Parents Know Better: Sorting on Match Effects in Primary School*

Marco Ovidi†

Università Cattolica Milano

October 2022

Abstract

I show that parents select schools by considering attributes of the student-school match that improve the learning outcomes beyond average school value-added. Using the centralized algorithm for primary school offers in London, I compare the achievement of primary school students in London who are as good as randomly assigned to the institution of choice. I find that only a small part of the match effects of parental choice can be explained by student's characteristics such as gender, ability, or socioeconomic status.

JEL Codes: H75, I21, I24, I28

Keywords: Centralised assignment, Deferred acceptance, School choice, School effectiveness

*I am especially grateful to Erich Battistin for careful support and guidance. I thank Marco Bertoni, Massimiliano Bratti, Davide Cipullo, Damon Clark, Lorenzo Cappellari, Tommaso Colussi, Francesca Cornaglia, Lucia Corno, Antonio Dalla Zuanna, Francesco Fasani, Emla Fitzsimons, Francois Gerard, Soledad Giardili, Ellen Greaves, Peter Hull, Attila Lindner, Marco Manacorda, José Montalbán Castilla, Lorenzo Neri, Barbara Petrongolo, Santiago Perez Vincent, Anna Raute, Camille Terrier, Olmo Silva, Laura van der Erve, Larissa Zierow and seminar participants at Queen Mary University, the Institute for Fiscal Studies, the LSE, Università Cattolica, the 6th SIdE WEEE Workshop, the 17th Brucchi Luchino workshop, the 22nd IZA Summer School in Labour Economics, the 10th IWAEE workshop, the 72nd ESEM, the 4th IZA workshop on Economics of Education, the II Padova Applied Economics Workshop, the 35th AIEL National Conference of Labour Economics, the 9th ICEEE, the 2021 RES Conference, the 25th SMYE, and the 2021 COMPIE conference for helpful discussions and comments. I acknowledge funding from Università Cattolica (D32 grant "EBAPP"). Any responsibility for the views expressed in the article rests solely with the author.

†Department of Economics and Finance,Università Cattolica, Largo Gemelli 1, 20123 Milano (Italy). Contact: marco.ovidi@unicatt.it.

1 Introduction

Although an increasing number of school districts around the world are expanding parental choice, the benefits of this expansion on educational productivity are strongly debated.[1] Should parents reward schools based on their causal impact on achievement (Rothstein, 2006), school choice may reallocate students to more effective schools and generate demand-side pressure on schools to improve their quality. However, empirical evidence suggests that parents respond mostly to indicators driven by student composition, such as test scores, which reflect peer quality rather than school quality (MacLeod and Urquiola, 2019).[2] Barseghyan et al. (2019) show that peer preferences weaken schools' incentives to improve.

A different channel through which school choice may improve student outcomes is the student-school match. In the presence of complementarities between student and school inputs, sorting across schools may depend on the specific educational needs of children, similarly to how workers sort into firms based on production complementarities (Lamadon et al., 2022). In the context of school choice, parents could select schools based on the "match effect" of attending the school of choice on student achievement beyond that school's value-added (VA, the average school causal impact across students).

Although match effects in education have recently received attention, available studies mainly investigate sorting based on student's characteristics such as gender, academic ability, or socioeconomic indicators. Bau (2022) find large effects of matching the instructional level to a student's socioeconomic status. Campos and Kearns (2022) find that a school choice expansion improves match quality, although positive achievement effects are driven by school quality improvements incentivised by competition for students. Abdulkadiroglu et al. (2020) show that match effects are negatively related to parental preferences conditional on peer quality and average school quality.

The novelty of my research is to document parents' sorting based on school match effects, over and beyond specific school effectiveness based on student's observable traits. I use administrative records on centralised assignment to primary schools in London to isolate

[1] Beyond England, which is the focus of this paper, choice among public-sector schools is allowed in many of the largest U.S. districts (Whitehurst, 2017) and urban districts such as Amsterdam (De Haan et al., 2021), Barcelona (Calsamiglia and Guell, 2018), Paris (Fack et al., 2019), and Beijing (He, 2017).

[2] One potential explanation is that school effectiveness on short-run test scores weakly correlates with that on longer-run outcomes (Jackson et al., 2021). In the UK, Gibbons et al. (2013) find that school value-added is capitalised on house prices in addition to peer quality.

quasi-experimental variation in admission to the most preferred institutions. I leverage data on previous cohorts of students to estimate a school's VA, and investigate whether attending the school of choice impacts student learning trajectories over and beyond the gain in school VA.

The ideal experiment would compare the learning outcomes of students randomly enrolled in otherwise identical schools in terms of VA, except for the preference rank assigned by parents. I take this idea to the data by leveraging the deferred acceptance mechanism (DA; Gale and Shapley, 1962) that matches students with school seats based on parental preferences and admission priorities. Seats at approximately 70% of London schools are rationed. In the case of excess demand, distance to school is used as a tie-breaker between applicants with equal priority, generating catchment boundaries that vary every year depending on the residence and preferences of all applicants. Uncertainty about the exact value of school catchments introduces variability in the admissions of students located close to the catchment boundary.

I use school offers to instrument enrolment conditional on assignment risk. I compare the achievement of students who, based on preferences, priorities, and distance, have equal chances of admission at the same schools of choice. DA maps assignment inputs into a scalar representing the assignment risk (Abdulkadiroglu et al., 2022). Same-risk applicants will have different offers if school seats are rationed, generating exogenous variation in the assignment. Consistent with this expectation, I show that applicants with an offer from the most preferred schools are statistically undistinguishable from applicants with the same assignment risk not receiving an offer from these schools.

I find that enrolling at the most preferred schools increases student learning over and beyond school VA. Comparing same-risk students by offer status conflates the match effect of most preferred schools and the gain in school VA. I contrast a student's actual achievement at a given school with the expected school VA on that student, which I estimate using administrative records from previous cohorts. The causal effect of enrolling at the first choice relative to a school ranked lower than second on Year 2 achievement in mathematics exceeds school VA by approximately 0.1 standard deviations (hereafter, σ). I find a similar effect size for the second-choice school. I show that only a small part of this match effect is explained by a student's characteristics considered in previous studies. I estimate match effects based on student's characteristics as the persistent effect of a school for a specific type

of student (Abdulkadiroglu et al., 2020), defined as the combination of covariates which are most predictive of academic achievement (gender, socioeconomic status, and baseline ability). Enroling at more-preferred schools generates a small increase in the match effect based on student's characteristics (about 0.02σ). The impact of accessing most-preferred institutions remains similar after accounting for this increase.

Sorting on match effects is more pronounced among parents from relatively advantaged backgrounds. Students not eligible for subsidised lunch, living in local areas with above-median education level, or with above-median achievement at school entrance exhibit larger effects. Parents with more resources may be in a position to make the most of school choice either because they are better informed on the suitability of different schools for their children or because they live in neighbourhoods with wider availability of high-quality options.

Several additional results suggest that average school quality is credibly accounted for in my comparison. First, VA gains from enrolling at the most preferred schools strongly predict causal effects on learning, in line with Angrist et al. (2016, 2017, 2021). I show that my results are unlikely driven by residual differences in school quality not captured by VA estimates. In addition, I document that positive match effects are found even among parents whose next-best school has higher average VA. Second, enroling at a more-preferred school is unrelated to school VA growth, implying that sorting into schools which are improving their quality over time does not explain my results. Third, I implement an alternative design to hold any school input constant. I compare students assigned to the same school who ranked that school with different preferences and find results similar to those obtained with my main research design.

I conclude that my results are driven by unobservable attributes that are specific to the student-school match. I show that parental rankings exhibit remarkable variability within the same school and narrow distance cells. Regardless of socioeconomic status, on average, parents rank schools by proximity to their residence and by peer quality. Access to high-performing schools is spatially segregated, with better-off parents nearly maximising both proximity and peer quality, while disadvantaged parents face steeper trade-offs. Despite the strong residential sorting, the fact that parents located at similar distance to a school often rank this school with different preferences is consistent with sorting patterns shown in my econometric analysis.

My results provide unique evidence of parents' sorting on match effects, implying that parental choice may increase allocative efficiency.[3] Differently from Campos and Kearns (2022), I study within-market effects of school choice in a context where parental choice is already a structural feature of the school market. My results may suggest that, after an initial market-level increase in aggregate school quality driven by competition among schools (Campos and Kearns, 2022), the benefits of school choice are displayed in the form of an improved match between students and schools.[4] Moreover, my results differ from findings by Abdulkadiroglu et al. (2020) in the context of New York City high schools. Substantial institutional differences may explain this contrast. First, primary schools in London are particularly small and assignment by distance implies relatively narrow choice sets, possibly easier to navigate for parents. Smaller and closer schools may improve the scope for parents to acquire information on their suitability for children's specific learning needs. Second, centralised assignment regulation leaves little scope for students' screening, and may therefore induce schools to compete on (vertical and horizontal) school quality instead.

A broad body of literature has investigated the impacts of attending the schools that parents prefer (e.g., Jackson, 2010; Pop-Eleches and Urquiola, 2013; Abdulkadiroglu et al., 2014; Deming et al., 2014; Hoekstra et al., 2018). In a meta-analysis, Beuermann and Jackson (2020) find a small and statistically insignificant effect on student achievement. In contrast to these works, I isolate student-school match effects from returns driven by average school quality. Methodologically, these studies consider admission cutoffs in isolation without fully exploiting school offer variation induced by centralised assignment (Abdulkadiroglu et al., 2022). Moreover, this study is the first to examine parental choice effects in primary schools. Since early stages are crucial for student development (e.g., Chetty et al., 2011; Heckman et al., 2013), this study fills an important gap in the literature.

Finally, many studies leverage data on submitted rankings to investigate parental preference for school attributes (e.g., Hastings et al., 2009; Burgess et al., 2015; Glazerman and Dotter, 2017; Burgess et al., 2019; Ainsworth et al., 2020). I describe parental preferences

[3] Kirkeboen et al. (2016) document positive returns to sorting into the field of study in higher education. In contrast, Kline and Walters (2016), Cornelissen et al. (2018), Walters (2018) find no or negative selection on gains in preschool programmes and high schools. Abdulkadiroglu et al. (2020) find positive selection on gains but very small in magnitude.

[4] This would not be inconsistent with the model in Campos and Kearns (2022) since aggregate school quality is shown to stabilise at a new, higher, equilibrium level, and competition effects cannot therefore be permanent.

accounting for the set of feasible schools, addressing recent concerns on the truthfulness of reported rankings under DA (Fack et al., 2019), and uniquely document substantial heterogeneity in parental rankings of the same school.

2 Institutional context

Primary education and school choice in London

Primary education in England spans seven grades, from age 5 to 11, and is organised into three phases. Students start with a reception year, which concludes the Early Year Foundation Stage (EYFS). During reception, students are assessed against several learning goals to inform teachers and parents of their readiness for Year 1. The second phase is Key Stage 1 (KS1), spanning two years. Teacher assessments in mathematics, science, and English are administered at the end of KS1. The final phase is Key Stage 2 (KS2), at the end of which students take externally marked standardised exams in mathematics and English. For all phases, the National Curriculum sets core knowledge and achievement objectives.

Parental choice among state-funded schools is established in England. Public-sector schools are the main provider of primary education, with less than 5% of students opting for private institutions.[5] Since the 1980s, the open enrolment policy has guaranteed parents the right of choosing a school for their children, as long as demand does not exceed capacity. Parents are required to rank up to six schools at the time of application. Local Authorities (LAs), the English school districts, assign a seat to all students at the highest-preference school available. School funding depends mostly on enrolment count, thus providing incentives to attract parental demand and fill capacity.[6]

Dissemination of data on school performance sparks competition for seats at high-performing schools. School Performance Tables, published annually since 1996, collect information on academic performance (both absolute achievement and achievement growth) and on intake composition of each state-funded school. Institutions with excellent test scores are typically sought after by parents, and they easily become oversubscribed (Burgess et al., 2015).

Admission criteria for oversubscribed schools have had an important impact on gentrifica-

[5] Author's own calculation from official 2019 data on students count by school phase and sector (link).
[6] Primary schools have a statutory class size cap of 30 students.

tion and urban development. When demand exceeds capacity, applicants are admitted mostly in order of proximity, generating fierce competition in the housing market to secure residence close to preferred institutions. The quality of surrounding schools is often mentioned in real estate advertising, and its impact on housing prices has been extensively documented by the economic literature (Machin, 2011; Gibbons et al., 2013; Battistin and Neri, 2017). The exact location of catchment boundaries, however, varies every year according to the supply of and the demand for school seats.

London is an ideal context to study school choice, with a dense supply of schools and high competition for popular institutions. Primary schools are typically small, enrolling approximately 50 students per cohort. Absolute achievement at KS2 outperforms the national average, driven by a dense right tail of schools serving exceptionally performing students. The fraction of students missing out on their top choices is systematically the highest countrywide (see Table I.2).

School assignment

School assignment is centrally regulated by the School Admissions Code. Applicants are admitted to their first choice as long as demand does not exceed capacity. Admission authorities must adopt and publish criteria to prioritize school applicants in case of oversubscription. National regulation leaves little discretion in setting priorities, explicitly banning a number of criteria such as selection by academic ability or interviews with parents and children. Few specific categories of students are prioritised and, within priority groups, distance to school is used as a tie-breaker. Schools are required to admit children with exceptionally disadvantaged backgrounds, constituting a very low share of students.[7] Moreover, applicants with siblings currently enrolled at the school are prioritised. Finally, exceptional admission criteria are permitted for religious schools, which typically set requirements based on faith.

School districts across England make offers through a deferred acceptance mechanism (DA, Gale and Shapley, 1962), matching students to the highest preference school with available seats. Since 2007, DA has been adopted nationwide for centralised school assignment

[7]The highest priority is given to children looked after by the LA, the approximately 0.5% of those under 18 years old in London in 2019 (official statistics here). In addition, priority is granted to children with a statement of special education needs (0.8% in my sample). The two groups are not mutually exclusive.

after the previously popular "Boston" mechanism was banned.[8] DA algorithms have proven less vulnerable to strategic preference reporting (Pathak and Sonmez, 2013). As long as parents act rationally, their rankings reveal the relative preference among listed institutions (i.e., the first choice is preferred to the second choice, the second to the third, etc., while no conclusion can be drawn regarding nonranked schools; see Fack et al., 2019).

Preferences, priorities, and school capacities are mapped into offers through the student-proposing DA algorithm. Each student is initially considered by their most preferred school. Applicants are ranked by priority and tie-breaker value and provisionally admitted up to capacity. In subsequent rounds, students who are rejected apply to their next-best choice and are ranked jointly with applicants provisionally admitted up to this point. Schools retain applicants up to capacity and rejects the rest, who in turn apply to their next-best choice. The algorithm stops when there is no more rejection. Some applicants may be left unassigned (3.2% in my sample, who are offered a nonranked school with spare capacity).

All parents in the country receive a single school offer in mid-April, deemed National Offer Day. Unsatisfied parents can join the waiting list at preferred schools with the same priority enjoyed in the centralised assignment. Although parents have the right to appeal in case of irregularities, the admission outcome is rarely overturned.[9]

3 Data

I exploit administrative data on applicants to state-funded primary schools in 2014 and 2015. Records include rank-order lists of schools submitted by parents to LAs and the school offered to each applicant by the assignment mechanism. Application data are matched to the National Pupil Database (NPD), including achievement records and socioeconomic characteristics of the universe of students in primary education. I observe the postcode of residence, granular information on residential location spanning an average of 15 properties and often corresponding to a single building in London. I compute the linear distance from

[8] This mechanism prioritises applicants ranking the school first, incentivising parents to nominate a "safe" first choice. Terrier et al. (2021) investigate the effects of the ban of the Boston mechanism on school choice outcomes.

[9] Among the 688 London primary schools with appeal data in 2015 (approximately 40% of the total), the 95% recorded no appeal resolved in parents' favour.

each applicant's postcode to all ranked schools.[10]

Assessments at the end of KS1 (Year 2) are the outcomes considered in my empirical investigations. Students are assessed by teachers at age 7, after three years of primary school. The results are grouped into three categories indicating whether students achieve below, at, or above expected standards. Three different subjects are assessed – English, separately for reading and writing, and mathematics.[11] Although teacher assessments are not standardised, detailed guidance is issued annually by the Government and external moderation is statutory, with LAs required to moderate a sample of at least 25% of schools (Department For Education, 2017). Students complete national tests in mathematics and reading and the end of KS1, with an optional writing test, which scores are not disclosed but are meant to inform teacher assessments. Burgess and Greaves (2013) find that approximately 80% of students are awarded the same achievement level in teacher assessments and standardised tests at the end of primary school (KS2), suggesting that teacher judgement is in line with externally-marked exams.[12]

To control for academic ability at entrance, I consider Early Years Foundation Stage Profile (EYFSP) teacher assessments. These evaluate 17 learning goals and are administered during the reception year. Similar to KS1 assessments, EYFSP results are grouped into three categories indicating whether students achieve below, at, or above expected standards. Moreover, I observe individual characteristics of students, including gender, free lunch eligibility, special education needs, language, and ethnicity group. At the local area (LSOA) level, income deprivation index (IDACI) measures the proportion of children in families that are considered deprived based on household income.

[10] Centroid coordinates for English postcodes are obtained from www.doogal.co.uk. For applicants without a postcode (approximately 3%), distance is imputed by exploiting the information on schools ranked by parents. I assign them the median distance among applicants ranking the same school with the same preference.

[11] The outcome variable I consider weights teacher assessments mirroring the scheme used by national authorities to compute average point scores for school accountability (9, 15, and 21 points for scoring below, at, or above the expected standards, respectively).

[12] Burgess and Greaves (2013) find that nonwhite students are more likely to receive teacher assessments below their standardised test score level. I control for ethnicity dummies in my analysis, which do not affect the results.

Sample description

I consider students who ranked at least one London primary school at application. The sample includes 199,220 applicants and 638,756 student-preference observations, with the average applicant ranking between 3 and 4 schools (Table I.2). Primary schools in London serve a mixed population, with 42% of students of white origin and 42% not speaking English at home (compared to 78% and 12% in the rest of England). Students in London are more likely to be eligible for free school meals than in the rest of the country. Despite the difficult context, primary schools in the capital outperform the average national institution in terms of academic achievement at Year 2, and the gap is wider than at primary school entrance.

Parents in London face striking competition for school seats. 17% of students miss out on the two most preferred institutions, 7 percentage points (p.p.) higher than in the rest of the country. Possibly anticipating higher competition, parents in London are 15 p.p. more likely to rank three schools or more at application.[13] A high proportion of parents (87%) complied with the centralised school offer, 3 p.p. lower than the national average, possibly reflecting higher propensity to opt for private schools (4% enrol at private schools against 2% in the rest of England).[14]

Replication of centralised school assignment

Centralised assignment breaking ties by distance implies that, when a school is oversubscribed, applicants located farther than a specific distance threshold are not admitted. This threshold, however, is not directly observed in the data since school offers also depend on parental preference and admission priorities (see Section 2). Panel A of Figure I.1 shows that although the probability of admission markedly decreases farther away from school, offer rates are not a deterministic function of distance. First, parents may rank the school differently conditional on distance, explaining why only 70% of applicants in the bottom decile receive an offer. Consistently, 35% of parents located closest to a school rank the school less than their first choice (Panel B). Second, particular categories of applicants are admitted with priority regardless of their location, partly explaining the non-negligible offer rate in the top decile of distance to school (approximately 0.2).

[13] In most English districts, the application form is restricted to three schools.
[14] A student is assumed to enrol in the private sector if not observed after application.

I replicate the assignment mechanism to trace catchment boundaries and identify applicants at the margin for admission.[15] The school catchment boundary is defined as the distance of the last admitted applicant. Replication is complicated by data availability since I have no information on demographics determining admission priorities – most importantly, whether a student has siblings at the school. Students with priority, however, are partially detectable in the data. First, I construct a proxy for siblings at the school based on postcode of residence and family-specific characteristics, and use this proxy to replicate school assignments. Second, I adjust offer replication to residual unobserved priorities.[16] Intuitively, if an applicant with offer resides beyond the initially estimated distance threshold, she must have priority. In Appendix A, I discuss how I achieve replication of school offers based on these ideas.[17]

My analysis rests on the assumption that the residual measurement error in admission priorities is not correlated with potential outcomes. Although school offers are perfectly replicated, some students with priority may remain undetected. This oversight would constitute a concern for my empirical strategy if students with unobserved priority were disproportionally located on either side of the catchment boundary and displayed markedly different potential outcomes. However, catchment boundaries are most likely unpredictable by parents, as I document below. Moreover, the validity of my design is supported by statistical balance in a number of characteristics associated with potential outcomes, including lagged achievement (see Section 4).

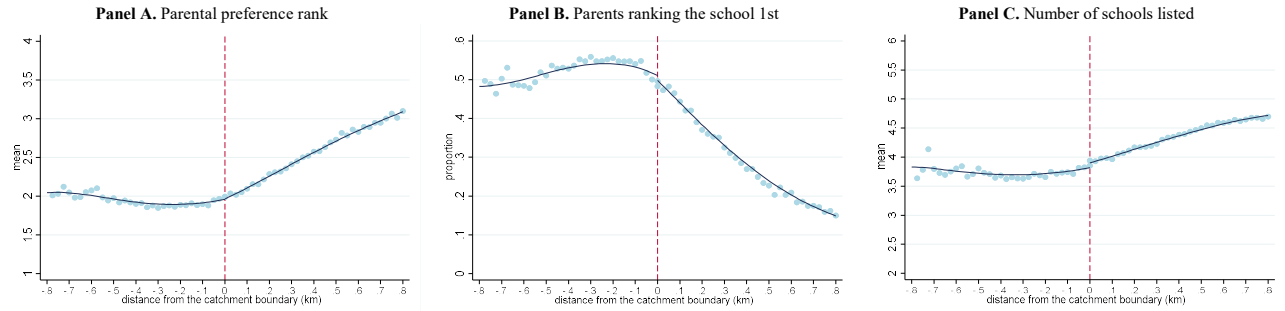
The distribution of parental rankings around the catchment boundary suggests that applicants are not able to anticipate the admission cutoff. Figure 1 shows that parental preference for the school (Panel A, a value of 1 denotes first choice), the share of parents ranking the school first (Panel B), and the number of ranked schools (Panel C) are continuous around the catchment boundary. As one might expect, the figure displays decreasing parental preference

[15] Other researchers have considered distance-based eligibility for policy interventions (Masi, 2018) or school admission (Gorman and Walker, 2021) to study school choice and its impacts. These studies, however, have not exploited the quasi-experimental variation arising from centralised assignment.

[16] These priorities are mainly religious criteria adopted by faith schools. I consider faith schools in offer replication but not in estimation as the measurement error in the catchment boundary is likely more serious.

[17] I replicate centralised assignment by running DA for all schools at the same time. In practice, the matching algorithm is run at the LA level and subsequently iterated up to 20 times to eliminate double offers across London LAs (Carter et al., 2020). The two procedures result in the same admission outcomes as long as 20 iterations are sufficient to sort all double admissions. In my algorithm, this number of iterations is always sufficient to eliminate double offers. Consistently, only one school is offered to each applicant.

Figure 1: Parental preference around the catchment boundary



Note. The figure plots parental preference rank (Panel A), the share of parents ranking the school first (Panel B), and the number of schools listed (Panel C) around the catchment boundary. Preference for the school varies from 1 to 6, indicating the first and sixth choice, respectively. Distance to school catchment boundary on the horizontal axis is defined by subtracting the distance of the last admitted candidate from an applicant's distance to school. Negative values indicate residence within the catchment. Markers represent average values in 25-metre-wide bins of distance from the boundary, and the solid line is a local linear fit of underlying observations, estimated separately on either side of the cutoff. A catchment boundary is defined for oversubscribed schools not admitting by faith. The sample is restricted to applicants within 800 metres from the catchment boundary and excludes last admitted applicants, who are used to define the school catchment. Reported values are averaged across the two cohorts considered. See Section 3 for details.

with distance to school (see Section 6). The decrease accelerates starting at a slightly shorter distance than the catchment boundary, suggesting that parents msy adjust their application behaviour based on their expectation about the cutoff realisation. However, the graph shows no discontinuity, suggesting that the exact location of the catchment boundary is, as expected, unpredictable by parents.[18]

4 Empirical framework

Parameter of interest

To fix ideas, I maintain the assumption that school offers are randomised and that compliance with these offers is perfect. I discuss below how these assumptions can be relaxed to fit the empirical context considered in this paper.

Following Abdulkadiroglu et al. (2020), student i's potential achievement at school s can be decomposed as:

$$Y_{is} = \nu_i + \alpha_s + \mu_{is}, \tag{1}$$

where ν_i is student ability, α_s is the school average causal effect (value-added, VA), and μ_{is} is the match effect for student i at school s. In a model where parents sort on their children's comparative advantage in the production of Y_{is} (Roy, 1951), μ_{is} is expected to increase with

[18]Figure A.4, which compares catchment boundaries across the two years considered, further reinforces this expectation. Although catchment boundaries are positively correlated over time, the figure shows idiosyncratic variation that is unlikely to be anticipated by parents.

the preference rank assigned by parents to the school.

Let r be the preference rank assigned to school s, with $r = \{1, \ldots, 6, \infty\}$ and $r = \infty$ denoting schools not ranked by the student. The school ranked r-th by student i is indexed by $s(i, r) = \{1, \ldots, S\}$. If students can receive offers only from listed institutions and Z_i denotes the preference rank for the school making an offer to student i, using equation (1) the observed outcome is:

$$Y_i = \sum_{z=1}^{6} \mathbb{1}(Z_i = z) \cdot Y_{is(i,z)} = \sum_{z=1}^{6} \mathbb{1}(Z_i = z) \cdot (\nu_i + \alpha_{s(i,z)} + \mu_{is(i,z)}). \tag{2}$$

Consider, for example, the comparison of students with offers from their first or second choice. Using equation (2), one can write:

$$E[\mu_{s(i,1)} - \mu_{s(i,2)}] = E[Y_i - \alpha_{s(i,1)} | Z_i = 1] - E[Y_i - \alpha_{s(i,2)} | Z_i = 2], \tag{3}$$

because offers are randomised. The last equation identifies the average match effect of attending the first choice relative to the second choice. The same reasoning extends to offers from schools ranked differently. The empirical analogue of the right-hand side of equation (3) requires knowledge of α_s. Following Deming et al. (2014), I estimate this term by constructing the average regression-adjusted test scores growth at school s (see Appendix B for details). By using the estimated value of α_s in equation (3), I rely on the assumption that this estimate is (nearly) unbiased on average. The empirical argument supporting this assumption is made in Angrist et al. (2016, 2017, 2021).

Offers from centralised assignment

School offers in London are not randomised. Under DA, school assignment depends on preferences submitted at the time of application, admission priorities, and distance to ranked schools. A simple comparison of students by offer status is likely biased as parents may choose residence and rank schools depending on the potential outcomes of their children. At the same time, the variables considered in admission are the only potential sources of self-selection of parents into the desired schools. As long as selection from these sources is controlled for, centralised school offers are independent of potential outcomes.

Conditional on assignment risk at school s, applicants are as good as randomly admitted to school s (Abdulkadiroglu et al., 2022). Specifically, for students with the same risk, the

only variation in school offers derives from the realisation of unpredictable admission cutoffs:

$$Z_i \perp Y_{is} \,|\, \boldsymbol{p_i}, \tag{4}$$

where the vector $\boldsymbol{p_i} \equiv [p_{i1}, \ldots, p_{iS}]'$ collects assignment risk at all schools, regardless of the preference rank. Because of (4), the parameter in equation (3) is identified by a weighted average of conditional versions of the quantities discussed above:

$$E[\mu_{is(i,1)} - \mu_{is(i,2)}] = \int \Big(E[Y_i - \alpha_{s(i,1)} | Z_i = 1, \boldsymbol{p_i}] - E[Y_i - \alpha_{s(i,2)} | Z_i = 2, \boldsymbol{p_i}] \Big) \, d\boldsymbol{p_i}. \tag{5}$$

The conditional independence in (4) follows from the tie-breaking embedded in centralised assignment. In the event of oversubscription, the assignment mechanism discriminates between applicants with equal preference and priorities by using distance to school. Tie-breaking generates year-specific catchment boundaries for each oversubscribed institution depending on the location, priorities and preferences of all applicants. Unless catchment boundaries are exactly anticipated by parents, centralised assignment generates uncertainty in admission outcomes conditional on the inputs considered by the matching mechanism.

The risks entering the left-hand-side term of equation (5) represent the probability of receiving an offer from the school ranked r-th conditional on assignment inputs:

$$p_{is(i,r)} \equiv P(Z_i = r | \boldsymbol{s_i}, \boldsymbol{\rho_i}, \boldsymbol{d_i}),$$

where the vector $\boldsymbol{s_i} = [s_{(i,1)}, \ldots, s_{(i,6)}]'$ collects the schools ranked by student i, and the vectors $\boldsymbol{\rho_i} = [\rho_{is(i,1)}, \ldots, \rho_{is(i,6)}]'$ and $\boldsymbol{d_i} = [d_{is(i,1)}, \ldots, d_{is(i,6)}]'$ denote student i's admission priorities and distance to ranked schools, respectively.[19] The probability $p_{is(i,r)}$ has two main building blocks. First, consider applicants with the same priority residing near the catchment boundary of school s who ranked school s at the top of their list. Intuitively, they face the same assignment risk, approximately 50% (Proposition 2 in Abdulkadiroglu et al., 2022). Second, the assignment risk of applicants ranking school s below the first choice depends on the probability of admission to schools ranked higher than school s. The estimation of assignment risk is detailed in Appendix C.[20]

[19] Assignment risk is scalar and coarsely distributed across applicants, addressing the empirical challenge of conditioning on the full set of assignment inputs. Preferences and priorities are too finely distributed for non-parametric type conditioning to be feasible.

[20] As expected, estimated assignment risk closely matches school offer. An application-level regression of the offer dummy on assignment risk estimates a slope coefficient of 0.995 ($\mathbb{R}^2 = 0.87$).

Imperfect compliance

I address non-compliance by using school offers as instrumental variables conditional on assignment risk. Identification through IV requires additional assumptions on parental choice behaviour. First, offers need to exert a strong impact on enrolment, a condition verified by first-stage results in Section 5 below.

Second, I must assume that receiving an offer from a marginally preferred school monotonically increases the preference rank for the school where students enrol. As formalised in Appendix D, this assumption implies that, for example, students with an offer from their second choice enrol at either their first or their second choice. Evidence in support of such monotonicity is provided in Section 5. In particular, Panel B of Figure 6 shows that only 1% of students enrol at a school ranked with lower preference than the one offered.[21]

Under these assumptions, I identify a local average treatment effect (LATE) of attending a school ranked with a higher preference. Consider the comparison between students enrolled in the first-choice school relative to a lower-ranked school, where D_{i1} and Z_{i1} indicate enrolment and offer at the first choice, respectively, and let $\tilde{Y}_i \equiv Y_i - \hat{\alpha}_{s(i)}$ be the VA-adjusted achievement of student i enrolled at school s. If C_i indicates school offer compliers, using Z_{i1} to instrument D_{i1} yields:

$$E[\mu_{is(i,1)} - \mu_{is(i,2)}|C_i = 1, \boldsymbol{p_i}] = \frac{E[\tilde{Y}_i|Z_{i1} = 1, \boldsymbol{p_i}] - E[\tilde{Y}_i|Z_{i1} = 0, \boldsymbol{p_i}]}{E[D_{i1}|Z_{i1} = 1, \boldsymbol{p_i}] - E[D_{i1}|Z_{i1} = 0, \boldsymbol{p_i}]} \quad , \tag{6}$$

where I assume for simplicity that students missing out on their first choice are offered their second choice.[22] I show in Appendix D that the average match effect among compliers can be obtained as a weighted average of the comparison in equation (6).

Estimation

I consider students with nondeterministic assignment risk at one or more of the listed schools (i.e., students displaying $p_{is(i,r)} < 1 \, \forall r$ and $\sum_r p_{is(i,r)} > 0$). Characteristics of the 22,268 applicants at risk of assignment resemble those in the full sample (Table C.1). The largest

[21] I consider student enrolment in the reception year. The interpretation of results with school mobility between reception and Year 2 (when achievement is observed) is discussed in Appendix G.

[22] Otherwise, equation (6) identifies a similar LATE where the counterfactual is a weighted average of potential outcomes at schools ranked second or lower (see equation 14 in Appendix D). Similar comparisons can be defined for schools ranked lower than the first choice.

Table 1: Balance tests

	Uncontrolled		Same-risk design		Within-school design	
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted outcome						
Mathematics						
Offered 1st choice	0.0989***	0.1146***	0.0233	0.0379	0.0205	0.0472**
	(0.0055)	(0.0065)	(0.0230)	(0.0250)	(0.0197)	(0.0216)
Offered 2nd choice	0.0561***	0.0733***	-0.0287	-0.0114	0.0205	0.0485**
	(0.0076)	(0.0083)	(0.0219)	(0.0247)	(0.0184)	(0.0207)
Offered 3rd choice		0.0669***		0.0471		0.0741***
		(0.0110)		(0.0287)		(0.0242)
Joint significance (p-value)	171.73 (0.000)	115.86 (0.000)	2.30 (0.100)	2.43 (0.063)	0.81 (0.443)	3.61 (0.013)
English						
Offered 1st choice	0.1026***	0.1377***	0.0339	0.0475*	0.0284	0.0577**
	(0.0060)	(0.0104)	(0.0251)	(0.0274)	(0.0216)	(0.0236)
Offered 2nd choice	0.0573***	0.0395***	-0.0260	-0.0098	0.0275	0.0582***
	(0.0082)	(0.0124)	(0.0238)	(0.0269)	(0.0200)	(0.0225)
Offered 3rd choice		0.0459***		0.0441		0.0813***
		(0.0153)		(0.0312)		(0.0263)
Joint significance (p-value)	159.03 (0.000)	106.98 (0.000)	2.51 (0.0816)	2.32 (0.073)	1.27 (0.280)	3.98 (0.008)
N	172,099	172,099	22,628	22,628	22,628	22,628
Tie-breaker controls			Y	Y	Y	Y
Assignment risk (ranked schools)					Y	Y
Assignment risk (all schools)			Y	Y		
School where enroled FEs					Y	Y

Note. The table shows estimates of covariate balance by offer status. It reports coefficients on dummies indicating offer from a student's first or second choice school (or third choice in columns 2, 4, 6) estimated from equation (7). Dependent variables are predicted KS1 score in mathematics and English obtained from a regression of student KS1 score on predetermined characteristics (gender, language, ethnicity, free school meal eligibility, special education needs, and baseline achievement), the latter regression has R2=0.35). The regressions in columns (1)-(2) consider all applicants and includes controls for cohort and dummies indicating first and second choice schools (and third choice in column 2). Columns (3)-(6) restrict the sample to students with nondetermistic assignment risk at one or more ranked schools. Columns (3)-(4) include a local linear polynomial of distance to the catchment boundary (the tie-breaker) at each ranked school, and dummies for assignment risk at all primary schools in London. Columns (3)-(4) include a local linear polynomial of distance to the catchment boundary (the tie-breaker) and assignment risk at each ranked school, and additionally control for student's school dummies.When the number of ranked schools is less than six, the corresponding control variables are set to 0, and dummies indicating missing preferences are included. Reported are, for each regression, F-tests of joint significance of coefficients on first and second choice offer dummies (and third choice in columns 2, 4, 6, p-values are reported in parentheses). Robust standard errors are reported in parentheses. See Section 4 for details. ***p<0.01. ** p<0.05. * p<0.1

differences include at-risk students being less likely of black origin (13.2% versus 16.6%) and residing in areas with higher levels of education (33.7% versus 36.5%). This mild selection is consistent with a nonzero chance of entering oversubscribed schools.

I start by testing covariates balance by offer status conditional on assignment risk. I estimate the following specification:

$$W_i = \sum_{z=1}^{2} \gamma_z \mathbb{1}(Z_i = z) + \boldsymbol{p_i} + f^d(\boldsymbol{d_i}) + u_{1i}, \tag{7}$$

where W_i is a baseline characteristic of student i. γ_1 and γ_2 represent the effect of being admitted to the first or second choice, respectively, vis-à-vis a school ranked lower than second choice or not ranked at all (approximately 9% of students are in the comparison group). I control for assignment risk by including dummies for student's chance of entering each school ($\boldsymbol{p_i}$, assignment risk is zero at nonranked schools). Following Abdulkadiroglu et al. (2022), $f^d(\boldsymbol{d_i})$ controls for tie-breaker value with a local linear polynomial of distance to the catchment boundary of each ranked school.[23] Since assignment risk is estimated, $f^d(\boldsymbol{d_i})$ adjusts for residual tie-breaker effects.

Same-risk students with and without school offer are observationally similar. I present a synthetic balance test by predicting student achievement based on predetermined characteristics (gender, language, ethnicity, special education needs, free school meal eligibility, baseline achievement, and local-area deprivation index), and by using predicted achievement as outcome in equation (7). Column (1), Panel A of Table 1 shows that students offered their first or second choice have about 0.1σ or 0.056σ higher predicted achievement in mathematics, respectively, than students offered their third choice or lower, with similar results for achievement in English. Controlling for assignment risk, this differences are substantially smaller and statistically undistinguishable from zero (see column 3). F-tests of the joint significance of γ_1 and γ_2 fail to reject the null hypothesis, in sharp contrast with the uncontrolled specification. A similar conclusion is reached when including offer from a student's third choice (columns 2 and 4). Balance tests on single student's covariates reported in Table I.1 suggest similar conclusions.

[23] Specifically, $f^d(d_i) = \sum_{s=1}^{6} b_{is(i,r)} * [d_{is(i,r)} + d_{is(i,r)} * \mathbb{1}(d_{is(i,r)} < \tau_{s(i,r)})]$, where $\tau_{s(i,r)}$ denotes the catchment boundary of school s and $b_{is(i,r)}$ indicates whether student i resides within a bandwidth around $\tau_{s(i,r)}$. See Appendix C for details on bandwidth selection. All controls are interacted with cohort dummies.

Match effects are estimated through the following first-stage and second-stage equations:

$$\mathbb{1}(D_i = r) = \sum_{z=1}^{2} \pi_{rz} \mathbb{1}(Z_i = z) + \boldsymbol{p_i} + f^d(\boldsymbol{d_i}) + u_{2i}, \qquad r = 1, 2, \tag{8}$$

$$\tilde{Y}_i = \sum_{r=1}^{2} \beta_r \mathbb{1}(D_i = r) + \boldsymbol{p_i} + f^d(\boldsymbol{d_i}) + u_{3i}. \tag{9}$$

where risk controls, $\boldsymbol{p_i}$ and f^d are defined analogously to equation (7). In equation (9), β_1 and β_2 estimate the impact of enrolling at the first or second choice on VA-adjusted achievement, respectively, relative to a school ranked third or lower.

5 Results

First stage

Receiving an offer from one of the most-preferred schools steeply increases the probability of enrolling at that school. Panel A of Table 2 reports first stage estimates of school offer coefficients from equations (8). Students offered their first choice versus a school ranked third or lower are 67 p.p. more likely to enrol at that school (Panel A, column 1). The corresponding estimate for the second choice is 77 p.p. Similar results are obtained from a model including offers from the second or third choice (Panel B).

First-stage results suggest that offer take-up responds to parental preference for the school. I find moderately smaller enrolment impacts of offer from the first choice than offer from the second choice (Panels B and C), likely reflecting the role of waiting lists at the most preferred institutions. This is visualised in Panel B of Figure A.3, showing larger enrolment rates among applicants located just beyond the catchment boundary of their first choice relative to lower-ranked schools. Moreover, all "off-diagonal" estimates in Table 2 (e.g., the impact of an offer from the second choice on first-choice enrolment) are negative, implying that an offer from a more-preferred school increases the likelihood of compliance with the centralised assignment. This is visualised in Panel A of Figure 6, plotting uncontrolled enrolment rates (blue bars). The 90% of students assigned their first choice comply with school offer, and this figure monotonically decreases when preference for the offered school decreases. Panel B shows that less than 1% of applicants enrol at a school with lower preference than the one assigned.

Table 2: First stage results

	Same-risk design			Within-school design		
	Enroled at 1st choice	Enroled at 2nd choice	Enroled at 3rd choice	Enroled at 1st choice	Enroled at 2nd choice	Enroled at 3rd choice
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A. Two-choice model			
Offered 1st choice	0.6723***	-0.0307***		0.5732***	-0.0233***	
	(0.0133)	(0.0084)		(0.0120)	(0.0085)	
Offered 2nd choice	-0.0823***	0.7711***		-0.1087***	0.7319***	
	(0.0112)	(0.0100)		(0.0094)	(0.0093)	
F-test	2133.78	3354.25		2007.59	3474.48	
			Panel B. Three-choice model			
Offered 1st choice	0.6467***	-0.0432***	-0.0204***	0.5288***	-0.0367***	-0.0117*
	(0.0149)	(0.0092)	(0.0064)	(0.0132)	(0.0098)	(0.0069)
Offered 2nd choice	-0.1126***	0.7562***	-0.0341***	-0.1553***	0.7178***	-0.0278***
	(0.0134)	(0.0111)	(0.0063)	(0.0112)	(0.0106)	(0.0058)
Offered 3rd choice	-0.0822***	-0.0404***	0.7198***	-0.1231***	-0.0372***	0.6897***
	(0.0157)	(0.0116)	(0.0149)	(0.0129)	(0.0104)	(0.0132)
F-test	1484.20	2265.46	948.79	1401.08	2361.57	1123.41
N	22,628	22,628	22,628	22,628	22,628	22,628
Tie-breaker controls	Y	Y	Y	Y	Y	Y
Assignment risk (ranked schools)				Y	Y	Y
Assignment risk (all schools)	Y	Y	Y			
School where enroled FEs				Y	Y	Y
Individual characteristics	Y	Y	Y	Y	Y	Y
Baseline achievement	Y	Y	Y	Y	Y	Y

Note. The table shows estimates of the impact of school offer on enrolment. It reports coefficients on school offer indicators from equations (8). Panel B reports estimates from specifications analogue to Panel A augmented by a third-choice offer dummy. When the number of listed schools is less than six, the corresponding control variables are set to 0, and dummies indicating missing preferences are included in the controls. All regressions control for language, ethnicity, FSM eligibility, SEN, gender, deprivation index and education in the area of residence, and baseline achievement. Reported are F-tests of joint significance of the two (Panel A) or three (Panel B) school offer indicators. Robust standard errors are reported in parentheses. See Section 5 for details. ***p<0.01. ** p<0.05. * p<0.1

Table 3: Match effects: VA-adjusted outcomes

	Mathematics				English			
	Same-risk design			Within- school design	Same-risk design			Within- school design
	KS1 score	VA-adjusted KS1 score	Match-adjusted KS1 score	KS1 score	KS1 score	VA-adjusted KS1 score	Match-adjusted KS1 score	KS1 score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A. Two-choice model				
Enroled in 1st choice	0.1197**	0.1073**	0.0985**	0.1260***	0.0566	0.0444	0.0361	0.0611
	(0.0479)	(0.0481)	(0.0482)	(0.0472)	(0.0454)	(0.0459)	(0.0462)	(0.0454)
Enroled in 2nd choice	0.1048**	0.0996**	0.0910**	0.0873**	0.0882**	0.0847**	0.0768*	0.0868**
	(0.0409)	(0.0409)	(0.0411)	(0.0380)	(0.0390)	(0.0392)	(0.0393)	(0.0361)
				Panel B. Three-choice model				
Enroled in 1st choice	0.1000*	0.0788	0.0694	0.1325**	0.0631	0.0407	0.0321	0.0814
	(0.0552)	(0.0555)	(0.0556)	(0.0582)	(0.0525)	(0.0531)	(0.0533)	(0.0556)
Enroled in 2nd choice	0.0826*	0.0677	0.0584	0.0937*	0.0955**	0.0805*	0.0723	0.1067**
	(0.0501)	(0.0502)	(0.0504)	(0.0486)	(0.0481)	(0.0483)	(0.0484)	(0.0460)
Enroled in 3rd choice	-0.0517	-0.0746	-0.0761	0.0138	0.0169	-0.0099	-0.0104	0.0429
	(0.0601)	(0.0606)	(0.0606)	(0.0554)	(0.0576)	(0.0578)	(0.0581)	(0.0526)
N	22628	22628	22628	22628	22628	22628	22628	22628
Tie-breaker controls	Y	Y	Y	Y	Y	Y	Y	Y
Assignment risk (all schools)	Y	Y	Y		Y	Y	Y	
Assignment risk (ranked schools)				Y				Y
School where enroled FEs				Y				Y
Individual characteristics	Y	Y	Y	Y	Y	Y	Y	Y
Baseline achievement	Y	Y	Y	Y	Y	Y	Y	Y

Note. The table shows estimates of match effects at the most preferred schools from equation (9). All regressions restrict the sample to students with nondetermistic assignment risk at one or more ranked schools. Dependent variables are KS1 score in mathematics (columns 1-4) or English (columns 5-8). Columns (2) and (6) report estimates of VA-adjusted KS1 score, obtained subtracting estimated school VA from student's score. Columns (3) and (7) report estimates of match-adjusted KS1 score, obtained subtracting estimated school match effects for a student's type from student's score. Controls in columns (1)-(3) and (5)-(7) include a local linear polynomial of distance to the catchment boundary (the tie-breaker) at each ranked school, and dummies for assignment risk at all primary schools in London. Columns 4 and 8 include a local linear polynomial of distance to the catchment boundary (the tie-breaker) and assignment risk at each ranked school, and additionally control for student's school dummies. Panel B reports estimates from specifications analogue to Panel A augmented by a third-choice offer dummy. When the number of ranked schools is less than six, the corresponding control variables are set to 0, and dummies indicating missing preferences are included in the controls. All regressions control for gender, language, ethnicity, free school meal eligibility, special education needs, deprivation index and education level in the neighbourhood of residence. Robust standard errors are reported in parentheses. See Section 5 for details on estimation, and Appendix B for details on estimated school VA and match effects. ***p<0.01. ** p<0.05. * p<0.1

Sorting on match effects

I start by documenting that enroling at a more-preferred school increases student's achievement. This can be seen in column 1, Panel A, of Table 3, reporting estimates of β_1 and β_2 from equation (9). Enroling at the first or second choice school increases achievement in mathematics by 0.12σ or 0.105σ, respectively. Estimates on student achievement in English are imprecise for the first choice (column 5). Including the third-choice school generates similar results, and the impact of enroling at one's third choice is not statistically distinguishable from zero (Panel B).

I next isolate the impact of enroling at a more-preferred school over and beyond average school quality by adjusting student's achievement for school value-added. I estimate a school's VA as regression-adjusted test scores growth at the school in previous cohorts (Deming et al., 2014, see Appendix B for details), and subtract estimated school VA from student's achievement. Column 2, Panel A of Table 3 shows that, when using VA-adjusted achievement

in mathematics as outcome in equation (9), the impact of first or second choice is only slightly smaller than unadjusted estimated (0.107σ or 0.1σ, respectively). I show in Table B.1 that estimated school VA strongly predicts student's achievement. Consistently with the results in Table 3, enroling at one's first choice against a school ranked third or lower slightly increases school VA (about 0.01σ, see Table B.2). Although one potential explanation of my results would be that parents sort into schools that are improving their quality over time, I show in columns (4)-(6) of Table B.1 that enroling at more-preferred schools has no impact on VA growth at a student's school.

Student's gender, ability, and socioeconomic status explain only a small part of the match effects I find. I group students in eight cells based on the predetermined characteristics that are most predictive of academic achievement, and estimate a school's persistent effects for each type of student following Abdulkadiroglu et al. (2020) (see Appendix B for details). Enroling at more-preferred schools increases school match effects based on student's characteristics by a slightly larger magnitude than average school quality (about 0.02σ, see Table B.2), implying a certain degree of parental sorting into schools based on these traits. After accounting for this increase, enrolment at one's first or second choice school improves student achievement in mathematics by 0.099σ or 0.091σ, respectively (column 3 of Table 3).

Enroling at more-preferred schools increases student's achievement even for parents who select institutions with lower average school quality with respect to their next-best school. I use deciles of estimated school match effect based on student's characteristics and define a student's match gain from enroling at her first or second choice as the difference in school match effect with respect to her second or third choice, respectively. Table 4 reports estimates from equation (9) augmented with interactions between erolment dummies and VA gain from enroling at one's first or second choice. Estimates of uninteracted enrolment dummies in column (1) imply an increase in student achievement in mathematics of about 0.1σ from enroling in the first or second choice, similar to my main results. In column (2), I include interactions with dummies for positive or negative VA gain. Estimates of the interaction between enrolment at one's first or second choice and negative VA gain are smaller than the uninteracted coefficient, implying that positive match effects are found even among parents which lose in terms of school quality from enroling at their most-preferred institutions. Results for English exhibit similar patterns (columns 3-4).

Table 4: Match effects: hold VA gains constant

	KS1 score in mathematics		KS1 score in English	
	(1)	(2)	(3)	(4)
Enroled in 1st choice	0.1171**	0.1417***	0.0540	0.0778
	(0.0479)	(0.0529)	(0.0454)	(0.0499)
Enroled in 2nd choice	0.1044**	0.0694	0.0893**	0.0820
	(0.0408)	(0.0682)	(0.0390)	(0.0649)
Enroled in 1st choice X	0.0048		0.0085	
school match gain	(0.0055)		(0.0053)	
Enroled in 2nd choice X	0.0053		0.0096	
school match gain	(0.0068)		(0.0067)	
Enroled in 1st choice X		-0.0107		-0.0017
1(school match gain > 0)		(0.0297)		(0.0279)
Enroled in 2nd choice X		0.0538		0.0290
1(school match gain > 0)		(0.0675)		(0.0640)
Enroled in 1st choice X		-0.0429		-0.0538*
1(school match gain < 0)		(0.0303)		(0.0281)
Enroled in 2nd choice X		0.0261		-0.0149
1(school match gain < 0)		(0.0677)		(0.0641)
N	22,628	22,628	22,628	22,628
Tie-breaker controls	Y	Y	Y	Y
Assignment risk (all schools)	Y	Y	Y	Y
Individual characteristics	Y	Y	Y	Y
Baseline achievement	Y	Y	Y	Y

Note. The table shows estimates of match effects at the most preferred schools. It reports estimates of school enrolment indicators from equation (9) augmented with itneractions between enrolment indicators and expected school match gain from entering most-preferred schools. Estimated specifications follow column 1 of Table 3. Match gains are computed as the difference between estimated school match effect decile of a student's first or second choice and her next-best alternative. Interactions in columns (2) and (4) consider dummies indicating positve or negative match gains. See Section 5 for details. Robust standard errors are reported in parentheses. ***p<0.01. ** p<0.05. * p<0.1

Alternative research design

I consider an alternative research design to reinforce the interpretation of my findings as match effects. My main design holds application choice constant, exploiting as-good-as-random variation in school offer. This design identifies the parameter of interest by exploiting the centralised assignment but needs to control for school VA to isolate match effects. One potential concern is that school quality is not fully controlled in my comparison. To eliminate this potential confounder, I compare students assigned to the same school, ranked by their parents with different preferences.

My alternative design holds school quality – and any school input – fixed, at the expense of a stronger identifying assumption. I exploit heterogeneity in application choices (see Section 6) and estimate empirical models similar to equation (9) augmented with a full set of dummies for the school where the student enrols.[24] A causal interpretation of the results requires the assumption that students who are as good as randomly admitted to *different* schools have similar potential outcomes. Although this analysis holds potentially important confounders constant, such as residential sorting with respect to listed schools, residual selection into application choice cannot be fully ruled out.

Balance tests are reassuring on the validity of my alternative design. Column (5) of Table 7 show that students enroled in the same school which they ranked with different preference have similar predicted achievement based on their predetermined characteristics. As expected, selection bias worsens when including the third choice in column (6), since these estimates compare students with increasingly different preference for the same school (i.e., ranking the same school first or second versus fourth or lower). Nonetheless, compared to uncontrolled estimates in column (2), selection bias is substantially reduced, especially for the first-choice offer.[25]

The results are qualitatively similar to the main specifications, suggesting that differences in school quality are not driving my results. Fore example, students enrolled at their first or second choice exhibit $0.126 - 0.087\sigma$ higher achievement in mathematics than students enrolled at the same school who ranked that school as their third choice or lower (see column

[24]Specifically, I consider students at risk of assignment and control for assignment risk at ranked schools. In contrast to the main empirical model, I do not control here for assignment risk at each school.

[25]In addition, columns (7)-(9) of Table 2 show that this design maintains a strong first stage.

Table 5: Heterogeneous match effects

	Match-adjusted KS1 score in mathematics										
	All	Free school meal eligible		White		Speaking English at home		Education in local area		Baseline achievement	
		Yes	No	Yes	No	Yes	No	Above median	Below median	Above median	Below median
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
						Panel A. Two-choice model					
Enroled in 1st choice	0.0985**	-0.0345	0.1086**	0.1537*	0.0351	0.0166	0.1855***	0.1641**	-0.0357	0.1913***	0.0558
	(0.0482)	(0.1805)	(0.0531)	(0.0841)	(0.0646)	(0.0856)	(0.0681)	(0.0674)	(0.0769)	(0.0721)	(0.0766)
Enroled in 2nd choice	0.0910**	0.1259	0.1118**	-0.0128	0.1544***	0.1740**	0.0526	0.1015*	0.0334	0.0997	0.0794
	(0.0411)	(0.1663)	(0.0451)	(0.0716)	(0.0568)	(0.0715)	(0.0584)	(0.0570)	(0.0653)	(0.0626)	(0.0655)
						Panel B. Three-choice model					
Enroled in 1st choice	0.1225*	-0.0562	0.0959	0.0888	0.0286	-0.0659	0.1807**	0.1425*	-0.0876	0.1445*	0.0274
	(0.0677)	(0.2049)	(0.0616)	(0.0993)	(0.0739)	(0.0985)	(0.0784)	(0.0805)	(0.0841)	(0.0842)	(0.0867)
Enroled in 2nd choice	0.0505	0.0964	0.0977*	-0.0799	0.1470**	0.0784	0.0473	0.0775	-0.0260	0.0466	0.0470
	(0.0622)	(0.2043)	(0.0556)	(0.0886)	(0.0691)	(0.0883)	(0.0714)	(0.0730)	(0.0761)	(0.0786)	(0.0780)
Enroled in 3rd choice	-0.0248	-0.0686	-0.0331	-0.1552	-0.0184	-0.2242**	-0.0126	-0.0515	-0.1617*	-0.1236	-0.0818
	(0.0740)	(0.2417)	(0.0675)	(0.1067)	(0.0839)	(0.1040)	(0.0871)	(0.0838)	(0.0983)	(0.0944)	(0.0969)
N	22,628	3,265	19,363	9,094	13,354	9,142	13,482	13,443	9,185	10,764	11,864
Tie-breaker controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Assignment risk (all schools)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Individual characteristics	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Baseline achievement	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note. The table shows estimates of match effects at the most preferred schools separately by student characteristics. Column (1) reports estimates in column (3) of Table 3. All other columns report estimates from similar specifications on different subsamples: student eligible or not eligible for free school meals (columns 2-3), students of white or other ethnic origin (columns 4-5), students speaking English or any other language at home (columns 6-7), students residing in local areas with above- or below-median deprivation (columns 8-9), or students with Year 0 achievement score above or below median (columns 10-11). Robust standard errors are reported in parentheses. See Section 5 for details. ***p<0.01. ** p<0.05. * p<0.1

4, Panel A of Table 3).

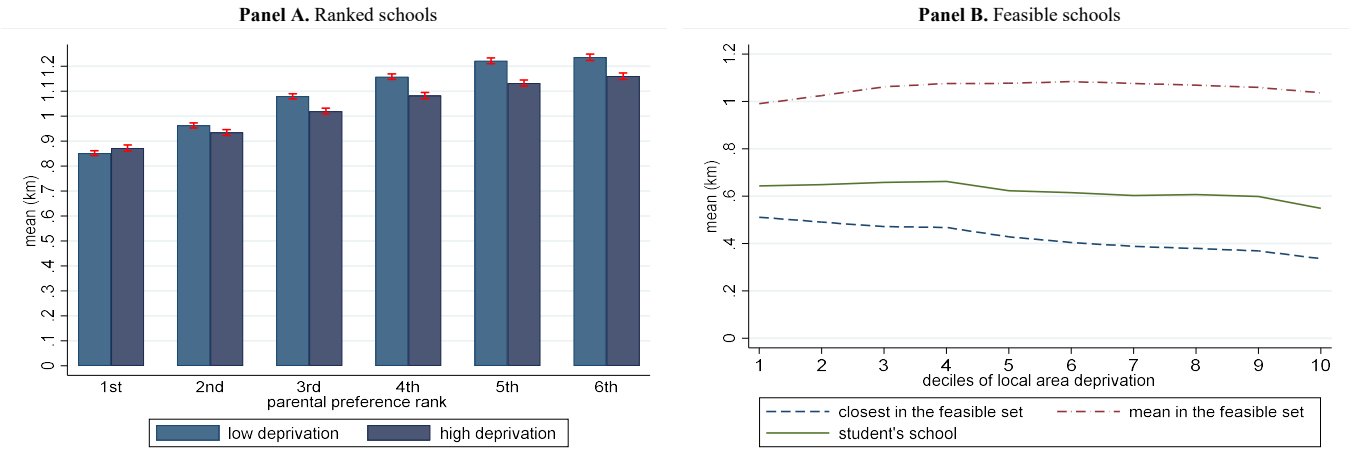
Overall, using different approaches and specifications, the results suggest that students enrolling at schools ranked by their parents with higher preference experience achievement gains above and beyond the impacts predicted by school VA. The findings point to the presence of heterogeneity in school effectiveness, and suggest that parents sort into schools that specifically enhance their children's achievement.

Heterogeneity analysis

I find suggestive evidence that sorting on match effects is more pronounced among parents of relatively advantaged socioeconomic backgrounds. Separate estimates by individual characteristics are presented in Table 5. Results are obtained from specifications similar to column (3), Panel A of Table 3, and full-sample estimates are reported in column (1) for ease of comparison.

Estimates are larger than average, at least at the first choice, for students not eligible for free school meals, residing in high-education areas, and with above-median baseline

Figure 2: Parental preferences and distance to school



Note. The figure plots parental preference rank against distance to school (Panel A) and distance to the school where the student enrols compared to average or lowest distance in the feasible choice set (Panel B). Bars in Panel A plot predicted values from equation (10), where preference rank dummies are interacted with an indicator variable equal to one if the deprivation index in the LSOA of residence is above the median. Controls include dummies for quintiles of peer quality and school value added. Superimposed in red are 95% confidence intervals of predicted values. In all panels, distance to school is measured in kilometres and computed as linear distance between student postcode and school postcode centroids. Schools farther than 2 kilometres from residence (the 90th percentile) are not considered. The deprivation index is based on average income in the LSOA of residence. See Section 6 for details.

achievement (columns 3, 8, and 10, respectively). All these characteristics may reflect higher socioeconomic status. In contrast, I cannot reject null effects among students residing in high-deprivation areas, with below-median baseline achievement or eligible for free school meals, and point estimates are even negative in the latter sub-group. Moreover, estimated match effects of one's first choice are larger than average for white students (column 4), and students speaking English at home (column 6), and I cannot reject null effects among students with other ethnic or linguistic backgrounds. Notably, however, most estimates by subgroup are not statistically different from each other.

Overall, the results suggest that parents with more resources may be in a position to better exploit school choice and improve match quality for their children.

6 Parental preference for schools

The results presented in Section 5 imply that parents select schools based on match effects, i.e., schools' specific effectiveness for their own children. In this section, I describe parental preferences for school attributes and discuss potential links with my results.

Parental preferences and school attributes

I start by using submitted rankings to describe parental preferences for geographical proximity, peer quality and school effectiveness. First, I plot the average attributes of listed institutions by parental preference rank, conditional on feasibility and number of schools listed. Specifically, I plot estimates of λ_1 to λ_6 from the following application-level regression:

$$A_{is(i,r)} = \sum_{p=1}^{6} \lambda_r \mathbb{1}(r = p) + X'_{is(i,r)}\zeta + u_{is}, \tag{10}$$

where $A_{is(i,r)}$ is an attribute of school s ranked r-th by student i. The parameter λ_r estimates the average level of attributes at schools ranked as r-th choice conditional on controls.[26]

Second, following Ainsworth et al. (2020), I compare the school where student enrols with other feasible institutions. This comparison complements the description of parental preferences by considering the supply side and provides insights into which attributes parents maximise. I define the individual feasible set as the collection of schools to which a student may or may not have applied that would have been accessible based on distance (see Appendix F).[27] The feasible set of the median applicant includes 6 schools within 2 kilometres (km) from their residence and 75% of parents could potentially access at least 4 schools.

Parents generally rank schools in order of proximity.[28] Panel A of Figure 2 plots distance to school by parental preference rank from equation (10), separately for students residing in areas with deprivation above or below the median.[29] The estimated distance to the first-choice school in better-off areas is approximately 850 metres, and all ranked schools are located within 1.2 km from residence. Similar distance to the first choice in high-deprivation neighbourhoods suggests that parents with lower socioeconomic status have a similar willingness to travel.

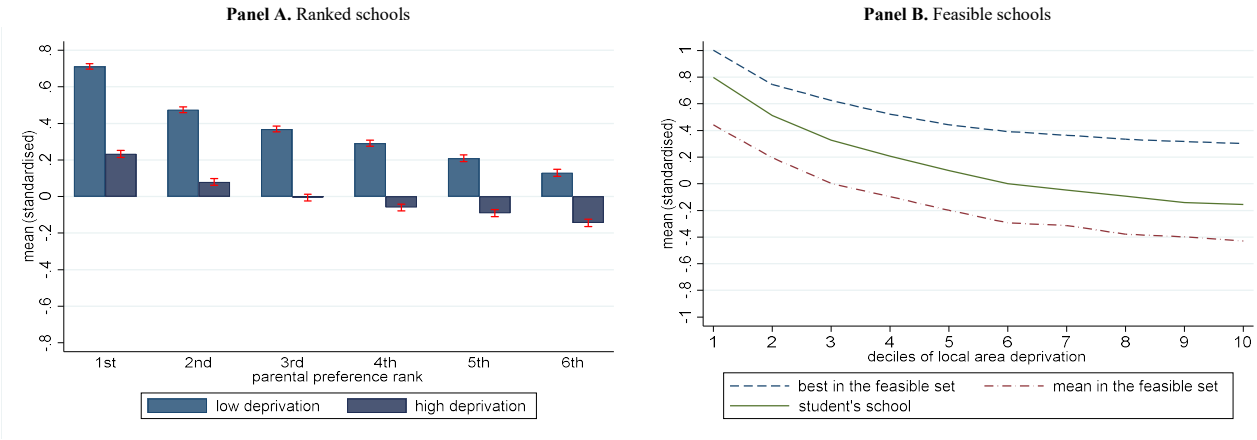
[26] In addition to dummies for number of schools listed and ex-post feasibility of the school, the vector $X'_{is(i,r)}$ includes attributes other than $A_{is(i,r)}$ (e.g., school VA and distance when considering peer quality).

[27] A similar idea is implemented in Burgess et al. (2015). School feasibility here is more precisely measured since it is based on school-specific boundaries obtained by replicating the centralised assignment.

[28] In this section, I consider applicants residing within 2 km from the school (the 90$^{\text{th}}$ percentile).

[29] Hastings et al. (2009) find different school preferences by socioeconomic status. Heterogeneity is explored here by interacting preference rank dummies in equation (10) with an indicator for above-median deprivation in the area of residence.

Figure 3: Parental preferences and peer quality



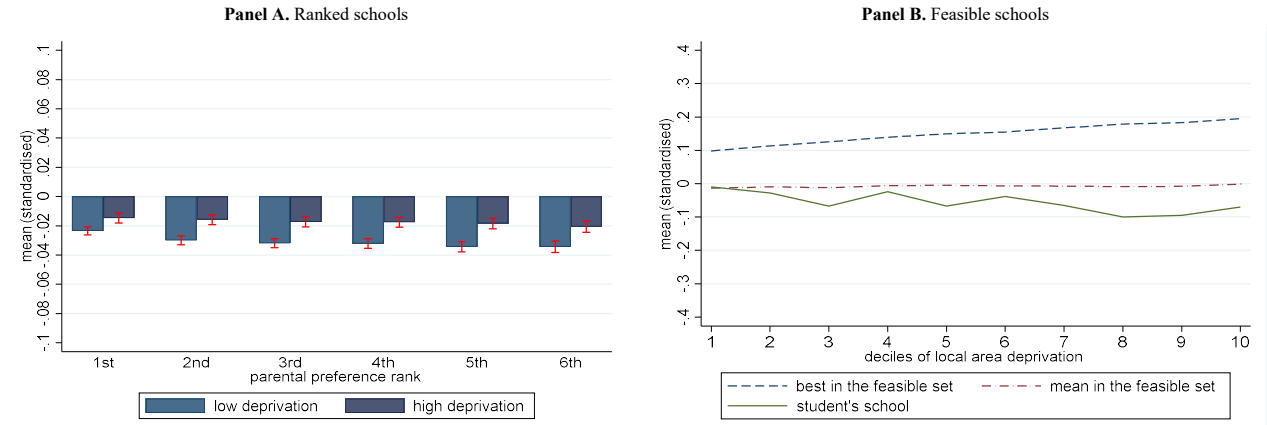
Note. The figure plots parental preference rank against peer quality (Panel A) and peer quality at the school where the student enrols compared to average or highest peer quality in the feasible choice set (Panel B). Bars in Panel A plot predicted values from equation (10), where preference rank dummies are interacted with an indicator variable equal to one if the deprivation index in the LSOA of residence is above the median. Controls include dummies for quintiles of school value added and distance to school. Superimposed in red are 95% confidence intervals of predicted values. In all panels, peer quality is measured by school-level final year test scores, averaged across 2007-2016 cohorts and across mathematics and English. Test scores are standardised to have zero mean and unit variance at the school level. The deprivation index is based on average income in the LSOA of residence. See Section 6 for details.

Parents nearly minimise the distance among available options. Panel B of Figure 2 compares distance to the school where the student enrols against the closest feasible institution and the average feasible distance by decile of local area deprivation (Ainsworth et al., 2020). Students with different socioeconomic statuses travel very similar distances to primary school, approximately 600 metres.[30] Parents in areas with higher deprivation face schools closer to the place of residence, likely reflecting higher population density. The 51% of applicants enrol at the closest accessible institution (see columns 8-10 of Table I.3). On average, parents "leave on the table" schools closer to the place of residence by 220 metres, possibly trading off proximity with other valued attributes. Interestingly, this figure is lower for applicants in low-deprivation areas who likely choose their residence close to desired schools.

On average, parents rank schools by peer quality, and access to high-performing institutions is spatially segregated. Panel A of Figure 3 plots the standardised final test scores at the school by parental preference rank from equation (10).[31] Estimated peer quality at the first choice in low-deprivation areas is 0.7σ above the average, decreasing to 0.2σ in worse-off neighbourhoods (similar to peer quality at the fifth choice in low-deprivation areas). Regardless of socioeconomic status, peer quality markedly decreases with preference rank. Most of the gap in peer quality by local deprivation is explained by the segregation

[30] This value differs from the distance values in Panel A since the latter are regression-adjusted estimates.
[31] I measure peer quality using KS2 test scores from 2006 to 2016 averaged across mathematics and reading.

Figure 4: Parental preferences and school VA



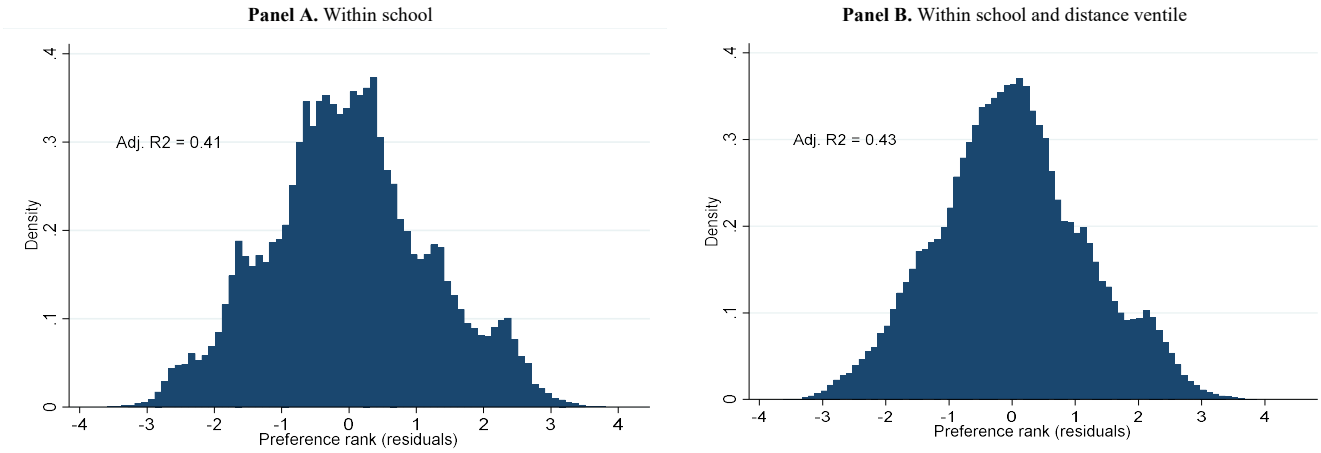
Note. The figure plots parental preference rank against estimated school value-added (Panel A) and value-added at the school where the student enrols compared to average or highest value-added in the feasible choice set (Panel B). Bars in Panel A plot predicted values from equation (10), where preference rank dummies are interacted with an indicator variable equal to one if the deprivation index in the LSOA of residence is above the median. Controls include dummies for quintiles of peer quality and distance to school. Superimposed in red are 95% confidence intervals of predicted values. In all panels, valu- added is estimated by regression-adjusted test scores growth at the school and averaged across subjects (see Appendix B). The deprivation index is based on average income in the LSOA of residence. See Section 6 for details.

of high-performing schools. Parents residing in areas with deprivation in the bottom decile face 0.8σ higher feasible peer quality than parents in the top decile, and the gap for schools where students enrol is approximately 1σ (see Panel B). Parents in areas with below-median deprivation leave little "on the table" in terms of peer quality, 0.27σ, compared to 0.43σ in areas with above-median deprivation (columns 5-7 of Table I.3). This gap may reflect steeper trade-offs between school test scores and distance for disadvantaged parents.

Conditional on distance and peer quality, parents do not respond to school VA. Panel A of Figure 4 shows nearly no association between school VA and parental preference rank, in line with findings in other contexts (MacLeod and Urquiola, 2019). Panel B shows that regardless of socioeconomic status, students enrol at schools with lower VA than the average feasible institution. Students could potentially access schools with 0.4σ larger VA (columns 2-4 of Table I.3), implying that achievement could substantially increase under alternative allocations if returns to school were homogeneous. However, the findings presented in Section 5 suggest that school effects are heterogeneous, offering a potential explanation for the weak correlation between parental preferences and average school VA.

Parental preferences lead to rationing of seats at schools with high peer quality. Attributes of oversubscribed schools are described in Table I.4. I define a school as oversubscribed if the number of applicants who are not offered a place at a more preferred institutions

Figure 5: Variability in parental preferences



Note. The figure plots residual parental preference rank after controlling for dummies indicating the number of schools listed, ex-post feasibility of the school, and school dummies (Panel A); or additionally including dummies for distance to school ventile (Panel B). The dependent variable is parental preference rank, ranging from 1 (first choice) to 6 (sixth choice). The adjusted R-squared index of the regression is reported in the top left. See Section 6 for details.

exceeds capacity. School seats are rationed in 58% of institutions.[32] Mirroring evidence from preference data, oversubscribed schools have markedly higher peer quality, 0.65σ in mathematics and 0.77σ in English, while the difference in VA is substantially narrower (0.16σ and 0.2σ, respectively).

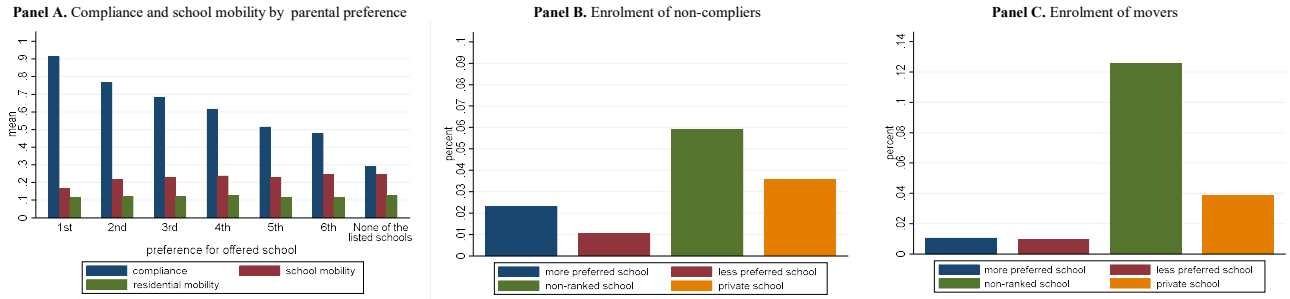
Heterogeneous parental preferences

Average patterns mask substantial disagreement in parental ranking of schools. The within-school dispersion is plotted in Figure 5, reporting the distribution of residual preference ranks from a regression on school dummies, ex-post feasibility, and number of ranked schools. Panel A shows that residuals exhibit remarkable variability, with an adjusted R-squared index of just 0.41. A substantial share of parents assigns preference ranks one or two positions away from the average for the same school. Including dummies for distance-to-school ventiles only slightly increases the adjusted R-squared (0.43, Panel B).[33] Therefore, even conditional on distance, a substantial share of variation in parental preferences is explained by unobservable attributes that are specific to the student-school match.

[32] I consider schools oversubscribed by at least 5 seats. The fraction of schools oversubscribed by one seat or more is 66%.

[33] Similar results are obtained when also considering nonranked feasible schools (available upon request). In this case, controlling for distance to school generates a larger increase in explained variability ($\mathbb{R}^2 = 0.5$), suggesting that distance is strongly associated with the decision of ranking a school.

Figure 6: Compliance with assignment and mobility



Note. The figure plots compliance with school offer and school and residential mobility by parental preference. Panel A plots compliance, school mobility and residential mobility rates by parental rank for school offered. Panel B plots the share of students who do not comply with school offer by preference for the school where they enrol in the reception year. Panel C plots the share of students who change schools from the reception year by preference for the school where they enrol in Year 2. Residential mobility is defined as changing home postcode from the previous academic year. See Section 6 for details.

Nevertheless, subsequent parental choice behaviour is consistent with preferences submitted at the time of application. The likelihood of moving children to a different institution by Year 2 decreases with the preference for the offered school (Panel A of Figure 6, red bars). Residential mobility, in contrast, is approximately orthogonal to school assignment (green bars).[34] Panel C shows that only 1% of students move to a school ranked with a lower preference, suggesting that parents consistently avoid less preferred schools even after initial enrolment. A similar fraction of students move to a school with higher parental preference, suggesting that centralised assignment is successfully enforced. In addition, I show in Appendix G that the decision to move to a different school responds to peer quality but not to school VA, in line with preferences submitted at the time of application.

In conclusion, although parents generally prefer schools that have higher peer quality and that are closer to the place of residence, their rankings exhibit substantial variability. Nevertheless, preferences expressed at the time of application predict school choice behaviour in subsequent years. Therefore, parental preferences appear to reflect solid tastes for available schools, and yet to a large extent, they are student-specific. The findings presented in Section 5 suggest that such heterogeneity is linked to unobservable school attributes that increase student achievement.

[34] I define residential mobility as an indicator variable equal to one if a student's home postcode changes from Year 0 to Year 2. This result suggests that parents who are willing to change their residence to secure a school place do so before applying.

7 Summary and conclusion

Expanding parental choice can be viewed as a zero-sum game if school quality is homogeneous across students or if parental choice is unrelated to student achievement. On the other hand, it may increase system-wide school productivity if students have specific educational needs and parents select schools on this basis. I have investigated this hypothesis in the context of primary school choice in London, a dense urban area featuring fierce competition for seats at popular schools. Identification is challenged by the possibility for parents to increase the chance of their children enrolling at desired schools through residential sorting. I show that centralised assignment breaking ties by distance can be used to isolate exogenous variation in admission, building on methods proposed by Abdulkadiroglu et al. (2022).

I show that the impact of enrolling at the most preferred schools exceeds the effect predicted by average school VA. The findings imply that returns to school are heterogeneous, and that parents leverage this heterogeneity to improve the quality of the student-school match. Consistently, I document substantial heterogeneity in parental preference for a given school, even conditional on residential distance. Since the information on specific educational needs of students is likely private, expanding parental choice may improve the efficiency of school seat allocation.

My results are in line with surveys of parents documenting that a school's local reputation or the particular needs of children are more important to them than distance or peer quality (Francis and Hutchings 2013; Montacute and Cullinane 2018). Moreover, parents with relatively advantaged backgrounds are more likely to report using multiple information sources, consistent with the results of my heterogeneity analysis. A potentially important source of match effects, which could not be assessed with the data at hand, is allocation to specific teachers. Ahn et al. (2021) find that teacher effectiveness depends on the individual characteristics of students.

Finally, my results refer to students without certainty of obtaining an offer. I find that parents who could not buy their way into the preferred schools through residential sorting select schools that are specifically effective for their own children. Among these parents, the effects are larger for families with a relatively advantaged socioeconomic background. Although external validity to other parents cannot be assessed by my research design, mo-

tivated and resourced parents who set their residence next to a particular school may make productivity-enhancing choices as I find for marginal families. This possibility may suggest an interesting direction for future work.

References

Abadie, A. and Cattaneo, M. D. (2018). Econometric methods for program evaluation. *Annual Review of Economics*.

Abdulkadiroglu, A., Angrist, J. D., Narita, Y., and Pathak, P. A. (2022). Breaking ties: Regression discontinuity design meets market design. *Econometrica*, 90(1):117–151.

Abdulkadiroglu, 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.

Abdulkadiroglu, A., Pathak, P. A., Schellenberg, J., and Walters, C. R. (2020). Do parents value school effectiveness? *American Economic Review*, 110(5):1502–1539.

Ahn, T., Aucejo, E., and James, J. (2021). The importance of matching effects for labor productivity: Evidence from teacher-student interactions. Working paper.

Ainsworth, R., Dehejia, R., Pop-Eleches, C., and Urquiola, M. (2020). Information and preferences in household demand for school value added. *NBER Working Paper n.28267*.

Angrist, J., Hull, P., Pathak, P. A., and Walters, C. R. (2021). Credible school value-added with undersubscribed school lotteries. *The Review of Economics and Statistics*. Forthcoming.

Angrist, J. D., Hull, P. D., Pathak, P. A., and Walters, C. R. (2016). Interpreting tests of school vam validity. *American Economic Review: Papers & Proceedings*, 106(5):388–392.

Angrist, J. D., Hull, P. D., Pathak, P. A., and Walters, C. R. (2017). Leveraging lotteries for school value added: testing and estimation. *Quarterly Journal of Economics*, 132(2):871–919.

Barseghyan, L., Clark, D., and Coate, S. (2019). Peer preferences, school competition, and the effects of public school choice. *American Economic Journal: Economic Policy*, 11(4):124–158.

Battistin, E. and Neri, L. (2017). School performance, score inflation and economic geography. *IZA Discussion Paper No. 11161*.

Bau, N. (2022). Estimating an equilibrium model of horizontal competition in education. *Journal of Political Economy*, 130(7):1717–1764.

Beuermann, D. W. and Jackson, C. K. (2020). The short and long-run effects of attending the schools that parents prefer. *Journal of Human Resources*. pages 1019-10535R1.

Burgess, S. and Greaves, E. (2013). Test scores, subjective assessment, and stereotyping of ethnic minorities. *Journal of Labor Economics*, 31(3):535–576.

Burgess, S., Greaves, E., and Vignoles, A. (2017). Understanding parental choices ofsecondary school in england usingnational administrative data. Department of Economics, University of Bristol Discussion Paper 17 / 689.

Burgess, S., Greaves, E., and Vignoles, A. (2019). School choice in England: evidence from national administrative data. *Oxford Review of Education*, 45(5):690–710.

Burgess, S., Greaves, E., Vignoles, A., and Wilson, D. (2015). What parents want: School preferences and school choice. *Economic journal*, 125:1262–1289.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Calsamiglia, C. and Guell, M. (2018). Priorities in school choice: the case of the Boston mechanism in Barcelona. *Journal of Public Economics*, 163:20–36.

Campos, C. and Kearns, C. (2022). The impact of neghbourhood on school choice: Evidence from Los Angeles' zones of choice. Working paper.

Carter, G., Pathak, P., and Terrier, C. (2020). Matching practices for primary and secondary schools — England. MiP Country Profile 30.

Chetty, R., Friedman, J., Hilger, N., Saez, E., Schanzenbach, D. W., and Yagan, D. (2011). How does your kindergarten classroom affect your earnings? evidence from project Star. *Quarterly Journal of Economics*, 126(4):1593–1660.

Cornelissen, T., Dustmann, C., Raute, A., and Schoenberg, U. (2018). Who benefits from universal child care? Estimating marginal returns to early child care attendance. *Journal of Political Economy*, 126(6):2356–2409.

De Haan, M., Gautier, P. A., Oosterbeek, H., and Klaauw, B. V. D. (2021). The performance of school assignment mechanisms in practice. Working paper.

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

Department For Education (2017). Key Stage 1: 2017 Assessment and Reporting Arrangements.

Fack, G., Grenet, J., and He, Y. (2019). Beyond truth-telling: Preference estimation with centralized school choice and college admissions. *American Economic Review*, 109(4):1486–1529.

Francis, B. and Hutchings, M. (2013). Parent power? Using money and information to boost children's chances of educational success. Sutton Trust report.

Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):9–15.

Gibbons, S., Machin, S., and Silva, O. (2013). Valuing school quality using boundary discontinuities. *Journal of Urban Economics*, 75:15–28.

Glazerman, S. and Dotter, D. (2017). Market signals: Evidence on the determinants and consequences of school choice from a citywide lottery. *Educational Evaluation and Policy Analysis*, 39(4):593–619.

Gorman, E. and Walker, I. (2021). Heterogeneous effects of missing out on a place at a preferred secondary school in England. *Economics of Education Review*, 81.

Hastings, J. S., Kane, T. J., and Staiger, D. O. (2009). Heterogeneous preferences and the efficacy of public school choice. Working paper.

He, Y. (2017). Gaming the Boston school choice mechanism in Beijing. Working paper.

Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103(6):2052–2086.

Hoekstra, M., Mouganie, P., and Wang, Y. (2018). Peer quality and the academic benefts to attending better schools. *Journal of Labour Economics*, 36(4):841–884.

Jackson, C. K. (2010). Do students benefit from attending better schools? evidence from rule-based student assignments in trinidad and tobago. *Economic Journal*, 120:1399–1429.

Jackson, C. K., Beuermann, D., Navarro-Sola, L., and Pardo, F. (2021). What is a good school, and can parents tell? Evidence on the multidimensionality of school output. *Review of Economic Studies*. Forthcoming.

Kirkeboen, L. J., 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. R. (2016). Evaluating public programs with close substitutes: The case of head start. *The Quarterly Journal of Economics*, 131(4):1795–1848.

Lamadon, T., Mogstad, M., and Setzler, B. (2022). Imperfect competition, compensating differentials andrent sharing in the U.S. labor market. *American Economic Review*, 112(1):169–212.

Machin, S. (2011). Houses and schools: Valuation of school quality through the housing market. *Labour Economics*, 18(6):723–729.

MacLeod, W. B. and Urquiola, M. (2019). Is education consumption or investment? Implications for school competition. *Annual Review of Economics*, 11:563–589.

Masi, B. (2018). A ticket to ride: The unintended consequences of school transport subsidies. *Economics of Education Review*, 63:100–115.

Montacute, R. and Cullinane, C. (2018). Parent power 2018: How parents use financial and cultural resources to boost their children's chances of success. Sutton Trust report.

Mountjoy, J. and Hickman, B. R. (2021). The returns to college(s): Relative value-added and match effects in higher education. *NBER Working Papers n. 29276*.

Pathak, P. A. and Sonmez, T. (2013). School admissions reform in Chicago and England: Comparing mechanisms by their vulnerability to manipulation. *American Economic Review*, 103(1):80–106.

Pop-Eleches, C. and Urquiola, M. (2013). Going to a better school: Effects and behavioral responses. *American Economic Review*, 103(4):1289–1324.

Rothstein, J. M. (2006). Good principals or good peers? Parental valuation of school characteristics, Tiebout equilibrium, and the incentive effects of competition among jurisdictions. *The American Economic Review,*, 96(4):1333–1350.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146.

Terrier, C., Pathak, P. A., and Ren, K. (2021). From Immediate Acceptance to Deferred Acceptance:effects on school admissions and achievement in England. *NBER Working Papers n.29600*.

Walters, C. R. (2018). The demand for effective charter schools. *Journal of Political Economy*, 126(6):2179–2223.

Whitehurst, G. J. (2017). New evidence on school choice and racially segregated schools. *Brookings Economic Studies Evidence Speaks Reports*, 2(33).

Appendix (for on-line publication only)

A School assignment replication

I replicate centralised school assignment by running a student-proposing DA algorithm starting from data on parental preferences, distance to school and school capacity.[1]

I start by constructing a proxy for siblings at the school, which constitutes the main source of unobserved admission priority. My proxy is based on postcode of residence (in my sample, there are approximately 4.5 students per postcode on average) and on family-specific observable characteristics.[2] I flag an applicant as supposed sibling if a student living in the same postcode with exactly the same family-specific observables is enrolled at the school of choice at the time of application. This variable provides an upper bound to the number of siblings at the school, which I adjust by ignoring (the very few) implausibly high values.[3] As a result, the 32% of first-choice applications, and 13% of all applications, are from students with supposed siblings at the school.

Several pieces of evidence suggest this measure provides a reasonable proxy to admission priority. First, one could expect strong incentives for parents to send their children at the same school. The 85% of applicants with potential siblings in just one school, indeed, rank it as first choice. Second, applicants with supposed siblings are extremely likely to be admitted. The 97% of applicants with supposed siblings obtain an offer from the most preferred school (compared to the 82% on average). Third, when the catchment boundary is estimated only based on distance, the 94% of students with a supposed sibling at their first choice who reside beyond the admission cutoff obtain an offer (compared to 66% on average). Fourth, if ranking schools involves an effort cost (Fack et al., 2019), one can expect parents with siblings at the school, who are extremely likely to obtain an offer, to express fewer preferences. The number of schools listed by supposed siblings is on average approximately 2.5, against 3.5 for all other applicants. Finally, the share of applicants with at least one sibling in primary school, 37%, is broadly in line with available statistics.[4]

[1] I proxy school capacity with the number of offers issued. This is a lower bound of the real capacity if a school is not oversubscribed. The distribution of school capacity looks as expected, with spikes around multiples of 30 (the statutory class size cap, see Figure A.1).

[2] I use free school meal eligibility, ethnic group, and language spoken at home, which depend on parental socioeconomic background. I observe a granular measure of ethnicity spanning 17 different groups.

[3] Specifically, I do not assign priority to students with a supposed sibling at more than two different schools (0.66% of applicants) or with more than four supposed siblings in total (3.34%). These numbers likely reflect postcode density rather than siblings at the school.

[4] 43% of sixth-grade students born in 2000-02 have a siblings in secondary school (Burgess et al., 2017).

Assignment replication proceeds in two steps. First, I replicate centralised assignment based on distance to school, parental preference, and the admission priority proxy for siblings at the school discussed above. Within priority groups, I rank applicants to a school in ascending order of distance and iteratively eliminate candidates who are eligible at schools ranked with higher preference. Catchment boundary is defined at each oversubscribed school as the distance of the last admitted applicant.[5] As some applicants with priority may remain undetected, this is not sufficient to fully replicate school offer. Panel B of Figure A.2 shows that predicted offer is wrong for almost 8% of applicants. This first step, however, provides useful information to complete the replication.

Second, I rely on the observation of the centrally assigned school offer and exploit the idea that, if an applicant located beyond the initially estimated catchment boundary receives school offer, she must have been admitted with priority. The catchment boundary estimated in the first step of my algorithm is an upper bound of the true threshold as some candidates admitted with priority remain undetected. Therefore, any school offer granted to applicants located beyond the initially estimated threshold reveals priority in admission. I flag these applicants and re-attempt the replication of school assignment by admitting them first. This procedure is iterated until no applicant with offer is found beyond the estimated catchment boundary.

In details, my algorithm works as follows.

1. Rank all applicants, regardless of their preference, by priority group and, within priority group, in ascending order of distance to school. Each student is ranked at up to 6 schools, depending on the number of schools listed. I start with two priority groups: applicants with supposed siblings, and all other students.

2. All applicants ranked within school capacity are eligible for admission at the school. If eligible at one school, the applicant is dropped from the list at all schools ranked with lower preference. This is executed sequentially preference by preference as follows.

   (a) Consider first-choice school. If an applicant is eligible, drop the applicant from the queue at schools ranked second to sixth.

   (b) Re-rank applicants at all schools considering only those retained after step (a).

[5] Catchment boundary is not defined for undersubscribed schools and religious schools.

(c) Repeat (a) and (b) analogously for second to fifth choice. In particular, if an applicant is eligible at the s-th choice, drop the applicant from the queue at all schools ranked lower than s. Retained applicants are re-ranked.

3. Repeat step 2 until no more applicant is dropped from the admission list. Assignment converges in at most 13 iterations.

4. Assign priority to applicants who are admitted to school according to administrative records but who would not receive offer based on steps 1-3.

5. Repeat steps 1-4 until no more applicants with priority are detected. The algorithm converges in less than 60 iterations.

   Steps 1-3 replicate the DA algorithm used by school districts to assign applicants to school seats. Steps 4 and 5 correct the replication by detecting applicants with unobserved priority. In each iteration, at the end of step 4, I store dummies indicating admission priority and correspondence between actual and replicated school offer. I also keep track of median catchment area boundary, defined as distance to school of the last applicant admitted. The algorithm is executed separately for both cohorts, and data are stacked across cohorts.

Algorithm performance after 60 iterations is illustrated in Figure A.2. Convergence is shown in Panel A, plotting the fraction of applicants with priority identified in each iteration. This fraction monotonically decreases to zero at an increasingly slow pace, and is virtually flat from iteration 45. At the last iteration, 0.00005% of applicants are flagged with unobserved priority. In total, approximately 9.5% of applications enjoy unobserved priority and these are disproportionally found at faith schools, as expected.[6] Panel B of Figure A.2, depicting errors in school assignment by iteration, shows my final assignment almost perfectly corresponds to the true school offer. At the last iteration, just 0.0002% of applicants are wrongly assigned and the correlation of predicted offer with the true offer is 99.9995%. School offer discontinuously drops at the estimated boundaries of preferred schools, as shown in Panel A of Figure A.3.

[6] At the median non-religious school, 7% of applicants have unobserved priority. This can reflect undetected siblings or, more likely, other less frequent priorities such as children of staff (commonly granted just to tenured teachers, and often subject to shortage in particular subjects).

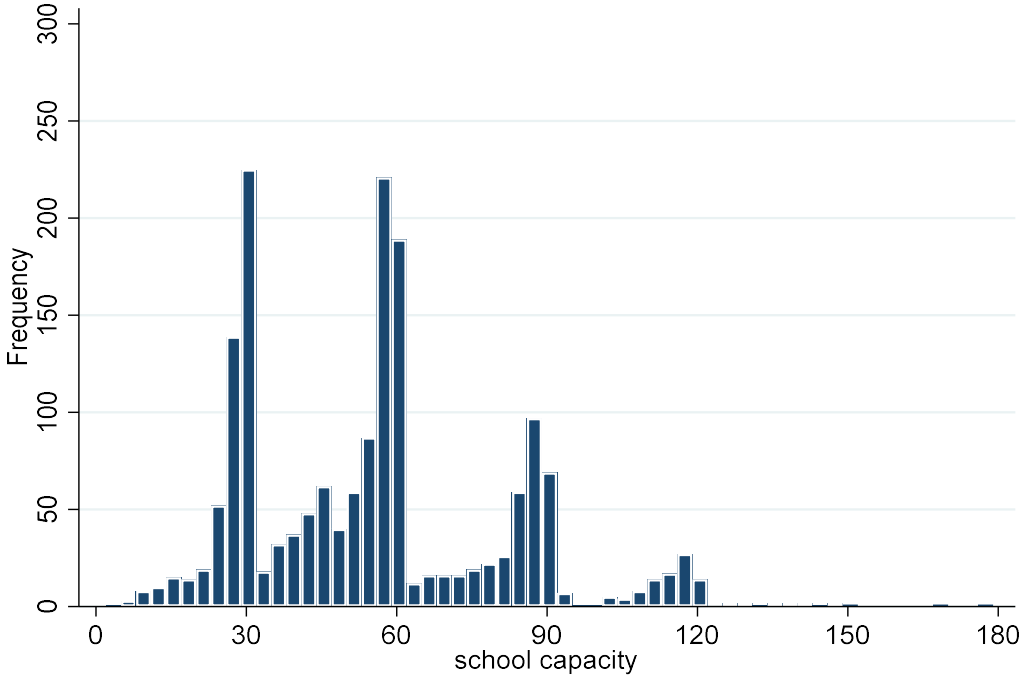
Panel B shows a similar, fuzzier, pattern for school enrolment, as expected.[7] Consistent with the idea that catchment boundary is initially overestimated, Panel C of Figure A.2 shows that median distance threshold decreases roughly monotonically as the algorithm is iterated. When assignment replication is concluded, the estimated median catchment boundary at oversubscribed schools is approximately 470 metres.[8]

Catchment boundary is stable over time on average, but the change in distance threshold exhibits remarkable dispersion. Among the 705 non-faith schools that are oversubscribed in both years 2014 and 2015, the catchment boundary shrinks by 10 metres on average. For half of institutions the year-on-year change is within 200 metres, and it is within 500 metres for the 80% of schools. Figure A.4 depicts catchment boundary in the two years considered, focusing on the 576 schools with distance thresholds within 1 km in both periods. Markers are weighted proportionally on enrolment count. Despite most schools lie around the 45-degree line, benchmarking an unchanged boundary, a substantial fraction of schools exhibit 2016 boundaries sensibly above or below the 2015 counterpart. These changes, caused by fluctuations in application and residential choices across cohorts, have no predictable direction and are unlikely to be anticipated by parents.

[7] The figures consider only applicants with marginal priority, that is the same admission priority of the last-admitted applicant. In addition, for schools ranked second or lower, it considers only applicants who are not offered a more-preferred school. This selection yields the sub-group of applicants for which offer is sharply determined by the catchment boundary.

[8] A minor concern is that catchment boundary is measured with error if the last admitted applicant enjoys unobserved admission priority. In that case, the correct cutoff is the distance of the applicant located immediately closer to the school. Density of applicants in London implies that such error is at most very small. Nevertheless, catchment boundary is constant for all applicants to a given school and measurement error cancels out when comparing students around the cutoff.

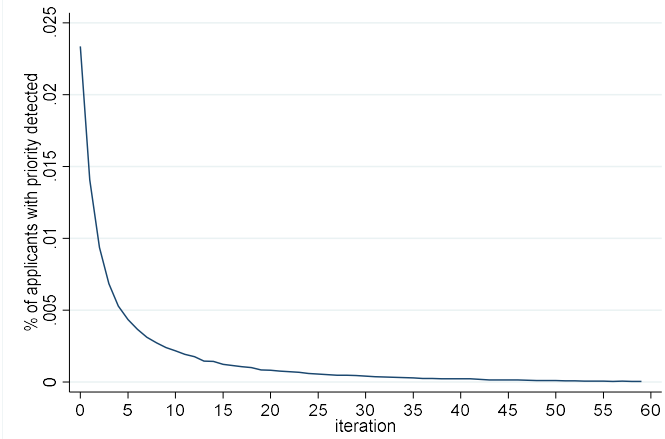
Figure A.1: School capacity



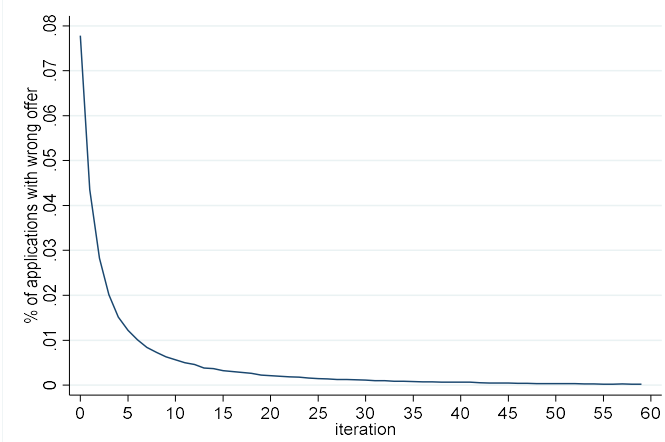
Note. The figure plots the distribution of school capacity in London primary schools. Capacity is approximated by the number of offers issued. Bars represent frequency counts in three-units-wide bins, computed using one observation per school-year. See Section 3 and Appendix A for details.

Figure A.2: Replication of school assignment

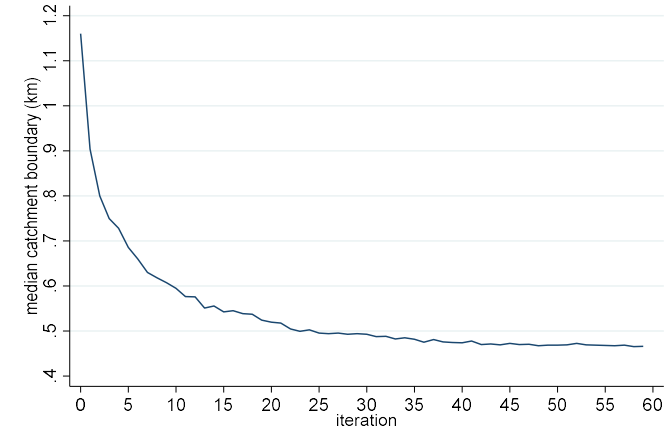
Panel A. Unobserved priorities



Panel B. Error in school offer replication

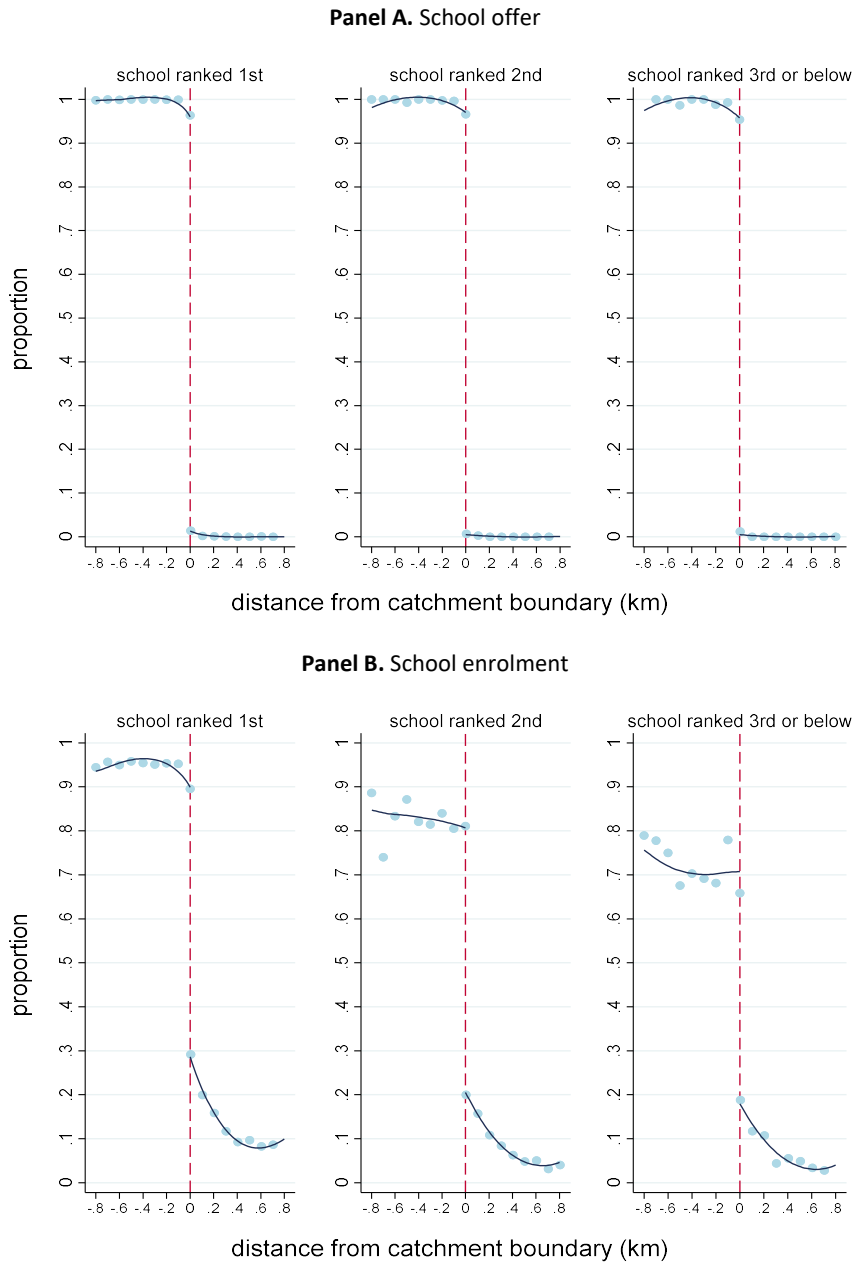


Panel C. Estimated catchment area boundary



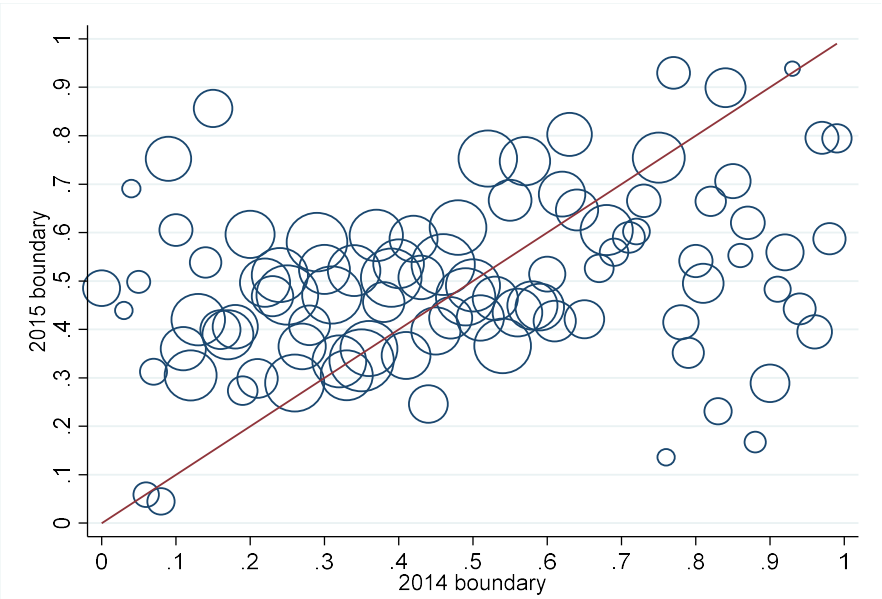
Note. The figure plots the fraction of applicants with admission priority detected (Panel A), the fraction of applicants with wrong predicted offer (Panel B), and median estimated catchment boundary (Panel C) by iteration of the school assignment replication algorithm. The sample includes all applicants to at least one London primary school in 2014 or 2015. Applicants are ranked by sibling status (proxied as detailed in Appendix A) and proximity in iteration 0. Applicants with offer beyond estimated boundary are then flagged as enjoying priority. Subsequent iterations rank pupils by priority as retrieved in the previous round, sibling status, and proximity. Reported are average figures across the two cohorts considered. See Section 3 and Appendix A for details.

Figure A.3: Tie-breaking



Note. The figure plots school offer (Panel A) and enrolment (Panel B) around catchment boundary for schools ranked first, second and third or below at application. Enrolment is measured at the reception year. Catchment boundary is defined only for oversubscribed non-faith schools. Distance to school catchment boundary on the horizontal axis is defined by subtracting the distance of the last admitted candidate from an applicant's distance to school. Negative values indicate residence within the catchment. Markers represent average values in 100-metre-wide bins of distance from the boundary, and the solid line is a local linear fit of underlying observations, estimated separately on either side of the cutoff. The sample is restricted to applicants within 800 meters from the catchment boundary and to applicants at risk of admission at the school, i.e. those with marginal admission priority and not eligible at any school ranked higher. See Section 3 and Appendix A for details.

Figure A.4: Year-on-year change in catchment boundaries



Note. The figure plots estimated catchment area boundary of oversubscribed, non-faith state schools in 2014 and 2015. Boundaries are traced by replicating the centralised assignment mechanism for all state primary schools in London. Reported is average 2016 boundary as function of 10-metre-wide bins of 2015 boundary. Markers show one observation per school, with size proportional to enrolment count. The 45-degree line, indicating unchanged catchment boundary, is reported in red. Sample is restricted to the 576 schools with catchment boundary within 1 kilometre in both years and issuing at least 30 offers (attracting the 46% of appications, and the 74% among oversubscribed schools). See Appendix A for details.

B Estimation of school VA and of school match effects based on student's characteristics

Estimates of average school VA are employed in my empirical analysis to hold school quality constant. Following Deming et al. (2014), I compute school average residuals from a student-level regression of standardised KS1 assessments on observable characteristics and baseline achievement. Specifically, I consider the 2009-2016 cohorts (N=688,748) of achievement data and estimate the following regression model for student i enrolled at school s in Year 2 at time t:

$$Y_{ist} = \eta_o + X_i'\eta_1 + \eta_2 Y_{is,t-2} + \delta_t + \epsilon_{ijt}, \tag{11}$$

where the vector of controls X_i' includes dummies for gender, language, ethnicity, free lunch eligibility, special education needs and local area of residence (MSOA), and δ_t are year dummies.[9] I additionally control for lagged achievement using EYFSP assessments.[10] In this model, estimates of η_1 and η_2 proxy the impact of individual ability on student achievement (ν_i in equation 1). I estimate school VA as institution-level average of the residuals from this regression ($\hat{\epsilon}_{ijt}$):

$$\hat{\alpha}_s = \frac{\sum_t \sum_i \hat{\epsilon}_{ist}}{\sum_t N_{st}},$$

where N_{st} is the number of students enrolled at school s.

Under heterogeneous school effects, I need additional assumptions for school VA estimates to be comparable across schools. By comparing equation (11) with (1), it can be seen that residuals from (11) capture both school VA and average match effects at the school:

$$\hat{\alpha}_s = \alpha_s + \bar{\mu}_s,$$

where $\bar{\mu}_s = 1/N \sum_i \mu_{is}$. VA estimates, therefore, are comparable across institutions only if the average match effect is constant across schools.[11] Under the hypothesis that match effects

[9] I consider subject-specific KS1 test scores for the 2009-2015 cohorts. For the 2016 cohort, where only categorised teacher assessments are available, I use average point scores as for outcomes used in my main analysis (see Section 3).

[10] Specifically, I consider total EYFSP scores available for cohorts the 2009-2014 cohorts. For 2015-2016 cohorts, I average results over the 17 learning goals assessed. Both KS1 outcomes and EYFSP controls are standardised by year. Since EYFSP are administered during the reception year, they could be influenced by school quality. However, they are meant to inform teachers and parents on student educational needs, and schools are not held accountable on the basis of EYFSP results. They represent therefore a good proxy of student academic ability at school entrance.

[11] See Mountjoy and Hickman, 2021 for a detailed discussion of this hypothesis, which they name "symmetric

are an increasing function of parental preference, this requires the average rank assigned by parents of enrolled students to be constant across schools. Variation in the latter quantity is likely limited in my empirical analysis, as only oversubscribed schools contribute to the identification. Nonetheless, with large differences in school popularity among parents, my estimates would constitute lower bounds of the true match effects since school VA would be inflated at schools with higher average parental preference. Results from my alternative design exploiting within-school variation are similar to the main findings (see Section 5), suggesting that, in practice, differences in average match effects across schools do not drive my results.

One σ higher VA improves KS1 score by about 0.16σ higher achievement, a comparable magnitude to what found in related studies (Angrist et al., 2017; Abdulkadiroglu et al., 2020; Angrist et al., 2021). The distribution of school VA is plotted in Figure B.1. Estimated school quality has approximately 35% correlation with absolute achievement, suggesting high-performing schools are not necessarily highly effective. VA is estimated separately by subject – mathematics, reading, and writing – and the latter two estimates are averaged to obtain school VA in English. To estimate match effects in equation (9), I consider school VA in the same subject of the outcome variable considered.

Estimated school VA strongly predicts student achievement, as shown in Table B.1. Impacts of enrolment at a more-preferred schools on student's school VA are reported in Table B.2.

Following Abdulkadiroglu et al. (2020), I additionally estimate match effects based on observable student characteristics as persistent school effects for specific types of students. First, I group students in eight covariate cells based on the observable characteristics that most strongly predict academic achievement. Specifically, I use gender, an indicator for above-median baseline ability, and a socioeconomic disadvantage indicator. I define a student as socioeconomically disadvantaged if either eligible for free school meals or residing in an area with deprivation index above the median among students eligible for free school meals. I show in table B.3 that these characteristics are highly predictive of academic achievement, and that further including other student characteristics such as ethnicity or language does not substantially increase the explained variability.

sorting on gains".

Second, I estimate match effects through an OLS regression interacting school indicators with each covariate cell ($1762X8$coefficients). Specifically, I estimate uninteracted school effects referring to low-ability, male, and non-disadvantaged students, and include interactions of school indicators with all other covariate cells. I compute bootstrapped standard errors of match effects with 300 replications. Third, I use standard errors to obtain linear Empirical Bayes posteriors of match effects following equation (5) in Angrist et al. (2021). The latter are minimum-MSE estimators of noisy regression coefficients. Table B.4 shows mean and standard deviation of estimated match effects.

Figure B.1: Estimated school VA

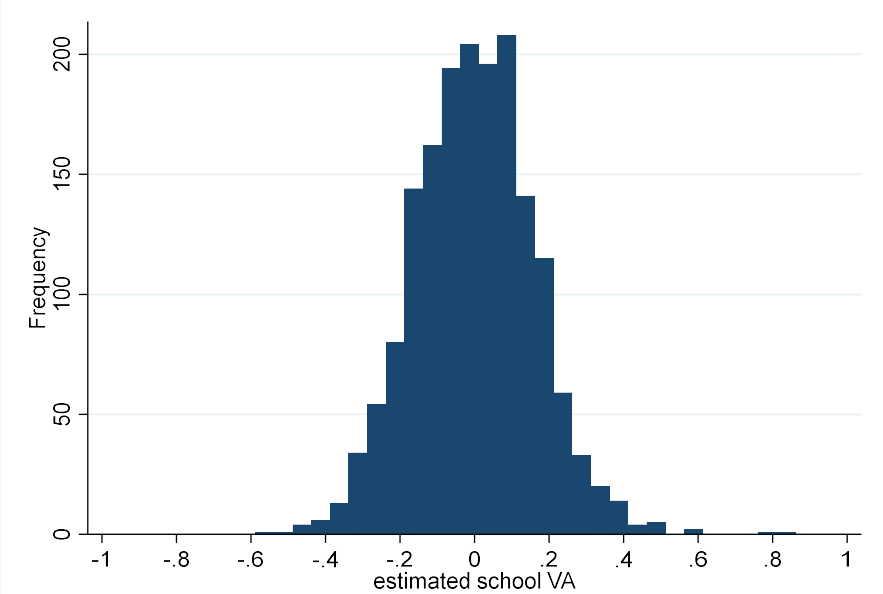


Table B.1: Validation of school VA estimates

	All students			Students at risk			Students at risk enroled for all three years		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				Panel A. KS1 score in mathematics					
Average school VA	0.3156***	0.3440***		0.3736***	0.4116***		0.4213***	0.4681***	
	(0.0116)	(0.0118)		(0.0316)	(0.0322)		(0.0332)	(0.0338)	
Growth of average school VA		0.3632***			0.4372***			0.5315***	
		(0.0151)			(0.0426)			(0.0456)	
School match effect based on student's characteristics			0.3130***			0.3600***			0.4019***
			(0.0106)			(0.0288)			(0.0303)
R2	0.3721	0.3749	0.3726	0.3723	0.3762	0.3727	0.3807	0.3858	0.3811
N	172,099	172,099	172,099	22,628	22,628	22,628			
				Panel B. KS1 score in English					
Average school VA	0.3653***	0.4069***		0.4163***	0.4673***		0.4566***	0.5141***	
	(0.0114)	(0.0116)		(0.0309)	(0.0316)		(0.0325)	(0.0330)	
Growth of average school VA		0.4659***			0.5575***			0.6647***	
		(0.0153)			(0.0435)			(0.0466)	
School match effect based on student's characteristics			0.3308***			0.3839***			0.4326***
			(0.0102)			(0.0283)			(0.0298)
R2	0.4358	0.4393	0.4358	0.4347	0.4396	0.4349	0.4439	0.4503	0.4444
N	172,099	172,099	172,099	22,628	22,628	22,628	19,635	19,635	19,635
Year FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y
Individual characteristics	Y	Y	Y	Y	Y	Y	Y	Y	Y
Baseline achievement	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note. The table shows the impact of estimated school VA on student achievement. Reported are estimates from regressions of KS1 score in mathematics (Panel A) or English (Panel B) on estimated school VA, VA growth, or match effect based on student's characteristics. Columns (1)-(3) consider all students, columns (4)-(6) restrict the sample to students with nondetermistic assignment risk at one or more ranked schools, columns (7)-(9) further restrict to students enroled in the primary school where observed in the reception year until the end of KS1 (three academic years). All regressions control for cohort dummies, gender, language, ethnicity, free school meal eligibility, special education needs, deprivation index and education level in the neighbourhood of residence, and baseline achievement. Robust standard errors are reported in parentheses. See Section 5 for details on estimation, and Appendix B for details on estimated school VA and match effects. ***p<0.01. ** p<0.05. * p<0.1

Table B.2: Impacts on school VA

	Average school VA	Growth in average school VA	School match effects based on student's characteristics
		Panel A. Two-choice model	
Enroled in 1st choice	0.0123*	0.0033	0.0216***
	(0.0065)	(0.0064)	(0.0083)
Enroled in 2nd choice	0.0043	-0.0005	0.0120*
	(0.0056)	(0.0052)	(0.0070)
		Panel B. Three-choice model	
Enroled in 1st choice	0.0220***	0.0064	0.0309***
	(0.0078)	(0.0074)	(0.0099)
Enroled in 2nd choice	0.0151**	0.0033	0.0212**
	(0.0073)	(0.0068)	(0.0090)
Enroled in 3rd choice	0.0252***	0.0090	0.0234**
	(0.0085)	(0.0079)	(0.0106)
N	22,628	22,628	22,628
Tie-breaker controls	Y	Y	Y
Assignment risk (all schools)	Y	Y	Y
Individual characteristics	Y	Y	Y
Baseline achievement	Y	Y	Y

Note. The table shows estimates of the impact of enroling at most-preferred schools on school value-added. Estimates and specifications follow column 1 of Table 3. Robust standard errors are reported in parentheses. See Section 4 for details. ***p<0.01. ** p<0.05. * p<0.1

Table B.3: Covariates for school match effect estimation

	KS1 score in mathematics			KS1 score in English		
	(1)	(2)	(3)	(4)	(5)	(6)
Free school meal eligible	-0.1715***		0.0068**	-0.1808***		-0.0527***
	(0.0023)		(0.0027)	(0.0021)		(0.0025)
Female	-0.1822***		-0.0623***	0.1083***		0.0188***
	(0.0018)		(0.0030)	(0.0017)		(0.0027)
Deprivation index in the area of residence	-0.0371***		0.0942***	-0.0286***		0.0861***
	(0.0051)		(0.0031)	(0.0046)		(0.0028)
Baseline achievement	0.6193***		-0.6947***	0.6060***		-0.5488***
	(0.0011)		(0.0096)	(0.0010)		(0.0083)
White		0.0537***	0.1220***		-0.0024	0.0855***
		(0.0033)	(0.0021)		(0.0032)	(0.0019)
Black		-0.1414***	-0.1524***		-0.0621***	-0.1736***
		(0.0037)	(0.0023)		(0.0035)	(0.0021)
Asian		0.1611***	-0.1869***		0.1558***	0.1037***
		(0.0039)	(0.0018)		(0.0036)	(0.0016)
Special Education Needs		-2.0757***	-0.1195***		-1.9837***	-0.1614***
		(0.0122)	(0.0053)		(0.0105)	(0.0049)
Speaking English at home		-0.0824***	0.5948***		-0.1208***	0.5879***
		(0.0025)	(0.0011)		(0.0024)	(0.0010)
Constant	0.1726***	0.0841***	0.1476***	0.0401***	0.1182***	0.0580***
	(0.0021)	(0.0030)	(0.0031)	(0.0019)	(0.0029)	(0.0028)
N	688,748	688,748	688,748	688,748	688,748	688,748
R2	0.4117	0.0989	0.4290	0.4628	0.0949	0.4762

Note. The table shows predictive power of student characteristics on academic achievement. Estimates are from regressions of KS1 score in mathematics (columns 1-3) or English (column 4-6). Controls included in columns (1), and (4) are student characteristics used to define covariate cells for school match effect estimation in Appendix B Subsequent columns include other student characteristics, and show that predictive power of student achievement does not improve (see the R2 coefficient reported at the bottom of the table). Robust standard errors are reported in parentheses. See Appendix B for details. ***p<0.01. ** p<0.05. * p<0.1

Table B.4: Estimates school match effects based on student's characteristics

	Student covariates			Estimated match effects: mathematics		Estimated match effects: mathematics	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Disadvantage	**Female**	**High-ability**	**Mean**	**SD**	**Mean**	**SD**
Average school VA	0	0	0	-0.0026	0.2775	-0.0023	0.2719
Interactions with covariate cells							
Cell 2	0	0	1	0.0113	0.1657	0.0011	0.1865
Cell 3	0	1	0	-0.0867	0.1331	-0.0812	0.1575
Cell 4	0	1	1	-0.0302	0.1837	-0.0283	0.1859
Cell 5	1	0	0	0.0465	0.1857	0.0551	0.2122
Cell 6	1	0	1	0.0388	0.2009	0.0545	0.2234
Cell 7	1	1	0	-0.0633	0.2003	-0.0506	0.2112
Cell 8	1	1	1	0.0168	0.1899	0.0175	0.1960

Note. The table shows estimated school match effects based on student's characteristics. Columns (1)-(3) describe the eight covariate cells in which students are grouped. Match effect estimates result from regressions of KS1 scores of school 1,762 school indicators with covariate cells. Standard errors are bootstrapped using 300 replications. Empirical Bayes (EB) posteriors are obtained for each covariate cell by multiplying a school's coefficient for the estimated variance of the coefficients across schools divided by the sum of the estimated variance and the squared standard error of a school's coefficient. EB estimates' mean and standard deviations across schools are reported in columns (4) and (5), respectively, for achievement in mathematics, and in columns (6) and (7), respectively, for achievement in English. See Appendix B for details.

C Estimation of assignment risk

I use information on preferences, priorities, and distance for all applications to estimate their risk of receiving an offer. Following Abdulkadiroglu et al. (2022), I proceed in three steps by considering the centralised assignment inputs one at a time.

Assignment risk depends, first, on admission priority. The key priority group to assess admission chance at a given school is the one of the last admitted student, defined as *marginal priority*. Applicants with higher than marginal priority (e.g., siblings of current students) receive an offer with certainty, while applicants with lower than marginal priority are never admitted. Assignment risk, therefore, is non-degenerate only for applicants with marginal priority. Based on priority status obtained in Section 3, students in my sample have marginal priority in 48% of applications.[12]

Conditional on priority, assignment risk depends on the value of the tie-breaker. Applicants located sufficiently close to the catchment boundaries face equal and non-degenerate chance of admission. On the contrary, applicants located comfortably within the boundary receive an offer with certainty, while those located well beyond the boundary are never admitted. To approximate this idea, applicants are grouped based on a narrow distance bandwidth around the cutoff.[13] Let τ_s be the catchment boundary at school ranked s-th and let δ denote the optimal bandwidth. Applicants can be partitioned in three groups based on their chance of admission:

- student i with marginal priority at the s-th choice is *conditionally seated* if $d_{is} \in [\tau_s - \delta, \tau_s + \delta]$,
- is *always seated* if her priority is higher than marginal or $d_{is} < (\tau_s - \delta)$,[14]
- is *never seated* otherwise.

[12] Admission priority is likely observed with significant error for applicants at faith schools (see Section 3). Therefore, I do not exploit faith schools in estimation. Specifically, I assign to faith school applications higher- or lower-than-marginal priority depending on offer status so that they do not generate assignment risk.

[13] I use catchment boundaries obtained in Appendix A and select the optimal data-driven bandwidth proposed by Calonico et al. (2014). Following Abdulkadiroglu et al. (2022), I select the minimum bandwidths across outcomes (achievement in mathematics and English) separately for each cohort, obtaining values of 333 and 243 metres in 2015 and 2016, respectively.

[14] Note that all applicants to undersubscribed schools are always seated as $\tau_s = \infty$.

Students in my sample are conditionally seated in approximately 12% of applications.

Finally, assignment risk depends on parental preferences. In particular, an applicant who would be admitted based on priority and distance, will not receive an offer if she is eligible at more-preferred institutions. I compute for each application the number of schools ranked higher than s where a student is conditionally (B^c_{is}) or always seated (B^a_{is}) based on the definition above.

The assignment risk of student i at the s-th-choice school is:[15]

- $p_{is} = 0$ if $B^a_{is} > 0$ or i is never seated,
- $p_{is} = 0.5^{B^c_{is}}$ if i is always seated,
- $p_{is} = 0.5^{(1+B^c_{is})}$ if i is conditionally seated.

[15]This formula derives as a particular case of Theorem 1 in Abdulkadiroglu et al. (2022) where the *most informative disqualification* is zero. In my context, indeed, tie-breakers are school-specific at all institutions.

Table C.1: Descriptive statistics (estimation sample)

	All applicants		Applicants at risk	
	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)
Baseline Characteristics				
FSM eligible	0.1545	0.3615	0.1443	0.3514
Not speaking English at home	0.4216	0.4938	0.4042	0.4907
White	0.4140	0.4926	0.4019	0.4903
Asian	0.1978	0.3984	0.2054	0.4040
Black	0.1662	0.3723	0.1324	0.3389
Special education needs	0.0066	0.0808	0.0064	0.0798
Female	0.4919	0.4999	0.4952	0.5000
Deprivation index (LSOA)	0.2921	0.1628	0.2851	0.1651
% with higher education (LSOA)	0.3365	0.1279	0.3645	0.1341
Baseline achievement	35.2023	7.2186	35.4323	7.2480
School choice variables				
N. of schools listed	3.1495	1.8750	3.6266	1.8248
Offered 1st choice	0.8472	0.3598	0.6921	0.4616
Enroled at offered school at reception year	0.9315	0.2526	0.9031	0.2959
N	172,099		22,628	

Note. The table shows descriptive statistics about applicants to at least one mainstream state-funded primary school in Greater London in 2014 and 2015. Columns (1)-(2), report averages and standard deviations for all applicants. Columns (3)-(4) consider applicants with non-deterministic assignment risk at one or more listed schools. See Section 5 and Appendix C for details.

D Imperfect compliance

I start here by maintaining the assumption that offers are randomly assigned, and relax this hypothesis below. I also maintain the assumption that students enrol to one of the ranked schools throughout this Appendix. The notation can be extended to include enrolment at non-ranked schools using the same reasoning.

Let $D_i = \{1, \ldots, 6\}$ be the preference rank for the school where student i enrols. With imperfect compliance, this might differ from the preference rank for the offered school, Z_i. The observed outcome for student i is:

$$Y_i = \sum_{d=1}^{6} \mathbb{1}(D_i = d) \cdot Y_{is(i,d)}$$

Let $D_i(z)$ denote potential enrolment as a function of school offer, representing the preference rank for the school where student i enrols if she receives an offer from her z-th choice. Let $\tilde{Y}_i \equiv Y_i - \hat{\alpha}_{s(i)}$ denote the VA-adjusted achievement of student i enrolled at school s (see Appendix B for details on estimation of school VA). The comparison in the left-hand-side of equation (3) yields:

$$\begin{aligned} E \quad & [\tilde{Y}_i | Z_i = 1] - E[\tilde{Y}_i | Z_i = 2] = \\ = \quad & E[\sum_{d=1}^{6} \mathbb{1}(D_i(1) = d) \cdot \tilde{Y}_{is(i,d)} | Z_i = 1] - E[\sum_{d=1}^{6} \mathbb{1}(D_i(2) = d) \cdot \tilde{Y}_{is(i,d)} | Z_i = 2] = \\ = \quad & E[\sum_{d=1}^{6} \mathbb{1}(D_i(1) = d) \cdot \tilde{Y}_{is(i,d)} - \sum_{d=1}^{6} \mathbb{1}(D_i(2) = d) \cdot \tilde{Y}_{is(i,d)}], \end{aligned} \tag{12}$$

where the latter equality uses random assignment of offers.

A plausible monotonicity assumption is imposed to identify the parameter of interest. Assume that students do not enrol at a school with lower parental preference than the one they are offered, $D_i(z) \leq z \, \forall z = 1, \ldots, 6$. This assumption implies $D_i(1) = 1$ and $D_i(2) = \{1, 2\}$. Let $C_i \equiv \prod_{z=1}^{6} \mathbb{1}(D_i(z) = z)$ denote compliance with school offer. The expression in (12) equals:

$$\begin{aligned} E[\tilde{Y}_{is(i,1)} - \tilde{Y}_{is(i,2)} | C_i = 1] \cdot P(C_i = 1) \quad & = \\ = E[\mu_{is(i,1)} - \mu_{is(i,2)} | C_i = 1] \cdot P(C_i = 1) \quad & , \end{aligned} \tag{13}$$

where the latter equality uses random assignment of offers. To rescale the reduced form comparison by the proportion of compliers, I consider 2SLS models where enrolment is in-

strumented by school offer. Let $D_{i1} \equiv \mathbb{1}(D_i = 1)$ and $Z_{i1} \equiv \mathbb{1}(Z_i = 1)$ be dummy variables indicating enrolment at and offer from the first choice, respectively. Instrumenting D_{i1} with Z_{i1} implements the following comparison:

$$\frac{E[\tilde{Y}_i|Z_{i1}=1] - E[\tilde{Y}_i|Z_{i1}=0]}{E[D_{i1}|Z_{i1}=1] - E[D_{i1}|Z_{i1}=0]} \quad = \quad E[\mu_{is(i,1)} - \sum_{z=2}^{6} \mathbb{1}(Z_i = z) \cdot \mu_{is(i,z)}|C_i = 1]. \quad (14)$$

Under the simplifying assumption that students missing out on the first choice are offered the second choice, the last equation identifies the average match effect of enrolling at the first choice relative to the second choice for school offer compliers. Similar comparisons can be defined for schools ranked less than first choice.

Although offers are not randomised, they are as-good-as random conditional on assignment risk. The average match effect among compliers is identified by a weighted average of conditional versions of the comparison in equation (14):

$$\int \frac{E[\tilde{Y}_i|Z_{i1}=1, \boldsymbol{p_i}] - E[\tilde{Y}_i|Z_{i1}=0, \boldsymbol{p_i}]}{E[D_{i1}|Z_{i1}=1, \boldsymbol{p_i}] - E[D_{i1}|Z_{i1}=0, \boldsymbol{p_i}]} \cdot d(\boldsymbol{p_i}) \quad = \quad E[\mu_{is(i,1)} - \mu_{is(i,2)}|C_i = 1]. \quad (15)$$

E Alternative approach to hold VA constant

Consider comparing unadjusted achievement of students with offer from their first or second choice. Using equation (2):

$$E[Y_i|Z_i = 1] - E[Y_i|Z_i = 2] \quad = \quad E[\alpha_{s(i,1)} - \alpha_{s(i,2)}] + E[\mu_{s(i,1)} - \mu_{s(i,2)}], \quad (16)$$

where I assume that school offers are randomised. The comparison in equation (16) conflates the match effect from attending the first relative to the second choice ($\mu_{is(i,1)} - \mu_{is(i,2)}$) and the gain in school VA ($\alpha_{s(i,1)} - \alpha_{s(i,2)}$). The parameter of interest is the last term of this equation, representing the average match effect from attending the first relative to the second choice.

One possible strategy to hold school VA constant is to adjust observed outcomes by school VA, as in equation (3). An alternative strategy is to focus on applicants whose counterfactual school has the same quality:

$$E[Y_i|Z_i = 1, \alpha_{s(i,1)} = \alpha_{s(i,2)}] - E[Y_i|Z_i = 2, \alpha_{s(i,1)} = \alpha_{s(i,2)}] = E[\mu_{s(i,1)} - \mu_{s(i,2)}]. \quad (17)$$

To implement this comparison, I consider un-adjusted achievement and augment equation (9) by adding an interaction of school enrolment with the expected VA gain. Let $\hat{\alpha}_{s(i,r)}$ be

the estimated VA of the school ranked r-th by student i. First, I compute the average VA of schools ranked lower than second choice, $\bar{\hat{\alpha}}_i = (\sum_{r=3}^{6} \hat{\alpha}_{s(i,r)})/4$. Expected VA gains from entering the top two choices are then obtained as $(\hat{\alpha}_{is(i,r)} - \bar{\alpha}_i)$, for $r = 1, 2$. Finally, match effects are estimated by θ_1 and θ_2 in:

$$Y_{is(i,r)} = \sum_{r=1}^{2} \theta_r D_{is(i,r)} + \sum_{r=1}^{2} \psi_r D_{is(i,r)} \times (\hat{\alpha}_{s(i,r)} - \bar{\hat{\alpha}}_i) + f^p(\boldsymbol{p_i}) + f^d(\boldsymbol{d_i}) + u_{4i}, \tag{18}$$

with analogue first stage equations. Estimates of θ_1 and θ_2 corresponds to the comparison in equation (17).

F Construction of individual feasible school set

I define the individual feasible school set exploiting school catchment boundaries obtained from replication of centralised school assignment (see Section 3 and Appendix A). I compute linear distance between student postcode and all primary schools, including those not ranked by parents. Specifically, I pair each student with all schools ranked by at least one applicant residing in the same school district. This mild restriction ensures computational feasibility, as there are approximately 200,000 applicants and 1,750 schools in my sample. I view this as a natural assumption since 93.2% of students enrol in the LA of residence.

I define a school as ex-post feasible if the student resides within the catchment boundary or if the school remained undersubscribed. I exclude religious schools from choice set since I do not accurately observe ex-post feasibility for these institutions. Moreover, even when applicants reside beyond the catchment boundary, a school is considered feasible if the student has a currently enrolled sibling (proxied as detailed in Appendix A) or the student is offered a seat. A school is included in the individual choice set if located within 2 km from student residence, corresponding to the 90th percentile of distance to school. The individual feasible school set is defined as the collection of ex-post feasible schools.

G Parental choice and school mobility

I have shown in Section 6 that the decision to change school after reception year is associated with parental preference for the offered school (see Figure 6). I further show here that

school mobility responds to peer quality rather than school VA, consistently with parental preferences for schools. Panel A of Table H.2 presents estimates from regressions of school mobility on relative attributes of the most preferred school and the school where student enrols. When a student's school falls short of the first choice in terms of peer quality by 1σ, the likelihood of moving to another school by Year 2 increases by 5-6 p.p. Controlling for school choice covariates or individual characteristics barely affects the estimates (columns 2 and 3). On the contrary, the estimated coefficient on school VA is much lower, at 0.8 p.p. in column 1, and it decreases when including further controls (columns 2 and 3). Residential mobility, likely involving larger costs, is almost unrelated to peer quality and VA (Panel B).

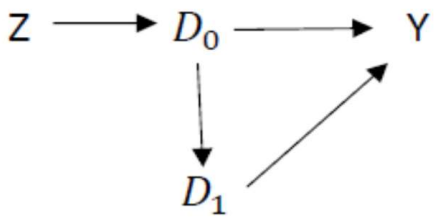
School mobility is more likely when students are assigned a school closer to the place of residence than their first choice. I estimate that the likelihood of moving to a different school increases by 1.2 p.p when a distance to school is 1km lower with respect to the first choice (Panel A of Table H.2). Although even smaller, the effect of distance to school on residential mobility is also statistically significant (Panel B). Findings suggest that parents willing to travel further distances to their first choice face particularly unsatisfactory schools at close proximity from residence, and are slightly more likely to move to different schools (or even local areas) when assigned to the latter.

Figure 6 suggests that students with offer from the most preferred schools are less likely to move to a different institution after the reception year. Table H.2 reports estimates from specifications analogue to Table 3 where school mobility indicator is the dependent variable. Panel A shows that, consistently across different specifications, students offered their first choice are 9 p.p. less likely to move to a different school with respect to peers offered their third choice or lower, decreasing to 5 p.p. for the second choice (mobility rate for students without offer from the two most preferred choices is approximately 0.2).

School mobility response to centralised assignment is unlikely to affect the interpretation of my results. Student achievement is observed at Year 2, the third year of primary education (see Section 3). Figure H.1 illustrates the relationship between student achievement (Y) and parental preference for the offered school (Z), the school where students initially enrols (D_0), and the school where student enrols at Year 2 (D_1) in a directed acyclic graph (DAG, Abadie and Cattaneo, 2018). The figure considers students with the same assignment risk so that school offer is good as randomly assigned (see Section 4) and affects student achievement

only through initial enrolment. In turn, initial enrolment may prompt parental response in the form of increased school mobility, impacting later enrolment. If school mobility has an impact on student achievement, the effect of initial enrolment combines two different channels: the direct impact of the school where the student initially enrols and the indirect impact of school mobility. A potential concern is that school mobility may constitute an alternative explanation of my results. I argue that my findings are not driven by school mobility for several reasons. First, almost no student moves to more-preferred schools than the one offered after the reception year (Figure 6, Panel C), implying that school mobility does not weaken the association of centralised offers with parental preference for the school where student enrols. Second, school mobility is unrelated to school VA (Table H.1, Panel A). Third, school mobility response to centralised offers is a parental decision. If, consistently with my results, this tends to improve the student-school match, my estimates would possibly be pushed downward.

Figure H.1: Offer, enrolment and outcome in a DAG



Note. The relationship between instrument, treatment and outcome in a directed acyclic graph. The graph includes parental preference rank for the school offered (Z), the school where student enrols at the reception year (D0), and the school where student enrols at Year 2 (D1); and student achievement (Y). See Appendix H for details.

Table H.1: School mobility and school attributes

	(1)	(2)	(3)
First choice VS student's school:			
	Panel A. Dep. Var.: School mobility		
Peer quality difference	0.0595***	0.0599***	0.0479***
	(0.0027)	(0.0032)	(0.0036)
School value added difference	0.0083***	0.0059**	0.0035
	(0.0026)	(0.0029)	(0.0030)
Distance difference	0.0125***	0.0126***	0.0124***
	(0.0011)	(0.0012)	(0.0012)
N	127,887	107,893	103,176
	Panel B. Dep. Var.: Residential mobility		
Peer quality difference	0.0001	0.0017	-0.0009
	(0.0020)	(0.0024)	(0.0028)
School value added difference	0.0038*	0.0014	0.0013
	(0.0020)	(0.0023)	(0.0024)
Distance difference	0.0061***	0.0060***	0.0060***
	(0.0009)	(0.0010)	(0.0010)
N	124,501	114,299	111,243
School choice controls		Y	Y
Individual characteristics			Y

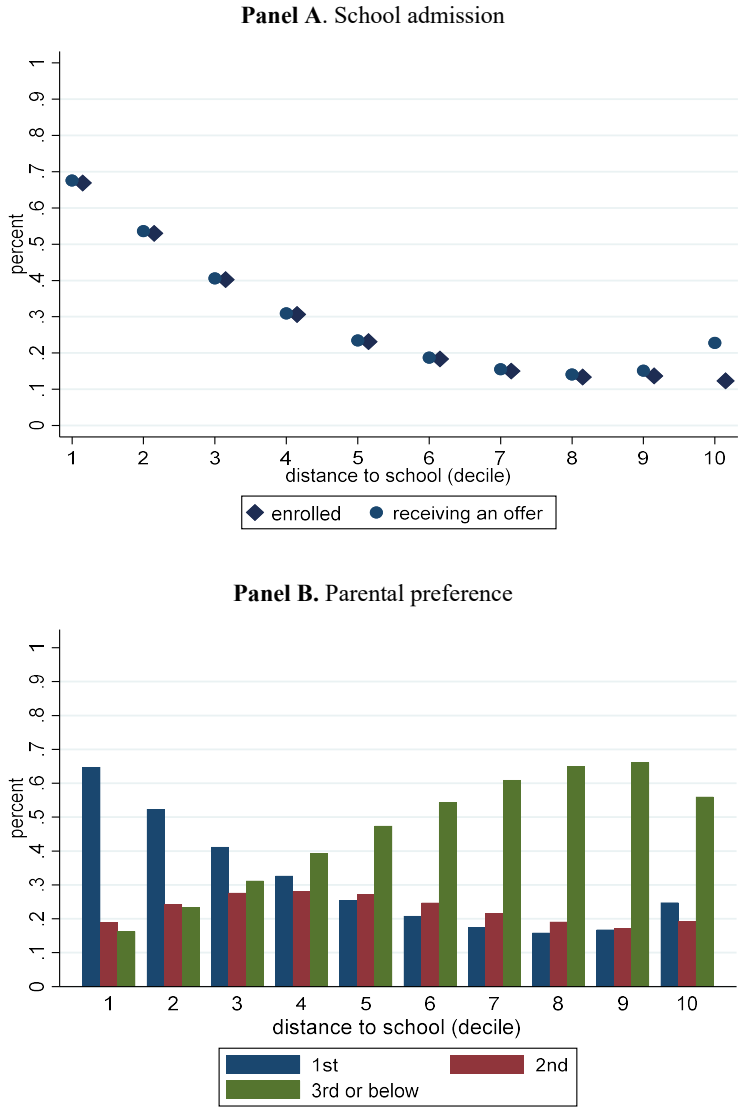
Note. The table shows correlation between school mobility and school attributes. The sample is restricted to first-choice applications. Panel A reports estimates from linear regressions of school mobility indicator, equal to one if a student moves to another school between reception year and Year 2. The dependent variable in Panel B is an indicator variable equal to 1 if a students moves residence (observed as home postcode). Independent variables are the differences between characteristics of the first choice and the school where student enrols. School value-added is estimated by regression-adjusted test scores growth at the school and averaged across subjects (see Appendix B). Peer quality is measured by school-level final year test scores, averaged across 2007-2014 cohorts and across mathematics and English. Distance, measured in kilometres and computed as linear distance between student postcode and school postcode centroids (schools farther than 2 kilometres from residence, the 90th percentile, are not considered).Control variables include year dummies. Column (2) adds dummies for n. of schools listed and ex-post feasibility of the school. Column (3) adds individual socioeconomic characteristics: gender, free lunch eligibility, special education needs, ethnicity, language, deprivation index and level of education in the area of residence. Robust standard errors are reported in parentheses. See Appendix G for details. *** p<0.01, ** p<0.05, * p<0.1

Table H.2: School mobility by offer status

	"preference" specification		"score" specification	
	(1)	(2)	(3)	(4)
		Panel A. Two-choice model		
Enroled in 1st choice	-0.0903***	-0.0906***	-0.0914***	-0.0911***
	(0.0139)	(0.0138)	(0.0125)	(0.0125)
Enroled in 2nd choice	-0.0570***	-0.0572***	-0.0506***	-0.0508***
	(0.0133)	(0.0133)	(0.0112)	(0.0112)
N	23,342	23,342	23,619	23,619
		Panel B. Three-choice model		
Enroled in 1st choice	-0.1150***	-0.1148***	-0.1053***	-0.1048***
	(0.0158)	(0.0158)	(0.0138)	(0.0138)
Enroled in 2nd choice	-0.0770***	-0.0768***	-0.0653***	-0.0653***
	(0.0157)	(0.0157)	(0.0128)	(0.0128)
Enroled in 3rd choice	-0.0602***	-0.0596***	-0.0391**	-0.0387**
	(0.0187)	(0.0187)	(0.0153)	(0.0153)
N	23,187	23,187	23,619	23,619
First and second choice FEs	Y	Y		
Tie-breaker controls	Y	Y	Y	Y
Assignment risk (ranked schools)	Y	Y		
Assignment risk (all schools)			Y	Y
Individual characteristics		Y		Y

The table shows the effect of school offer on school mobility. The dependent variable is an indicator equal to one if student moves to a different school from the reception year by the end of Year 2. Reported are estimated coefficients on first, second, or third choice offer from specifications analogue to Table 3. Robust standard errors are reported in parentheses. See Appendix G for details. *** p<0.01, ** p<0.05, * p<0.1

Figure I.1: School admission and distance to school



Note. The figure plots school offer and enrolment rates, and parental preference assigned to the school, by distance to residence. The sample includes all applicants to at least one London primary school in 2014 or 2015. Offer is reported by markers in Panel A, while diamonds represent enrolment measured at the reception year. Bars in Panel B represent the share of parents ranking the school first, second and third or below. Distance bins are deciles of within-school distribution of applicants. Outliers in the top 5% of the aggregate distance distribution are excluded. See Section 3 for details.

H Additional tables and figures

Table I.1: Balance tests (additional outcomes)

	Uncontrolled		Same-risk design		Within-school design	
	(1)	(2)	(3)	(4)	(5)	(6)
Demographics						
Free school meal eligible						
Offered 1st choice	-0.0404***	-0.0412***	-0.0117	-0.0191	-0.0107	-0.0177
	(0.0032)	(0.0039)	(0.0131)	(0.0142)	(0.0112)	(0.0122)
Offered 2nd choice	-0.0213***	-0.0220***	-0.0008	-0.0096	-0.0102	-0.0175
	(0.0043)	(0.0049)	(0.0122)	(0.0138)	(0.0105)	(0.0117)
Offered 3rd choice		-0.0067		-0.0239		-0.0195
		(0.0064)		(0.0165)		(0.0141)
Joint significance (p-value)	85.33 (0.000)	53.71 (0.000)	0.46 (0.630)	0.99 (0.397)	0.63 (0.531)	1.08 (0.356)
Special Education Needs						
Offered 1st choice	-0.0089***	-0.0121***	-0.0027	-0.0040	-0.0040	-0.0055*
	(0.0010)	(0.0013)	(0.0028)	(0.0032)	(0.0025)	(0.0029)
Offered 2nd choice	-0.0081***	-0.0115***	-0.0005	-0.0021	-0.0012	-0.0028
	(0.0013)	(0.0015)	(0.0033)	(0.0037)	(0.0027)	(0.0032)
Offered 3rd choice		-0.0112***		-0.0041		-0.0041
		(0.0018)		(0.0038)		(0.0034)
Joint significance (p-value)	36.67 (0.000)	28.48 (0.000)	0.52 (0.592)	0.67 (0.569)	1.42 (0.241)	1.31 (0.271)
White						
Offered 1st choice	0.0720***	0.0748***	-0.0131	-0.0208	0.0015	-0.0040
	(0.0040)	(0.0046)	(0.0147)	(0.0161)	(0.0124)	(0.0137)
Offered 2nd choice	0.0277***	0.0307***	-0.0347**	-0.0439***	-0.0231**	-0.0288**
	(0.0056)	(0.0061)	(0.0136)	(0.0154)	(0.0114)	(0.0128)
Offered 3rd choice		0.0203**		-0.0248		-0.0152
		(0.0082)		(0.0186)		(0.0152)
Joint significance (p-value)	188.48 (0.000)	119.91 (0.000)	3.33 (0.036)	2.75 (0.041)	2.74 (0.064)	2.16 (0.091)
Black						
Offered 1st choice	-0.0428***	-0.0472***	-0.0186*	-0.0194	0.0107	0.0095
	(0.0033)	(0.0039)	(0.0108)	(0.0119)	(0.0092)	(0.0102)
Offered 2nd choice	-0.0226***	-0.0268***	-0.0166*	-0.0177	-0.0006	-0.0019
	(0.0044)	(0.0049)	(0.0101)	(0.0114)	(0.0088)	(0.0100)
Offered 3rd choice		-0.0252***		-0.0028		-0.0033
		(0.0065)		(0.0136)		(0.0118)
Joint significance (p-value)	93.98 (0.000)	59.14 (0.000)	1.96 (0.140)	1.31 (0.269)	0.89 (0.412)	0.62 (0.605)
Deprivation in the local area						
Offered 1st choice	-0.0207***	-0.0166***	-0.0011	-0.0026	0.0103***	0.0102**
	(0.0011)	(0.0013)	(0.0040)	(0.0043)	(0.0038)	(0.0042)
Offered 2nd choice	-0.0038**	-0.0004	-0.0023	-0.0041	0.0011	0.0010
	(0.0015)	(0.0017)	(0.0037)	(0.0041)	(0.0034)	(0.0038)
Offered 3rd choice		0.0096***		-0.0049		-0.0002
		(0.0022)		(0.0051)		(0.0047)
Joint significance (p-value)	244.71 (0.000)	157.74 (0.000)	0.19 (0.824)	0.45 (0.720)	4.40 (0.012)	2.94 (0.032)
% with higher education in the local area						
Offered 1st choice	0.0206***	0.0195***	-0.0000	0.0004	-0.0038*	-0.0049**
	(0.0009)	(0.0010)	(0.0022)	(0.0024)	(0.0021)	(0.0024)
Offered 2nd choice	0.0105***	0.0099***	0.0014	0.0019	0.0013	0.0001
	(0.0011)	(0.0013)	(0.0020)	(0.0023)	(0.0019)	(0.0022)
Offered 3rd choice		0.0000		0.0014		-0.0031
		(0.0017)		(0.0027)		(0.0027)
Joint significance (p-value)	325.96 (0.000)	201.39 (0.000)	0.31 (0.732)	0.29 (0.834)	3.07 (0.046)	2.45 (0.061)
Baseline achievement						
All subjects						
Offered 1st choice	0.1508***	0.1769***	0.0463	0.0631	0.0503	0.0848**
	(0.0091)	(0.0106)	(0.0369)	(0.0400)	(0.0311)	(0.0339)
Offered 2nd choice	0.0799***	0.1082***	-0.0333	-0.0133	0.0438	0.0799**
	(0.0124)	(0.0137)	(0.0353)	(0.0398)	(0.0289)	(0.0323)
Offered 3rd choice		0.1118***		0.0542		0.0956**
		(0.0181)		(0.0466)		(0.0383)
Joint significance (p-value)	150.58 (0.000)	103.17 (0.000)	2.07 (0.126)	1.84 (0.137)	1.74 (0.175)	3.25 (0.021)
Mathematics						
Offered 1st choice	0.1189***	0.1401***	0.0237	0.0406	0.0525	0.0709**
	(0.0090)	(0.0105)	(0.0383)	(0.0417)	(0.0329)	(0.0358)
Offered 2nd choice	0.0564***	0.0801***	-0.0586	-0.0385	0.0250	0.0443
	(0.0124)	(0.0136)	(0.0367)	(0.0414)	(0.0303)	(0.0339)
Offered 3rd choice		0.0908***		0.0545		0.0511
		(0.0181)		(0.0478)		(0.0397)
Joint significance (p-value)	97.02 (0.003)	67.07 (0.000)	2.22 (0.109)	1.94 (0.121)	1.29 (0.276)	1.42 (0.235)
N						
Tie-breaker controls			Y	Y	Y	Y
Assignment risk (ranked schools)					Y	Y
Assignment risk (all schools)			Y	Y		
School where enroled FEs					Y	Y

Note. The table shows estimates of covariate balance by offer status. Estimates and specifications follow Table 1. Robust standard errors are reported in parentheses. See Section 4 for details. ***p<0.01. ** p<0.05. * p<0.1

Table I.2: Descriptive statistics (working sample)

	London	Rest of England	Difference (1 - 2)
	(1)	(2)	(3)
Baseline Characteristics			
FSM eligible	0.1517	0.1362	0.0155***
Not speaking English at home	0.4185	0.1199	0.2986***
White	0.4165	0.7765	-0.3600***
Asian	0.1946	0.0776	0.1170***
Black	0.1618	0.0223	0.1395***
Special education needs	0.0079	0.0064	0.0015***
Female	0.4900	0.4896	0.0004
Exceeding expectations at Year 0: mathematics	0.1330	0.1250	0.0080***
Achievement outcomes			
Exceeding expectations at Year 2: mathematics	0.2667	0.2145	0.0522***
Exceeding expectations at Year 2: reading	0.3023	0.2626	0.0397***
Exceeding expectations at Year 2: writing	0.2032	0.1602	0.0430***
School choice variables			
N. of schools listed	3.2069	--	
Ranked 1 choice	0.2684	0.3801	-0.1117***
Ranked at least 3 choices	0.5759	0.4343	0.1416***
Ranked 6 choices	0.2132	--	
Offered 1st choice	0.8283	0.8944	-0.0662***
Offered one of the top three choices	0.9422	0.9684	-0.0261***
Offered one of ranked choices	0.9684	--	
Enroled at offered school at reception year	0.8692	0.8975	-0.0283***
Not enroled at state schools at reception year	0.0372	0.0193	0.0179***
N	199,220	1,035,825	1,235,045

Note. This table shows descriptive statistics about applicants to any mainstream state-funded primary school in England (column 1) or to at least one primary school in Greater London (column 2) in 2014 and 2015. Columns (1) and (2) report averages computed using one observation per student, column (3) reports the mean difference between (1) and (2). All statistics are conditional on non-missing observations. See Section 3 for details. *** p<0.01, ** p<0.05, * p<0.1

Table I.3: School where student enrols VS other feasible schools

	N. of feasible schools	School value-added (standardised)			Peer quality (standardised)			Distance (km)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		% in best feasible school	Mean percentile rank in feasible set	Left "on the table"	% in best feasible school	Mean percentile rank in feasible set	Left "on the table"	% in best feasible school	Mean percentile rank in feasible set	Left "on the table"
All students	6.84	23.37	0.434	0.131	41.37	0.578	0.340	50.57	0.591	0.223
Deprivation in local area above median	8.63	17.02	0.450	0.189	30.97	0.576	0.434	45.27	0.553	0.253
Deprivation in local area below median	5.43	29.46	0.419	0.090	52.40	0.579	0.269	55.58	0.629	0.193

Note. The table compares the school where student enrols to other feasible schools. Column (1) reports the number of schools in the feasible set. Columns (2), (5), and (8) report the share of students enroling in the best feasible school according to a given attribute (the highest peer qualty or value-added, the lowest distance). Columns (3), (6), and (9) report the mean percentile rank of the school where student enrols in the feasible set, where feasible schools are ranked according to a given attribute (100 = school with highest peer qualty or value-added, or with the lowest distance). Columns (4), (7), and (10) report the difference between the school where student enrols and the best feasible school according to a given attribute. Columns (2)-(4) consider school value-added, estimated by regression-adjusted test scores growth at the school and averaged across subjects (see Appendix B). Columns (5)-(7) consider peer quality, measured by school-level final year test scores, averaged across 2007-2016 cohorts and across mathematics and English. Columns (8)-(10) consider distance, measured in kilometres and computed as linear distance between student postcode and school postcode centroids (schools farther than 2 kilometres from residence, the 90th percentile, are not considered). Reported are averages among all students, or separately for those residing in a LSOA with deprivation above or below median. See Section 6 for details.

Table I.4: Oversubscribed schools

	Popular schools	Not popular schools	Difference (1-2)
	(1)	(2)	(3)
Peer quality			
Sixth grade mathematics score	0.3313	-0.3182	0.6495***
Sixth grade reading score	0.3838	-0.3901	0.7738***
School quality			
School value added in mathematics	0.0652	-0.0949	0.1601***
School value added in reading	0.0815	-0.1209	0.2024***
School type			
Religious school	0.2178	0.1378	0.0799***
Academy school	0.1378	0.1739	-0.0361***
Community school	0.5405	0.6089	-0.0685***
Peer composition			
% FSM eligible students	0.1822	0.2789	-0.1057***
% white students	0.4813	0.3625	0.1188***
Income deprivation in student loca area (LSOA)	0.3140	0.4010	-0.0870***
N	2039	1386	3425
N (schools)	1180	865	2045

Note. This table shows characteristics of London primary schools by oversubscription status in 2014 and 2015. Columns (1) and (2) report average characteristics of oversubscribed and undersubscribed schools respectively, while mean difference is reported in column (3). Observations are at the school-year level. A school is coded as oversubscribed in a given year if applicants missing out on any higher-preference school exceed capacity by at least 5 seats. Peer quality is measured by school-level final year test scores across 2007-2016 cohorts, and standardised to have zero mean and unit variance. School value-added is estimated by regression-adjusted test scores growth at the school (see Appendix B). A school is defined as religious if it admits by faith. Peer composition variables are computed as average characteristic among a school's intake across grades 0-6 in 2014. Deprivation index is based on average income in the LSOA of residence. See Section 6 for details. ***p<0.01. ** p<0.05. * p<0.1