the heading. In Panel B, the list includes the type of offer plus interactions with covariates (lunch status, race, gender, baseline test scores, and tier dummies). All models control for the main effects of the baseline covariates listed in Table A2. Models using the exam offer include school-by-tier-by-year fixed effects and school-by-tier-by-year running variable controls, as described in the text. Models using the Noble offer also control for Noble risk sets. The table reports first-stage F-statistics and over-identification test p-values. Robust standard errors are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8. Exam School Effects on Charter-enrolled Applicants in a Large Urban District

Test Timing		Math	English
(1)		(2)	(3)
Within one year after enrollment		-0.212***	0.054
		(0.070)	(0.098)
	N	4,069	4,106
All available years		-0.161***	-0.021
		(0.051)	(0.057)
	N	5,499	5,534

Notes: This table reports 2SLS estimates of the effects of years of exam school exposure for 2003-15 applicants to exam schools in a large urban district. The applicants used to compute these estimates were enrolled in charter schools at the time of application. Charter-originating exam school applicants are likely to attend charters when not offered an exam school seat. Outcomes are test scores on a statewide assessment, standardized to the district mean and to have unit standard deviation. All students were tested within the end of the first year after enrollment; some were also tested in later grades. Standard errors clustered at the applicant level are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A1. Differential Attrition

	PLAN		ACT	
	Math	Reading	Math	Reading
	(1)	(2)	(3)	(4)
	Panel A: Any exam offer			
Clears qualifying cutoff	0.015	0.006	0.004	-0.000
	(0.010)	(0.010)	(0.010)	(0.010)
Mean	0.842	0.842	0.830	0.831
N	19,407	19,906	21,383	20,241
	Panel B: Magnet lottery offer			
Magnet lottery offer	0.036***	0.038***	0.036***	0.036***
	(0.013)	(0.013)	(0.013)	(0.013)
Mean	0.801	0.801	0.782	0.782
N	13,236	13,236	13,236	13,236
	Panel C: Noble lottery offer			
Noble lottery offer	0.014	0.014	0.022*	0.022*
	(0.011)	(0.011)	(0.011)	(0.011)
Mean	0.694	0.694	0.669	0.669
N	8,608	8,608	8,608	8,608

Notes: This table reports estimates of the effects of offer receipt on follow-up data availability for 2009-11 applicants. All models include the set of baseline covariates listed in Table A2. Models in Panel A also include school-by-tier-by-year fixed effects and school-by-tier-by-year running variable controls, as described in the text. Models in Panels B and C also control for magnet and Noble risk sets, respectively. Robust standard errors are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A2. Covariate Balance for Exam, Magnet, and Noble Network Research Designs

	Exam school RD (qualifying cutoff)		Exam school RD (school-specific offers)		Magnet lottery		Noble lottery	
	Mean	Offer effect	Joint F-test (DF=9)	p-value	Mean	Offer effect	Mean	Offer effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	0.375	-0.013	1.971	0.038	0.361	-0.000	0.373	-0.005
		(0.013)				(0.014)		(0.013)
Hispanic	0.436	0.027*	1.280	0.242	0.456	-0.000	0.568	0.007
		(0.014)				(0.016)		(0.014)
White	0.107	0.002	1.595	0.110	0.105	-0.002	0.024	-0.000
		(0.008)				(0.010)		(0.006)
Asian	0.067	-0.010	1.803	0.062	0.063	0.001	0.027	0.002
		(0.007)				(0.008)		(0.006)
Female	0.594	0.010	0.869	0.553	0.591	-0.015	0.549	0.008
		(0.015)				(0.018)		(0.017)
Free/reduced price lunch	0.794	0.022**	0.984	0.451	0.829	0.008	0.894	0.019*
		(0.011)				(0.013)		(0.011)
Baseline math	0.905	-0.025	0.698	0.711	0.530	-0.035	0.076	0.011
		(0.019)				(0.030)		(0.030)
Baseline reading	0.799	-0.053***	0.765	0.649	0.517	-0.028	0.110	0.005
		(0.017)				(0.027)		(0.027)
N	17,826		18,805		10,500		7,000	

Notes: This table reports estimates of offer effects on covariates for 2009-11 applicants. Columns 1 and 2 report any-exam offer effects; columns 3-4 show F statistics for school-specific offers; columns 5-6 show magnet offer effects; columns 7-8 show Noble offer effects. Sample sizes in columns 1 and 3 count the number of observatons in the bandwidth for ACT math, for which baseline scores are available. Sample sizes in columns 5 and 7 count the number of observations with baseline scores. Fixed effects and running variable controls are as in Table 2. Models in columns 5-6 and 7-8 control for magnet and Noble risk sets, respectively. Robust standard errors are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A3. Characteristics of Compliers and Affirmative-Action-Affected Applicants

	Exam Applicants	Exam Enrolled	AA Affected (Improved Offer)	AA Affected (More Selective)	Exam Compliers
	(1)	(2)	(3)	(4)	(5)
Black	0.424	0.356	0.319	0.288	0.398
Hispanic	0.418	0.338	0.571	0.560	0.464
White	0.097	0.194	0.05	0.065	0.114
Asian	0.048	0.091	0.045	0.066	0.014
Female	0.545	0.598	0.612	0.604	0.597
Free lunch	0.728	0.542	0.739	0.690	0.701
Reduced lunch	0.098	0.117	0.096	0.103	0.132
Special education	0.101	0.068	0.034	0.039	0.015
Tier 1	0.237	0.213	0.647	0.550	0.272
Tier 2	0.263	0.220	0.333	0.426	0.222
Tier 3	0.273	0.260	0.014	0.022	0.273
Tier 4	0.226	0.308	0.007	0.000	0.235
Baseline math	0.397	1.215	1.129	1.616	0.803
Baseline reading	0.408	1.410	1.351	1.342	0.969
N	45,513	7,795	1,548	917	13,852

Notes: This table reports summary statistics for exam applicants, exam school students enrolled in 9th grade, applicants for whom affirmative action matters, and complier characteristics for the exam school RD design. The AA-affected sample includes: (a) improved offer: applicants for whom affirmative action produces an offer from a school they've ranked above at which would otherwise have been offered a seat, (b) more selective: gets an offer from a more selective school (Jones, Northside, Payton, and Young) than otherwise. Negative estimates are reported as zeros.

A Theoretical Appendix: Risk in Serial Dictatorship

SD, a version of DA without priorities, works like this:

> Order applicants by tie-breaker. Proceeding in order, offer each applicant his or her most preferred school with seats remaining.

Like any DA-type mechanism, SD generates a set of cutoffs, denoted τ_s for school s. For any school s that ends up full, cutoff τ_s is given by the tie-breaker, normalized to the unit interval, of the last student offered a seat at s. Otherwise, $\tau_s = 1$. Finite-market cutoffs are typically random, that is, they depend on the distribution of lottery draws. In large "continuum" markets, however, cutoffs are constant, a result that motivates use of the continuum model.[25]

In SD, applicants seated at school s qualify there and are (necessarily) disqualified at schools they like better. The building blocks for SD risk at school s are therefore (a) the cutoff at s and (b) cutoffs at schools preferred to s. The latter are characterized by a quantity we call *most informative disqualification* (MID), which tells us how the tie-breaker distribution among type θ applicants to s is truncated by offers at schools θ prefers to s. Formally, let Θ_s denote the set of applicant types who list s and let

$$B_{\theta s} = \{s' \in S \mid s' \succ_\theta s\} \text{ for } \theta \in \Theta_s \tag{17}$$

denote the set of schools type θ prefers to s (The discussion here ignores tier, which can be absorbed into preference lists). For each type and school, $MID_{\theta s}$ is a function of cutoffs at schools in $B_{\theta s}$, specifically:

$$MID_{\theta s} \equiv \begin{cases} 0 & \text{if } B_{\theta s} = \emptyset \\ \max\{\tau_b \mid b \in B_{\theta s}\} & \text{otherwise.} \end{cases} \tag{18}$$

$MID_{\theta s}$ is zero when school s is listed first since all who list s first compete for a seat there. The second line reflects the fact that an applicant who lists s second is seated there only when disqualified at the school they've listed first, while applicants who list s third are seated there when disqualified at their first and second choices, and so on. Moreover, anyone who fails to clear cutoff τ_b is surely disqualified at schools with lower (less forgiving) cutoffs. For example, applicants who fail to qualify at a school with a cutoff of 0.5 are disqualified at schools with cutoffs below 0.5. We can therefore quantify the truncation induced by disqualification at schools preferred to s by

[25] Abdulkadiroğlu et al. (2017a) explores alternative justifications of the continuum model.

recording the most forgiving cutoff among them.

Type θ cannot be seated at s when $MID_{\theta s} > \tau_s$ because those qualified at s can do better (they qualify at the school that determines $MID_{\theta s}$). Assignment risk when $MID_{\theta s} \leq \tau_s$ is therefore the probability that

$$MID_{\theta s} < R_i \leq \tau_s.$$

We summarize these facts in the following proposition, which is derived in Abdulkadiroğlu et al. (2019).

Proposition 1 (Global Score in Serial Dictatorship): *Consider serial dictatorship in a continuum market, using a single tie-breaker, R_i, distributed according to $F_R(r|\theta)$ for type θ. For all s and $\theta \in \Theta_s$, we have:*

$$p_s(\theta) \equiv E[D_{is}|\theta_i = \theta] = \max\{0, F_R(\tau_s|\theta) - F_R(MID_{\theta s}|\theta)\}.$$

SD assignment risk, which is positive only when when the cutoff at s exceeds $MID_{\theta s}$, is given by the size of the group with R_i between $MID_{\theta s}$ and τ_s. This is

$$F_R(\tau_s|\theta) - F_R(MID_{\theta s}).$$

With lottery tie-breaking (and a uniformly distributed lottery number), the SD risk formula simplifies to $\tau_s - MID_{\theta s}$. With non-lottery tie-breaking, as for CPS exam schools, the SD propensity score depends on the conditional distribution function, $F_R(\cdot|\theta)$, evaluated at τ_s and $MID_{\theta s}$.

This result does not immediately resolve omitted variables bias. With non-random tie-breakers like test scores, conditional tie-breaker distributions, $F_R(.|\theta)$, are likely to depend on θ, so the score in Proposition 1 need not have coarser support than does θ. This is in spite of the fact many applicants with different values of θ share the same $MID_{\theta s}$. Second, $F_R(.|\theta)$ is typically unknown. Finally, while control for the propensity score eliminates confounding from type, assignments are a function of tie-breakers as well as type, and non-lottery tie-breakers are likely to be correlated with potential outcomes.

Abdulkadiroğlu et al. (2019) address these challenges by evaluating risk for applicants close to cutoffs. Proposition 1 identifies the relevant cutoffs in markets with many schools and types. Begin

by defining intervals around each cutoff. Given bandwidth δ, these intervals are defined by

$$t_{is}(\delta) = \begin{cases} n & \text{if } R_i > \tau_s + \delta \\ a & \text{if } R_i \leq \tau_s - \delta \\ c & \text{if } R_i \in (\tau_s - \delta, \tau_s + \delta]. \end{cases} \tag{19}$$

We collect the set of these for all schools in the vector

$$T_i(\delta) = [t_{i1}(\delta), ..., t_{is}(\delta), ..., t_{iS}(\delta)]'.$$

We then have:

PROPOSITION 2 (Local Score in Serial Dictatorship): *Consider serial dictatorship in a continuum market. Assume that cutoffs τ_s are distinct. For each $s \in S$ and $\theta \in \Theta_s$ such that $MID_{\theta s} \neq 0$, suppose $MID_{\theta s} = \tau_{s'}$ for $s' \neq s$. For $T = [t_1, ..., t_s, ..., t_S]' \in \{n, a, c\}^S$, all $\delta > 0$, and all w,*

$$P[D_i(s) = 1 | \theta_i = \theta, T_i(\delta) = T, W_i = w] = 0 \text{ if } \tau_{s'} > \tau_s.$$

Otherwise,

$$\lim_{\delta \to 0} P[D_i(s) = 1 | \theta_i = \theta, T_i(\delta) = T, W_i = w] = \begin{cases} 0 & \text{if } t_s = n \text{ or } t_{s'} = a \\ 1 & \text{if } t_s = a \text{ and } t_{s'} = n \\ 0.5 & \text{if } t_s = c \text{ or } t_{s'} = c. \end{cases}$$

When $MID_{\theta s} = 0$, risk is 0 when $t_s = n$; 1 when $t_s = a$; and 0.5 otherwise.

The Proposition establishes a key conditional independence result: limiting SD assignment risk depends only on tie-breaker proximity to the cutoff at s and to $MID_{\theta s}$; risk is otherwise unrelated to applicant characteristics. Type θ applicants with tie-breakers near either $MID_{\theta s}$ or the cutoff at s face risk of one-half. Applicants with $t_s = a$ and $t_{s'} = n$ have tie-breakers strictly between $MID_{\theta s}$ and τ_s, meaning they're disqualified at s' but qualified at s. Finally, applicants with $t_s = n$ or $t_{s'} = a$ cannot be seated at s, either because they're disqualified there or because they qualify at s'.

It remains to note that qualification at *any* exam school involves only a single cutoff, $q(\theta_i)$. Near this cutoff, qualification risk is 0.5. Below and above $q(\theta_i)$, risk is degenerate.

In the empirical (as opposed to theoretical) world, almost all applicants necessarily have tie-

breaker values that are strictly above or below any particular randomization cutoff. But applicants with tie-breakers close to either $MID_{\theta s}$ or the cutoff at s are special, because it is these applicants for whom qualification is (almost) randomly assigned.

B Data Appendix for Chicago

This document describes data processing used to construct analysis files. Chicago Public Schools (CPS) is the source of the application files for two school sectors: (i) Exam schools, and (ii) Magnet schools. The Noble Network provided us with (iii) the Noble application file. CPS is also the source of: (iv) the enrollment file which contains student demographic and school attendance information, and (v) the outcomes data files which include the Plan test score file, and the Prairie State Achievement Examination (PSAE) file. This appendix describes these data sets and the procedures used to construct the sample of the main empirical analyses.

B.1 Data Sources

Exam School Application File

The exam school application file contains a record for each student consisting of an application id number, CPS id number, name, gender, race, date of birth, the tier (from 2009 onwards), address, special education status, application year, preferences over nine exam schools, and the composite score for admission. Each record also includes the school where the student receives an offer (if any). This data set covers students with application years from 2002-2012. The analysis sample only includes applicants from 2009-2011. Students enroll in the fall of the following year. Cutoffs for this period are published in the CPS website: http://cps.edu/AccessAndEnrollment/Documents/SEHS_CutoffScores.pdf. We exclude duplicate observations, and applicants who were missing the application id number from the analysis.

Magnet School Application File

The magnet application file contains a record for all applications to magnet schools. The records consist of an application id, a CPS id number, the set of magnet schools the students apply to, first name, last name, date of birth, application year and offer status. Magnet schools use separate lotteries for applicants in proximity to the school, applicants with a sibling in the school and finally, separate lotteries for each tier. These different lottery groups are indicated in the application file. This data set covers students from 2001-2012. The analysis sample only includes applicants from 2009-2012.

Noble Network Application File

The Noble Network application file contains a record for each student consisting of the CPS ID number, gender, name, race, date of birth, application year, the set of noble schools the student applies to, the lottery number for oversubscribed schools, offer status for each Noble school, and whether it is a lottery offer or a sibling offer. This data set covers students with application years from 1998-2015. The analysis sample only includes applicants from 2009-2011.

CPS Enrollment File

The CPS enrollment file spans school years 2007-2008 through 2012-2013. Each record contains a start of the school-year (October) snapshot for each student enrolled in Chicago Public Schools, with unique student identifier (the CPS ID), the student's grade and school, and demographic information. The variables of interest in the enrollment file are grade, year, date of birth, sex, race, special education (SPED), limited English proficiency (LEP) status, disability status, subsidized lunch eligibility, and school. Students are coded as attending an exam school, a magnet school, or a Noble school if their enrollment in October is at each school sector, respectively. We transform the enrollment file into a wide-format layout for each student where we compute the grade and exam school years attended for a given year.

Outcome files

The Plan test score file spans school years from 2010 to 2014. It includes scores for four subjects: English, Math, Reading, and Science, the grade and the year in which the test was taken (mostly 10th grade), and the school of the student. We standardize scores among CPS test-takers by year.

The ACT test scores come from the Prairie State Achievement Examination (PSAE) files from 2010 to 2014. PSAE is a two-day standardized test taken by all High School Juniors in the U.S. state of Illinois through 2014. On the first day, students take the ACT. On the second day, students take a WorkKeys examination and Illinois State Board of Education-developed science examination. Students were evaluated in four subjects: Math, Reading, Science, and Writing. Like PLAN, the file has the grade and year of test. Scores are standardized among all CPS test-takers by year.

B.2 Merging Files

These five data sets were merged using the CPS ID. This implies that students who had not enrolled in CPS at the moment they applied either to Exam, Magnet, or Noble schools are dropped from the analysis. Panel A of Table B1 shows for each year and each school sector the proportion of applicants who are missing the CPS ID.

The master file corresponds to the enrollment data set of high school students from 2010 to 2013 for students.[26] This file is merged with the application files for each school sector. Panel B of Table B1 shows the application rates for each school sector in the period of analysis. We keep the students that applied either to an Exam, Magnet, or Noble school. We replace with zero the relevant variables for the analysis of a school sector for those students who did not apply to it.[27] All our estimations control for the set of school sectors to which a student applied in order to partial out potential differences. Table B2 shows the final sample in our analysis and the number of applicants for each school sector by year.

We then merge this data set with the outcome data: Plan and ACT Math and Reading test scores. Table B3 shows the merging rates between the two files for each outcome: the inverse of attrition rates. To generate the variable of years of exposure to each type of school we take the number of years enrolled in Exam, Magnet, or Noble schools between the year of application and the year in which the respective test was taken. This data is the analysis file.

[26] Since students enrolled in the fall of the following year to the year of application.

[27] For example: applicants to Noble who did not apply to an Exam school have a score of zero in the admission test to Exam schools. Similarly, applicants to Exam schools who did not apply to Noble have a lottery number of zero for the different Noble campuses that run a lottery.

Table B1. Merging Application and Enrollment Files

Year	Exam Schools	Magnet Schools	Noble Schools
Panel A: Fraction of Applicants missing CPS ID			
2009	11.4%	0.0%	-
2010	10.9%	0.0%	-
2011	9.4%	0.0%	-
Panel B: Application Rates for each type of School			
2009	39.8%	9.8%	6.3%
2010	46.9%	11.7%	8.3%
2011	45.4%	9.6%	9.4%

Table B2. Size of Analysis Sample

Year		Applied to:		
	N	Exam	Magnet	Noble
2009	15,351	14,456	3,523	2,500
2010	16,465	15,326	3,092	3,003
2011	17,279	15,936	5,046	3,671

C Data Appendix for the Anonymous Large Urban District

This appendix describes the data and the procedures used to construct the analysis files for the anonymous large urban district. The Public Schools Office is the source of the application files for exam schools. The State provided us with the enrollment files and statewide assessment data.

Exam School Application File

The exam school application file contains a record for each student consisting of an application id number, state id number, name, gender, race, date of birth, application year, grade of application,

Table B3. Attrition Rates

Year	Plan		ACT	
	Math	Reading	Math	Reading
2009	82.4%	82.4%	79.2%	79.2%
2010	81.1 %	80.1%	78.0%	78.0%
2011	77.3%	77.4%	75.9%	75.9%

preferences over the exam schools, and the composite score for admission. Each record also includes the school where the student receives an offer (if any). This data set covers students with application years from 1995-2017. The analysis sample includes applicants from 2004-2016 for which both baseline and outcome test scores are available. Students enroll in the fall of the same year. We exclude duplicate observations, and applicants who were missing the application id number from the analysis.

Enrollment File

The State enrollment file spans school years 2001-02 through 2017-18. Each record contains both a start of the school-year snapshot and an end of the school-year snapshot for each student enrolled in public schools. The variables include a unique student identifier, the student's grade, and school, and demographic information. The relevant demographics are grade, year, sex, race, low-income status, special education status (SPED), and native status. The end of the year enrollment file determines the school of the student before the application. The start of the year enrollment files determines the school of the student after the application. We compute the grade and exam school years attended for a given year.

Outcome files: Statewide Assessment Test

The test score file spans school years from 2002 to 2018. It includes scores for two subjects: English, and Math, the grade and the year in which the test was taken. We standardize scores among test-takers by year and grade.

C.1 Merging Files

These three data sets were merged using the State identifier. Students who could not be found in the State file for any year prior to application were dropped from the analysis. Column 1 of Table C1 shows for each year and each school sector the proportion of applicants who are missing this ID. The master file corresponds to the application file to exam schools. The using data sets are

- The enrollment file of high school students from 2004 to 2017. Column 2 of Table C1 shows the matching rate between the application and the enrollment files by year. On average, approximately 85% of the applicants are found in the enrollment file.

Table C1. Merging Application, Enrollment, and Outcomes Files

Year	Missing ID (1)	Match Enrollment (2)	Match Outcomes (3)
2004	10.9%	82.3%	92.4%
2005	10.4%	82.8%	93.7%
2006	7.2%	84.1%	95.8%
2007	8.2%	83.2%	95.3%
2008	5.4%	82.7%	94.3%
2009	5.6%	83.1%	95.0%
2010	4.8%	84.2%	95.0%
2011	4.9%	83.9%	95.1%
2012	3.7%	84.9%	95.9%
2013	3.3%	85.2%	96.0%
2014	3.1%	85.1%	96.2%
2015	5.0%	85.2%	94.1%
2016	5.3%	85.0%	95.7%
2017	3.3%	88.8%	93.5%

- The outcome file covers 2002 to 2018. Column 3 of Table C1 shows the matching rate between the application and the outcome files by year. On average, we have test scores for approximately 94% of the exam applicants.