ISSN 2704-7407

UNIVERSITÀ CATTOLICA DEL SACRO CUORE
Dipartimento di Economia e Finanza

Working Paper Series

Parents Know Better: Sorting on Match Effects in Primary School

Marco Ovidi

Working Paper n. 121

November 2022

UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Parents Know Better: Sorting on Match Effects in Primary School

Marco Ovidi

Università Cattolica del Sacro Cuore

Working Paper n. 121
November 2022

Dipartimento di Economia e Finanza
Università Cattolica del Sacro Cuore
Largo Gemelli 1 - 20123 Milano – Italy
tel: +39.02.7234.2976 - fax: +39.02.7234.2781
e-mail: dip.economiaefinanza@unicatt.it

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

Marco Ovidi†

Università Cattolica Milano

November 2022

Abstract

I show that parents select schools by considering attributes of the student-school match that improve the learning outcomes beyond average school quality. Using the centralized algorithm for offers to primary school in London, I compare the achievement of students who are as good as randomly enroled in schools ranked differently in their application. Enroling at the most-preferred school versus an institution ranked lower increases achievement by 0.10 SD beyond school value-added among students with similar characteristics. 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 many 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 vary with 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 expected achievement gain from attending the school of choice beyond that school's average quality (or value-added, VA, the average causal impact across students).

Available studies mainly investigate match effects 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 effects on achievement 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 the expected school effectiveness as predicted from student's observable traits. I use administrative records from the centralised assignment of students to primary schools in London to isolate quasi-experimental variation in admission to the most preferred

[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., 2022), 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.

institutions. I leverage data on previous cohorts of students to estimate school VA, and investigate whether attending the school of choice has a causal impact on student learning trajectories beyond that school's VA.

The ideal experiment would compare the learning outcomes of students randomly enroled at 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 width of school catchments introduces variability in the admissions of students located close to the catchment boundary.

I use school offers generated by the centralised DA assignment to instrument school enrolment conditional on the assignment risk. Specifically, 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 design, I show that applicants with an offer from the most preferred schools are statistically undistinguishable from applicants with the same assignment risk and not receiving an offer from these schools.

The description of parental rankings of schools provides motivating evidence for my analysis. Preferences for schools, on average, respond to peer quality and proximity to residence. 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. However, parental rankings exhibit remarkable variability conditional on distance and peer quality. For example, I show that about 40% of parents rank the same school at least one position away from the average in their 100-meter distance cell. While parental rankings may be noisy measures of their preferences, I show that parents avoid schools ranked lower than the one offered even three years after the application, suggesting that they represent solid and robust preferences for schools.[3]

Enroling at the most preferred schools increases student learning. Enrolment at the most-preferred school, as opposed to enrolment at a school ranked second or lower, has a causal effect on Year 2 math scores of 0.10 standard deviations (hereafter, σ). Similarly, students not receiving an offer from the most preferred school but enroled at the second most preferred school have higher Year 2 math scores compared to what they would have scored at the school ranked third or lower. 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 VA at that school.

Only a small part of these match effects is explained by student characteristics considered in previous studies. Following (Abdulkadiroglu et al., 2020), I estimate school VA as function of student characteristics as the persistent effect of a school for a specific type of student. Specifically, I consider cells of covariates which are most predictive of academic achievement (gender, socioeconomic status, and baseline ability) and estimate school value-added on students in each cell. 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.

Match effects are more pronounced among parents from relatively more advantaged backgrounds. Students not eligible for subsidised lunch, living in local areas with above-median education, 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.

One concern with the interpretation of my results as match effects is that estimated VA does not fully capture the expected achievement growth at the school. I show that average school quality is credibly accounted for in my comparison. Specifically, I first show that estimated school VA strongly predicts student achievement. Second, I find that enroling at a more-preferred school is unrelated to school VA growth, implying that sorting into schools

[3] 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., 2022). I describe parental preferences 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.

which are improving their quality over time does not explain my results. Finally, I implement an alternative design to hold any school input constant. In this design, 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.

My results provide evidence of parents' sorting on match effects, implying that parental choice may increase allocative efficiency.[4] Abdulkadiroglu et al. (2020) find positive selection on gains but very small in magnitude, and show that parental preferences do not respond to match effects based on student's characteristics. My causal parameters are identified for a policy-relevant group of students, which are those with uncertain admission outcome and complying with offers from the centralised mechanism. These students would be admitted to their preferred schools if capacity at these schools was expanded based on parental rankings. The focus on such margin rather than the full student population is one important difference with Abdulkadiroglu et al. (2020), which may concur to explain why I reach different conclusions. Institutional differences between my context (London primary schools) and their context (New York City high schools) may additionally play a role. Primary schools in London are small and assignment by distance implies relatively narrow choice sets, which are possibly easier to navigate for parents. Smaller and closer schools to choose from may facilitate the acquisition of information on their suitability for children's specific learning needs. Besides, centralised assignment regulation leaves little scope for students' screening, and may therefore induce schools to compete on (vertical and horizontal) school quality instead. Moreover, I view my results as complementary to findings in Campos and Kearns (2022). Differently from their work, 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 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.[5]

A broad body of literature has investigated the impacts of attending the schools that

[4] 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.

[5] 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 persist. Consistently, I find the average VA across schools to be roughly constant throughout the period I consider (Figure D.2).

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 with this body of research, I isolate student-school match effects from returns driven by average school quality. Methodologically, the studies above consider admission cutoffs in isolation without fully exploiting school offer variation induced by centralised assignment (Abdulkadiroglu et al., 2022). Moreover, my study is the first to examine parental choice effects in primary schools, filling an important gap in the literature. Early school years are crucial for child's development (e.g., Chetty et al., 2011; Heckman et al., 2013) and, differently from later stages, the choice of primary school depends entirely on parents.

2 Primary school choice in London

Data

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 (age 7). 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.

I exploit administrative data on applicants to state-funded primary schools in London in 2014 and 2015. Records include rank-order lists of schools submitted by parents to their school district (local authority, hereafter, LA) and the school offered to each applicant. 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, a granular information on residential location spanning an average of 17 addresses in England and often corresponding to a single building in London. I compute the linear distance from each applicant's postcode to all ranked schools.[6] In

[6] Centroid coordinates for English postcodes are obtained from www.doogal.co.uk. For applicants without a

addition, I use NPD records since 2006 to observe schools' academic performance.

KS1 assessments in 2017 and 2018 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.[7] Three different subjects are assessed – English, reading and writing, and mathematics. 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 at the end of KS1, with an optional writing test, which scores are not disclosed since 2016 but are meant to inform teacher assessments. Using administrative records from seven previous cohorts (2009-2015), I show in Figure H.1 a very strong agreement between test scores and teacher assessments, suggesting that the latter are a reliable measure of achievement (in line with findings by Burgess and Greaves, 2013).[8] For both English and mathematics, the correlation coefficient between KS1 test scores and teacher-awarded level is 0.94.

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 at a student's primary school.[9] Similar to KS1 assessments, EYFSP results are grouped into three categories indicating whether students achieve below, at, or above expected standards. Moreover, I observe a number of student characteristics 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.

I observe 199,180 applicants to at least one of the 1,739 primary schools in London. Most

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.

[7]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).

[8]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.

[9]Using these assessments as baseline is a potential limitation since they are administered during the first year of primary school and may therefore incorporate some school effect. EYFSP assessments are intended as a baseline measure of a student's starting point in formal education. Nevertheless, I acknowledge this potential drawback by additionally discussing results not controlling for this measure in Section 5 below.

parents make use of school choice, with the average student listing around 3 schools and the 58% of parents listing at least three institutions (Table H.1). 83% of students are offered their first choice, missing out on preferred schools substantially more frequently than in the rest of the country (the national average is 89%). Almost all students are offered one of the listed schools (97%). A high proportion of parents (86%) comply with the centralised school offer, and only 4% enrol at a private institution in the reception year (they are 7% by Year 2).

Institutional setting and parental preferences

Parental choice among state-funded schools is an established feature of the school market. Public-sector schools are the main provider of primary education, with less than 5% of students opting for private institutions.[10] 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, inside or outside their district, and to submit their list to the LA of residence. LAs 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.[11]

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 and on intake composition of each state-funded school. Institutions with excellent test scores are typically sought after by parents, as can be seen in application data. Panel A of Figure 1 plots the standardised final test scores at the school by parental preference rank separately for students residing in areas with deprivation above or below the median.[12] Regardless of socioeconomic status, peer quality markedly increases with parental preference.[13]

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

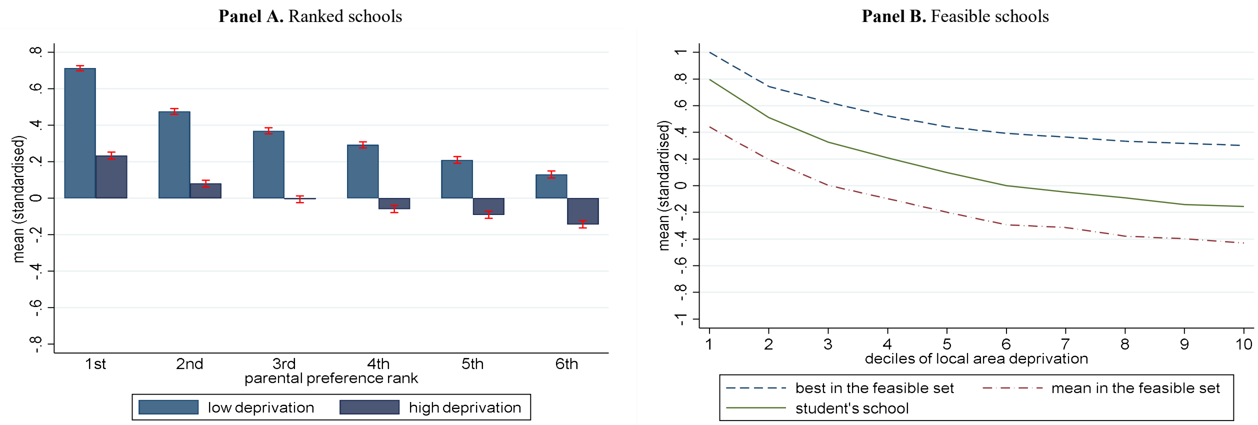
[10] Author's own calculation from official 2019 data on students count by school phase and sector (link).

[11] Primary schools have a statutory class size cap of 30 students.

[12] I plot average parental preference rank conditional on the number of schools listed and other controls from equation (10), see Appendix A for details. I measure peer quality using KS2 test scores from 2006 to 2016 averaged across mathematics and reading.

[13] The properties of the assignment algorithm imply that the ranking among two schools listed by parents at application shall reflect their preference order (see Section 3).

Figure 1: Parental preferences and peer quality

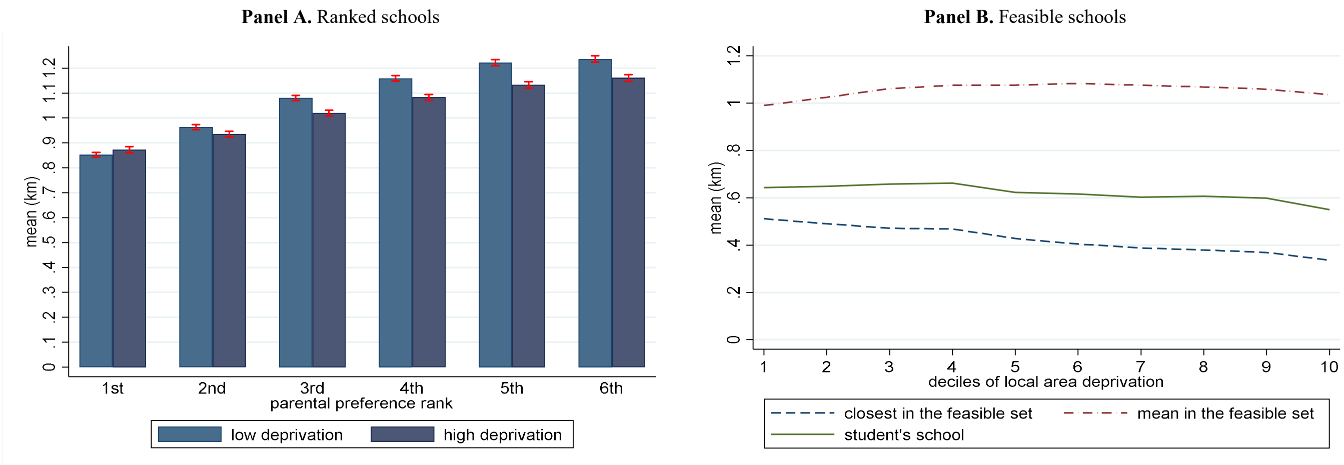


Note. This figure plots parental preference rank against peer quality (Panel A) and peer quality at the school where the student enrols compared to average or highest peer quality in the feasible choice set (Panel B). Bars in Panel A plot predicted values from equation (10), where preference rank dummies 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 2 and Appendix A for details.

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.

Residential sorting implies that the access to high-performing institutions is spatially segregated. Schools ranked by parents in high-deprivation areas exhibit remarkably lower peer quality than those ranked by better-off parents at each preference rank (Panel A of Figure 1). Most of this gap is explained by differences in peer quality of school that parents could access based on distance. Following Ainsworth et al. (2022), Panel B compares the school where a student enrols with other feasible institutions by decile of local area deprivation. Peer quality at feasible schools is 0.8σ higher in areas with deprivation in the bottom decile compared to 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 A.1). This gap may reflect steeper trade-offs between school test scores and distance for disadvantaged parents.

Figure 2: Parental preferences and distance to school



Note. This 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 2 and Appendix A for details.

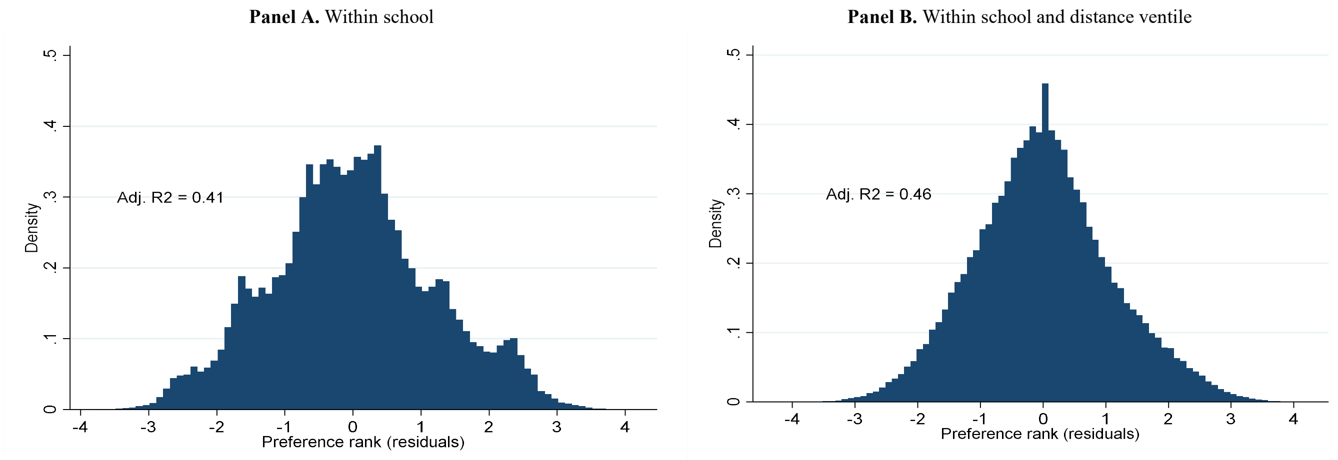
Parents nearly minimise distance among available options. Students with different socioeconomic statuses travel very similar distances to primary school, approximately 600 metres (Panel B of Figure 2).[14] Parents in areas with higher deprivation face schools closer to the place of residence, likely reflecting higher population density. The 51% of applicants enrol at the closest accessible institution (see columns 8-10 of Table A.1). 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.

Parental preferences lead to rationing of seats at schools with high peer quality. Attributes of oversubscribed schools are described in Table H.3. I define a school as oversubscribed if the number of applicants who are not offered a place at a more preferred institutions exceeds capacity. Schools oversubscribed by at least 5 seats are the 61% of institutions in 2014 and the 58% in 2016.[15] Mirroring evidence from preference data, oversubscribed schools have markedly higher peer quality than institutions with spare capacity, 0.65σ in mathematics and 0.77σ in English.

[14] This value differs from the distance values in Panel A since the latter are regression-adjusted estimates.

[15] The fraction of oversubscribed schools is 72% in 2014 and 66% in 2015. In both years, a remarkable share of schools (45%) is oversubscribed by at least 20 seats. School capacity is proxied using school offers (see Section 3).

Figure 3: Variability in parental preferences



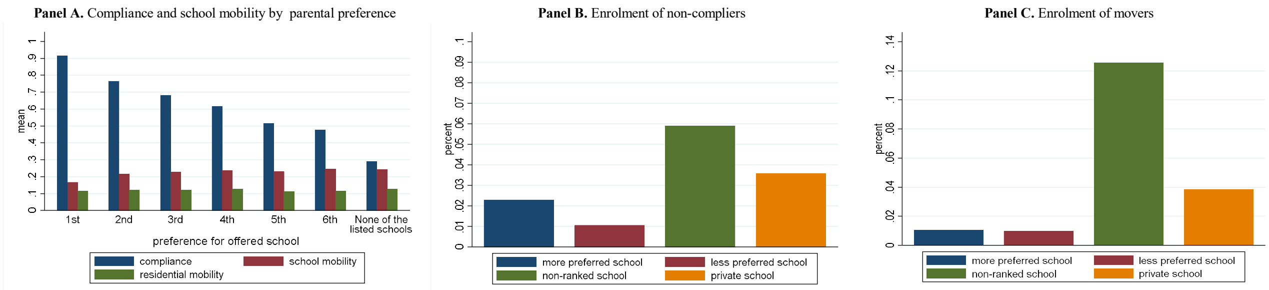
Note. This figure plots residual parental preference rank after controlling for dummies indicating the number of schools listed, ex-post feasibility of the school, and school dummies (Panel A); or additionally including school-by-distance-bin dummies, with 100-meter wide distance bins (Panel B). The 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 2 for details.

Motivating evidence

Average patterns mask substantial disagreement in parental ranking of schools. Panel A of Figure 3 reports the distribution of residual preference ranks from a regression on school dummies, ex-post feasibility, and number of ranked schools. As reported in the figure, the adjusted $\mathbb{R}^2$ index is just 0.41, showing that parental rankings vary not only conditional on peer quality, but on any other school input that is common across students. One obvious explanation is that parents may reside at different distances from the school. Controlling for school-by-distance-cell dummies increases only slightly the explained variability (Panel B, adjusted $\mathbb{R}^2$=0.46). The figure shows that about 40% of parents rank a given school at least one position away from the average in their 100-meter distance cell. Overall, I find remarkable variability in parental rankings of the same school among parents residing at a similar distance. Therefore, a substantial share of variation in parental preferences is explained by unobserved attributes that are specific to the student-school match.

Nevertheless, subsequent parental choice behaviour is consistent with preferences submitted at the time of application. First, an offer from a more-preferred school increases the likelihood of compliance with the centralised assignment. This can be seen in Panel A of Figure 4, 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

Figure 4: Compliance with assignment and mobility



Note. This 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 2 for details.

for the offered school decreases. Less than 1% of applicants enrol at a school with lower preference than the one assigned, consistently with the rankings submitted at application (see Panel B). Second, the likelihood of moving children to a different institution by Year 2 decreases with the preference for the offered school (Panel A of Figure 4, red bars).[16] Only 1% of students move to a school ranked with a lower preference than the one where they initially enrol, suggesting that parents consistently avoid schools ranked with lower preference at application even after initial enrolment (see Panel C). 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 B 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.

Overall, although parents generally prefer schools with higher peer quality and that are closer to the place of residence, their rankings exhibit substantial variability and yet they strongly predict school choice behaviour in subsequent years. Therefore, parental preferences appear to reflect solid tastes for available schools that are to a large extent student-specific. I investigate in the next sections whether such heterogeneity is linked to attributes of the student-school match that increase student achievement.

[16] Residential mobility, in contrast, is approximately orthogonal to school assignment (green bars). 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.

3 The centralised school assignment algorithm

Institutional setting

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. A small number of children with exceptionally disadvantaged backgrounds are guaranteed admission at their preferred school.[17] Applicants with siblings currently enroled at the school are also 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 after the previously popular "Boston" mechanism was banned.[18] 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).

In particular, 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 are 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

[17] 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.

[18] 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.

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.[19]

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 H.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. 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).

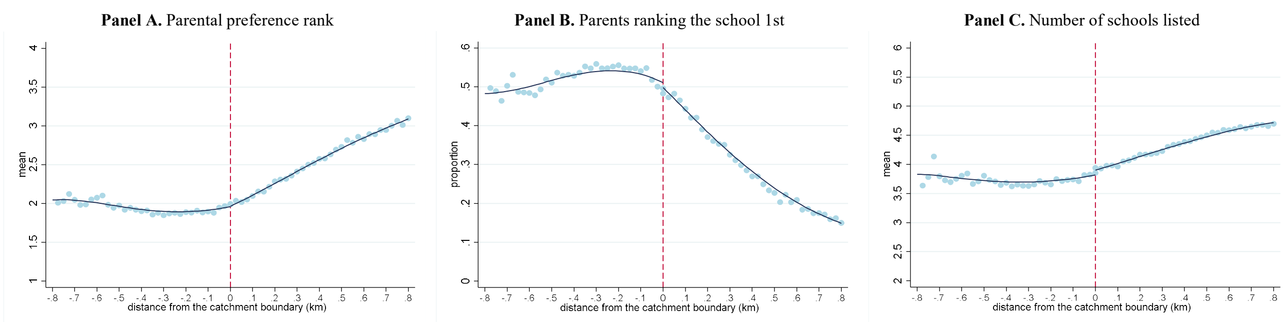
I replicate the assignment mechanism to trace catchment boundaries and identify applicants at the margin for admission.[20] The school catchment boundary is defined as the distance of the last admitted applicant. Replication is complicated by data availability since I have no information on demographics determining admission priorities – most importantly, whether a student has siblings at the school.[21] 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 assign-

[19] 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.

[20] 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.

[21] In practice, most school districts break ties by linear distance. I show in Table G.2 that my results are robust to the exclusion of LAs breaking ties by walking distance instead.

Figure 5: 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.

ments. Second, I adjust offer replication to residual unobserved priorities.[22] Intuitively, if an applicant with offer resides beyond the initially estimated distance threshold, she must have priority. In Appendix C, I discuss how I achieve replication of school offers based on these ideas.[23]

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 (see Panel B of Figure C.2), 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 5 shows that parental preference for the school (Panel A, a value of 1 denotes first choice), the share of parents ranking the

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

[23]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.

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 with distance to school (consistently with Figure 2). The decrease accelerates starting at a slightly shorter distance than the catchment boundary, suggesting that parents may 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.[24]

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 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:

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

[24]Figure C.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.

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 D for details). By using the estimated value of α_s in equation (3), I rely on the assumption that this estimate is 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 right-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.[25] 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 E.[26]

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 supported by first-stage results in Section 5.

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

[25] 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.

[26] 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 ($\mathbb{R}^2 = 0.87$).

in Appendix F, 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 was provided in Section 2. In particular, Panel B of Figure 4 shows that only 1% of students enrol at a school ranked with lower preference than the one offered.[27]

Under these assumptions, I identify a local average treatment effect (LATE) of attending a school ranked with a higher preference. Specifically, consider the comparison between students enroled in the first-choice school relative to a lower-ranked school, where D_{i1} and Z_{i1} indicate enrolment at and offer from the first choice, respectively, and let $\tilde{Y}_i \equiv Y_i - \hat{\alpha}_{s(i)}$ be the VA-adjusted achievement of student i enroled 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.[28] I show in Appendix F 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 estimation sample (Table E.1).[29] The largest 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:

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

[27] 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 B.

[28] 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 F). Similar comparisons can be defined for schools ranked lower than the first choice.

[29] I consider in estimation the 172,099 students for which I observe KS1 and EYFSP assessments and individual characteristics.

Table 1: Balance tests

	Uncontrolled			Same-risk design			Within-school design		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Predicted outcome									
Mathematics									
Offered 1st choice	0.0822***	0.0989***	0.1146***	0.0360*	0.0233	0.0379	0.0106	0.0205	0.0472**
	(0.0043)	(0.0055)	(0.0065)	(0.0210)	(0.0230)	(0.0250)	(0.0178)	(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)	365.98 (0.000)	171.73 (0.000)	115.86 (0.000)	2.94 (0.086)	2.30 (0.100)	2.43 (0.063)	0.35 (0.552)	0.81 (0.443)	3.61 (0.013)
English									
Offered 1st choice	0.0858***	0.1026***	0.1185***	0.0454**	0.0339	0.0475*	0.0151	0.0284	0.0577**
	(0.0046)	(0.0060)	(0.0070)	(0.0230)	(0.0251)	(0.0274)	(0.0195)	(0.0216)	(0.0236)
Offered 2nd choice		0.0573***	0.0746***		-0.0260	-0.0098		0.0275	0.0582***
		(0.0082)	(0.0090)		(0.0238)	(0.0269)		(0.0200)	(0.0225)
Offered 3rd choice			0.0683***			0.0441			0.0813***
			(0.0119)			(0.0312)			(0.0263)
Joint significance (p-value)	341.03 (0.000)	159.03 (0.000)	106.98 (0.000)	3.89 (0.0486)	2.51 (0.0816)	2.32 (0.073)	0.60 (0.439)	1.27 (0.280)	3.98 (0.008)
N	172,099	172,099	172,099	22,628	22,628	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

where W_i is a baseline characteristic of student i. I present empirical models including offers from the top one, two, or three ranked schools ($C = 1, 2, 3$). For example, when $C = 2$, 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. 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.[30] 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) of Table 1 shows that students offered their first choice have about 0.08σ higher predicted achievement in mathematics than students offered an institution ranked with lower preferences, with similar results for achievement in English. When compared to students offered their third choice or lower, peers with an offer from the first or second choice have 0.1σ or 0.06σ higher predicted achievement (column 2), respectively, and similar results hold when including offers from a student's third choice (column 3). Controlling for assignment risk, these differences are substantially smaller and mostly undistinguishable from zero (columns 4-6). F-tests of the joint significance of offer dummies fail to reject the null hypothesis at the 5% level, in sharp contrast with the uncontrolled specification. Moreover, I show in Table H.2 that controlling for assignment balances single covariates used to predict achievement.

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

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

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

[30]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 E for details on bandwidth selection. All controls are interacted with cohort dummies.

where risk controls, $\boldsymbol{p_i}$ and f^d are defined analogously to equation (7). In equation (9), the parameters β_r estimate the impact of enroling at a student's r-th choice on VA-adjusted achievement relative to a school ranked lower.

5 Results

First stage

Receiving an offer from one of the most-preferred schools increases steeply the probability of enroling 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 second or lower are 71 p.p. more likely to enrol at that school (Panel A, column 1), decreasing to 67 p.p. when compared to those offered a school ranked third or lower (Panel B). The corresponding estimate for the second choice is 77 p.p. (column 2). Similar results are obtained from a model including offers from a student's third choice (Panel C).

First-stage results suggest that offer take-up responds to parental preference for the school. I find a smaller first stage impact of an offer from a student's first choice compared to the second choice (Panels B and C), likely reflecting the role of waiting lists at the most preferred institutions. This is consistent with Panel B of Figure C.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 with respect to an offer from the third choice or lower) are negative, implying that an offer from a more-preferred school increases the likelihood of compliance with the centralised assignment. This is consistent with Panel A of Figure 4 (see Section 2).

Match effects

I start by documenting that enroling at a more-preferred school increases student's achievement in mathematics. This can be seen in Table 3, reporting estimates of the β_z coefficients from equation (9) when considering offers from a student's top one (Panel A), two (Panel B), or three (Panel C) choices. Dependent variable is a student's KS1 score, implying that estimates conflate match effects and the impact of school value-added. Enroling at the first

Table 2: First stage results

	Same-risk design			Within-school des	
	Enroled at 1st choice	Enroled at 2nd choice	Enroled at 3rd choice	Enroled at 1st choice	Enroled at 2nd choice
	(1)	(2)	(3)	(4)	(5)
			Panel A. One-choice model		
Offered 1st choice	0.7081***			0.6256***	
	(0.0116)			(0.0106)	
F-test	3742.28			3478.96	
			Panel B. Two-choice model		
Offered 1st choice	0.6720***	-0.0301***		0.5732***	-0.0233***
	(0.0134)	(0.0084)		(0.0120)	(0.0085)
Offered 2nd choice	-0.0817***	0.7710***		-0.1087***	0.7319***
	(0.0112)	(0.0100)		(0.0094)	(0.0093)
F-test	2133.78	3354.25		2007.59	3474.48
			Panel C. Three-choice model		
Offered 1st choice	0.6467***	-0.0426***	-0.0204***	0.5289***	-0.0368***
	(0.0149)	(0.0092)	(0.0064)	(0.0132)	(0.0098)
Offered 2nd choice	-0.1126***	0.7561***	-0.0341***	-0.1550***	0.7178***
	(0.0134)	(0.0111)	(0.0063)	(0.0112)	(0.0106)
Offered 3rd choice	-0.0822***	-0.0404***	0.7198***	-0.1228***	-0.0374***
	(0.0157)	(0.0116)	(0.0149)	(0.0129)	(0.0104)
F-test	1484.20	2265.46	948.79	1401.08	2361.57
N	22,628	22,628	22,628	22,628	22,628
Tie-breaker controls	Y	Y	Y	Y	Y
Assignment risk (ranked schools)				Y	Y
Assignment risk (all schools)	Y	Y	Y		
School where enroled FEs				Y	Y
Individual characteristics	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 (8). Panels A, B, and C consider offer from the top, the top two, and the top three ranked schools, respectively. Controls in include a local linear polynomial of distance to the catchment boundary at each ranked school, and dummies for assignment ri schools in London. Columns (4)-(6) include a local linear polynomial of distance to the catchment boundary and assignm ranked school, and additionally control for student's school dummies. When the number of listed schools is less than six, th control variables are set to 0, and dummies indicating missing preferences are included in the controls. All regressions contr ethnicity, FSM eligibility, SEN, gender, deprivation index and education in the area of residence. Reported are F-tests of join school offer indicators. Robust standard errors are reported in parentheses. See Section 5 for details. ***p<0.01. ** p<0.05. * p

Table 3: Total achievement effects of enroling at most-preferred schools

	KS1 score in mathematics				KS1 score in English			
	Uncontrolled	Same-risk design			Uncontrolled	Same-risk design		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A. One-choice model				
Enroled in 1st choice	0.0758***	0.1161**	0.1145**	0.0647	0.0764***	0.0763	0.0647	0.0103
	(0.0070)	(0.0504)	(0.0498)	(0.0404)	(0.0070)	(0.0505)	(0.0491)	(0.0381)
				Panel B. Two-choice model				
Enroled in 1st choice	0.0949***	0.1531**	0.1533***	0.1197**	0.0999***	0.1072*	0.0933	0.0566
	(0.0089)	(0.0597)	(0.0590)	(0.0479)	(0.0089)	(0.0600)	(0.0585)	(0.0454)
Enroled in 2nd choice	0.0683***	0.0707	0.0740	0.1048**	0.0788***	0.0591	0.0546	0.0882**
	(0.0125)	(0.0514)	(0.0506)	(0.0409)	(0.0125)	(0.0515)	(0.0503)	(0.0390)
				Panel C. Three-choice model				
Enroled in 1st choice	0.1169***	0.1611**	0.1530**	0.1000*	0.1253***	0.1404**	0.1197*	0.0631
	(0.0103)	(0.0689)	(0.0680)	(0.0552)	(0.0103)	(0.0699)	(0.0681)	(0.0525)
Enroled in 2nd choice	0.0912***	0.0797	0.0737	0.0826*	0.1057***	0.0963	0.0842	0.0955**
	(0.0136)	(0.0633)	(0.0623)	(0.0501)	(0.0136)	(0.0641)	(0.0626)	(0.0481)
Enroled in 3rd choice	0.0819***	0.0210	-0.0007	-0.0517	0.0975***	0.0871	0.0691	0.0169
	(0.0182)	(0.0753)	(0.0741)	(0.0601)	(0.0182)	(0.0758)	(0.0736)	(0.0576)
N	172,099	22,628	22,628	22,628	172,099	22,628	22,628	22,628
Tie-breaker controls		Y	Y	Y		Y	Y	Y
Assignment risk (all schools)		Y	Y	Y		Y	Y	Y
Individual characteristics			Y	Y			Y	Y
Baseline achievement				Y				Y

Note. The table shows estimates of match effects at the most preferred schools from equation (9). Dependent variables are KS1 score in mathematics (columns 1-4) or English (columns 5-8). Colums (1) and (5) consider all applicants, columns (2)-(4) and (6)-(8) restrict the sample to students with nondetermistic assignment risk at one or more ranked schools. Controls in columns (2)-(4) and (6)-(8) 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. Panels A, B, and C consider enrolment at the top, the top two, and the top three ranked schools, respectively. 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 year dummies. Columns (1) and (5) include dummies for the top (Panel A), the top two (Panel B), or the top three ranked schools (Panel C). Columns (3) and (7) control for gender, language, ethnicity, free school meal eligibility, special education needs, deprivation index and education level in the neighbourhood of residence. Columns (4) and (8) additionally include dummies for baseline student achievement. Robust standard errors are reported in parentheses. See Section 5 and Section 6 for details. ***p<0.01. ** p<0.05. * p<0.1

choice increases achievement in mathematics by 0.12σ compared to a school ranked second or lower (Panel A, column 2). When controlling for student's characteristics, the result is virtually unchanged (column 3). The impact of enroling at the first choice increases at 0.15σ or 0.16σ when compared, respectively, to a school ranked third or lower (Panel B) or fourth or lower (Panel C). The impact of enroling at the second choice with respect to a school ranked lower is about 0.07σ, although imprecisely estimated. The impact of enroling at a student's third choice is small in magnitude and not statistically different from zero. Results for achievement in English in columns (6)-(7) are slighlty lower and often not statistically different from zero.

Estimates in columns (4) and (8) additionally control for a student's baseline achievement. On the one hand, estimates are lower but fairly robust to this inclusion. On the other hand, as detailed in Section 2, these assessments are administered at the primary school of enrolment during the reception year and may therefore partly reflect school effects. The fact that the impacts of enroling at the first or second choice are substantially more similar when controlling for baseline achievement is in line with this hypothesis. In what follows, I show both results from specifications including baseline achievement or just student's demographics.

I next isolate the match effects of enroling at a more-preferred school 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 D for details), and subtract estimated school VA from student's achievement to define $\tilde{Y}_i$ used in equation (3).[31] Columns (1)-(2) of Table 4 show that, when using VA-adjusted achievement in mathematics in equation (9), the impact of enroling at most-preferred schools is very similar to the unadjusted estimates. For example, the effect of enroling at a student's first choice with respect to a school ranked third or lower (Panel B) is $0.11 - 0.14\sigma$ compared to $0.12 - 0.15\sigma$ in Table 3. Consistently with these results, enroling at one's first choice against a school ranked third or lower slightly increases school VA (about 0.01σ, columns 1-2 of Table D.5). Moreover, I show in columns (3)-(4) of Table D.5 that enroling at more-preferred schools has no impact on VA growth at a student's school, implying that parental sorting into schools that are improving their quality over time does not drive my results.

[31] I show in Table D.4 that estimated school VA strongly predicts student's achievement (see Appendix D for details).

Table 4: Match effects of enroling at most-preferred schools

	Mathematics						English					
	Same-risk design (dep. var.: VA-adjusted KS1 score)		Same-risk design (dep. var. match-adjusted KS1 score)		Within- school design (dep. var. KS1 score)		Same-risk design (dep. var.: VA-adjusted KS1 score)		Same-risk design (dep. var. match-adjusted KS1 score)		Within- school design (dep. var. KS1 score)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
						Panel A. One-choice model						
Enroled in 1st choice	0.1048**	0.0550	0.1000**	0.0507	0.0941**	0.0734*	0.0542	-0.0001	0.0491	-0.0043	0.0317	0.0089
	(0.0498)	(0.0406)	(0.0496)	(0.0407)	(0.0478)	(0.0380)	(0.0493)	(0.0385)	(0.0490)	(0.0387)	(0.0476)	(0.0365)
						Panel B. Two-choice model						
Enroled in 1st choice	0.1409**	0.1073**	0.1320**	0.0985**	0.1698***	0.1260***	0.0809	0.0444	0.0722	0.0361	0.1083*	0.0611
	(0.0590)	(0.0481)	(0.0587)	(0.0482)	(0.0590)	(0.0472)	(0.0586)	(0.0459)	(0.0583)	(0.0462)	(0.0590)	(0.0454)
Enroled in 2nd choice	0.0688	0.0996**	0.0611	0.0910**	0.1257***	0.0873**	0.0510	0.0847**	0.0442	0.0768*	0.1272***	0.0868**
	(0.0506)	(0.0409)	(0.0505)	(0.0411)	(0.0472)	(0.0380)	(0.0504)	(0.0392)	(0.0502)	(0.0393)	(0.0471)	(0.0361)
						Panel C. Three-choice model						
Enroled in 1st choice	0.1318*	0.0788	0.1225*	0.0694	0.2331***	0.1325**	0.0970	0.0407	0.0883	0.0321	0.1874**	0.0814
	(0.0681)	(0.0555)	(0.0677)	(0.0556)	(0.0731)	(0.0582)	(0.0683)	(0.0531)	(0.0678)	(0.0533)	(0.0730)	(0.0556)
Enroled in 2nd choice	0.0586	0.0677	0.0505	0.0584	0.1878***	0.0937*	0.0690	0.0805*	0.0622	0.0723	0.2048***	0.1067**
	(0.0624)	(0.0502)	(0.0622)	(0.0504)	(0.0611)	(0.0486)	(0.0628)	(0.0483)	(0.0623)	(0.0484)	(0.0611)	(0.0460)
Enroled in 3rd choice	-0.0239	-0.0746	-0.0248	-0.0761	0.1337*	0.0138	0.0419	-0.0099	0.0421	-0.0104	0.1672**	0.0429
	(0.0744)	(0.0606)	(0.0740)	(0.0606)	(0.0687)	(0.0554)	(0.0737)	(0.0578)	(0.0734)	(0.0581)	(0.0689)	(0.0526)
N	22628	22628	22628	22628	22628	22628	22628	22628	22628	22628	22628	22628
Tie-breaker controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Assignment risk (all schools)	Y	Y	Y	Y			Y	Y	Y	Y		
Assignment risk (ranked schools)					Y	Y					Y	Y
School where enroled FEs					Y	Y					Y	Y
Individual characteristics	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Baseline achievement		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 1-6) or English (columns 7-12). Columns (1)-(2) and (7)-(8) report estimates of VA-adjusted KS1 score, obtained subtracting estimated school VA from student's score. Columns (3)-(4) and (9)-(10) report estimates of match-adjusted KS1 score, obtained subtracting estimated school match effects based on student's characteristics from student's score. Controls include a local linear polynomial of distance to the catchment boundary at each ranked school, and dummies for assignment risk at all primary schools in London. Columns (5)-(6) and (11)-(12) 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 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. Even columns additionally include baseline student achievement. Robust standard errors are reported in parentheses. See Section 5 for details on estimation, and Appendix D for details on estimated school VA and school match effects. ***p<0.01. ** p<0.05. * p<0.1

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 (Abdulkadiroglu et al., 2020; see Appendix D 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, columns 5-6 of Table D.5), implying a small degree of parental sorting into schools based on these traits. Table 4 shows that, after accounting for this increase, estimated impacts of enroling at most-preferred schools are only slightly smaller than the unadjusted estimates in Table 3. For example, enrolment at one's first-choice school improves student achievement in mathematics by 0.10-0.13σ with respect to a school ranked third or lower (columns 3-4, Panel B of Table 4).

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 effects based on student's characteristics with respect to her second or third choice, respectively. Table 5 reports estimates from equation (9) augmented with interactions between erolment dummies and match gain from enroling at one's first or second choice. Estimates of uninteracted enrolment dummies in column (1) imply a match effect from enroling in the first or second choice on student achievement in mathematics of about 0.10σ, similar to the main results. In column (2), I include interactions with dummies for a positive or negative match gain. Estimates of the interaction between enrolment at one's first or second choice and indicators for 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).

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-

Table 5: Match effects and VA gains

	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

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 Figure 3) and estimate empirical models similar to equation (9) augmented with a full set of dummies for the school where the student enrols.[32] 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 support the validity of my alternative design. Columns (7)-(8) of Table 1 show that students enroled at the same school which they ranked with different preferences have similar predicted achievement based on their predetermined characteristics. This conclusion holds when considering single covariates in columns (7)-(8) of Table H.2. As expected, selection bias worsens when including offers from a student's third choice (column 9 of Tables 1 and H.2) 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 (3), selection bias is substantially reduced, especially for the first-choice offer.[33]

The results are qualitatively similar to the main specifications, suggesting that differences in school quality are not driving my results. Estimated impacts of enroling at most-preferred schools conditional on enrolment are reported in columns (5)-(6) and (11)-(12) of Table 4. For example, enroling at a student's first choice increases achievement in mathematics by $0.13 - 0.17\sigma$ with respect to peers enroled at the same school who ranked that school as their third choice or lower (columns 3-4, Panel B).

Overall, using different approaches and specifications, the results suggest that students

[32]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.

[33]In addition, columns (4)-(6) of Table 2 show that this design maintains a strong first stage.

Table 6: 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. Dependent variables is match-adjusted KS1 score, obtained subtracting estimated school match effects for a student's type from student's score. 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 baseline 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

enroling at schools ranked by their parents with higher preference experience achievement gains above and beyond the impacts predicted by school VA.[34] 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 6. Results are obtained from specifications similar to column (2), Panel A of Table 4, 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 achievement (columns 3, 8, and 10, respectively). All these characteristics may reflect higher

[34]See Appendix G for additional robustness checks.

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 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 enroling 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 enroling 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 distance to residence. 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 directly assessed with the data at hand, is the allocation to specific teachers. Ahn et al. (2021) find that teacher effectiveness depends on the individual characteristics of students. Investigating the channels through which match effects operate is 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. (2022). Why do households leave school value added on the table? The roles of information and preferences. *American Economic Review*. Forthcoming.

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.

Chetty, R., Friedman, J. N., and Rockoff, J. (2016). Using lagged outcomes to evaluate bias in Value-Added Models. *American Economic Review: Papers & Proceedings*, 106(5):393–399.

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. (2022). The performance of school assignment mechanisms in practice. *Journal of Political Economy*. Forthcoming.

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 Description of parental preferences

I use student-school level data for each institution listed to describe parental preferences as summarised in Section 2. 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.[1]

Second, following Ainsworth et al. (2022), 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.[2] 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. Attributes of feasible schools compared to student's school are described in Table A.1.

I define the individual feasible school set exploiting school catchment boundaries obtained from replication of centralised school assignment (see Section 3 and Appendix C). 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

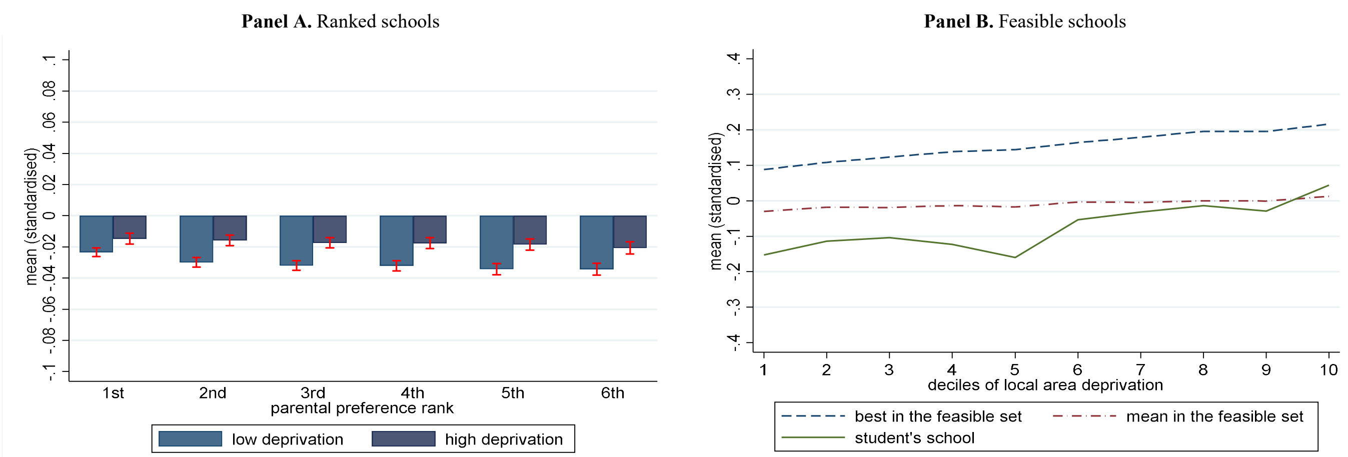
[1] In addition to dummies for number of schools listed and ex-post feasibility of the school, the vector $X'_{is(i,r)}$ includes attributes other than $A_{is(i,r)}$ (e.g., school VA and distance when considering peer quality).

[2] 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.

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 enroled sibling (proxied as detailed in Appendix C) 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.

Conditional on distance and peer quality, parental preference only modestly increases with estimated school VA (see Appendix D for details on school VA estimation). Panel A of Figure A.1 shows that schools ranked higher by parents have slightly higher VA, especially the first choice, consistent with the small gains in school VA from enroling at most-preferred schools illustrated in Section 5 (about 0.01σ). The result that parental preferences do not respond to school VA as much as they respond to peer quality and distance is in line with findings in other contexts (MacLeod and Urquiola, 2019). Panel B shows that students could potentially access schools with even substantially larger VA (with 0.15σ larger VA, columns 2-4 of Table A.1), implying that achievement could substantially increase under alternative allocations if returns to school were homogeneous. However, the findings presented in Section 5 imply that school effects are heterogeneous, offering one potential explanation for the weak correlation between parental preferences and average school VA.

Figure A.1: 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 2 and Appendix A for details.

Table A.1: 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	22.92	0.434	0.152	41.37	0.578	0.352	50.57	0.591	0.204
Deprivation in local area above median	7.13	16.40	0.454	0.186	30.97	0.576	0.445	45.26	0.553	0.240
Deprivation in local area below median	4.39	29.18	0.413	0.127	52.40	0.579	0.275	55.91	0.629	0.178

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 2 and Appendix A for details.

B Parental choice and school mobility

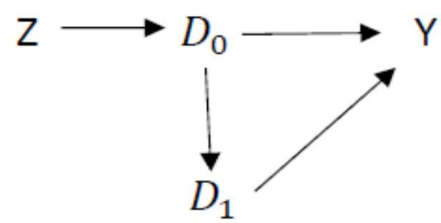
I have shown in Section 2 that the decision to change school after reception year is associated with parental rankings submitted at application (see Figure 4). I further show here that school mobility responds to peer quality more than school VA, consistently with parental preferences for schools. Panel A of Table B.1 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 lower, around 3 p.p. in column 1 and marginally statistically significant, 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 p.p when distance to school is 1km lower with respect to the first choice (Panel A of Table B.1). 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 4 suggests that students with offer from the most preferred schools are less likely to move to a different institution after the reception year. Table B.2 reports estimates from specifications analogue to Table 4 where school mobility indicator is the dependent variable. Panel A shows that, consistently across different specifications, students offered their first choice are $6-10$ p.p. less likely to move to a different school with respect to peers offered their third choice or lower, decreasing to $4-6$ p.p. for the second choice.

School mobility response to centralised assignment is unlikely to constitute a concern for my results. I consider student achievement at Year 2 as the outcome of interest and enrolment at the reception year as the treatment of interest. I have shown that my instrument, centralised school offer conditional on assignment risk, decreases school mobility after the reception year. Figure B.1 illustrates the relationship between student achievement (Y), 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). Among students with the same assignment risk, 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. 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 confound my results. First, control students may gain access at preferred schools in later grades. However, I show that almost no student moves to more-preferred schools than the one offered after the reception year (Figure 4, Panel C). Second, school mobility 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 B.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 B.1: School mobility and school attributes

	(1)	(2)	(3)
First choice VS student's school:			
	Panel A. School mobility		
Peer quality difference	0.0618***	0.0618***	0.0471***
	(0.0028)	(00032)	(0.0036)
School value added difference	0.0291*	0.0213	0.0157
	(0.0154)	(0.0173)	(0.0174)
Distance difference	0.0102***	0.0104***	0.00742**
	(0.0028)	(0.0031)	(0.0032)
N	126,159	106,403	105,725
	Panel B. Residential mobility		
Peer quality difference	0.0025	0.0031	0.0016
	(0.0021)	(0.0024)	(0.0027)
School value added difference	-0.0036	-0.0010	-0.0006
	(0.0119)	(0.0131)	(0.0134)
Distance difference	0.0068***	0.0075***	0.0067***
	(0.0022)	(0.0024)	(0.0025)
N	123,567	104,267	103,617
School choice controls		Y	Y
Individual characteristics			Y

Note. The table shows correlations 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 as school match effect based on student characteristics (see Appendix D). Peer quality is measured by school-level final year test scores, averaged across 2007-2014 cohorts and across mathematics and English. Distance 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). 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 B for details. *** p<0.01, ** p<0.05, * p<0.1

Table B.2: School mobility by offer status

	Same-risk design			Within-school design		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A. School mobility			
Offered 1st choice	-0.0972***	-0.0971***	-0.0972***	-0.0614***	-0.0610***	-0.0604***
	(0.0143)	(0.0144)	(0.0144)	(0.0116)	(0.0116)	(0.0116)
Offered 2nd choice	-0.0611***	-0.0612***	-0.0619***	-0.0396***	-0.0397***	-0.0394***
	(0.0135)	(0.0135)	(0.0135)	(0.0105)	(0.0105)	(0.0105)
N	22628	22628	22628	22628	22628	22628
			Panel B. Residential mobility			
Offered 1st choice	-0.0075	-0.0071	-0.0078	-0.0162	-0.0160	-0.0157
	(0.0128)	(0.0128)	(0.0128)	(0.0110)	(0.0110)	(0.0110)
Offered 2nd choice	-0.0249**	-0.0249**	-0.0253**	-0.0111	-0.0112	-0.0107
	(0.0120)	(0.0120)	(0.0120)	(0.0092)	(0.0092)	(0.0092)
N	22628	22628	22628	22628	22628	22628
Tie-breaker controls	Y	Y	Y	Y	Y	Y
Assignment risk (all schools)	Y	Y	Y			
Assignment risk (ranked schools)				Y	Y	Y
School where enroled FEs				Y	Y	Y
Individual characteristics		Y	Y		Y	Y
Baseline achievement			Y			Y

The table shows the effect of receiving an offer from most-preferred schools 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 (Panel A) or if the students moves residence in the same period (Panel B, defined as a postcode change). Reported are estimated coefficients on offer indicators from the first or second choice from specifications analogue to Table 4. Robust standard errors are reported in parentheses. See Appendix B for details. *** p<0.01, ** p<0.05, * p<0.1

C 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.[3]

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.[4] I flag an applicant as supposed sibling if a student living in the same postcode with exactly the same family-specific observables is enroled 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.[5] 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.[6]

[3] 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 C.1).

[4] 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.

[5] 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.

[6] 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.[7] As some applicants with priority may remain undetected, this is not sufficient to fully replicate school offer. Panel B of Figure C.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).

[7] 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 C.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.[8] Panel B of Figure C.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 C.3.

[8] 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).

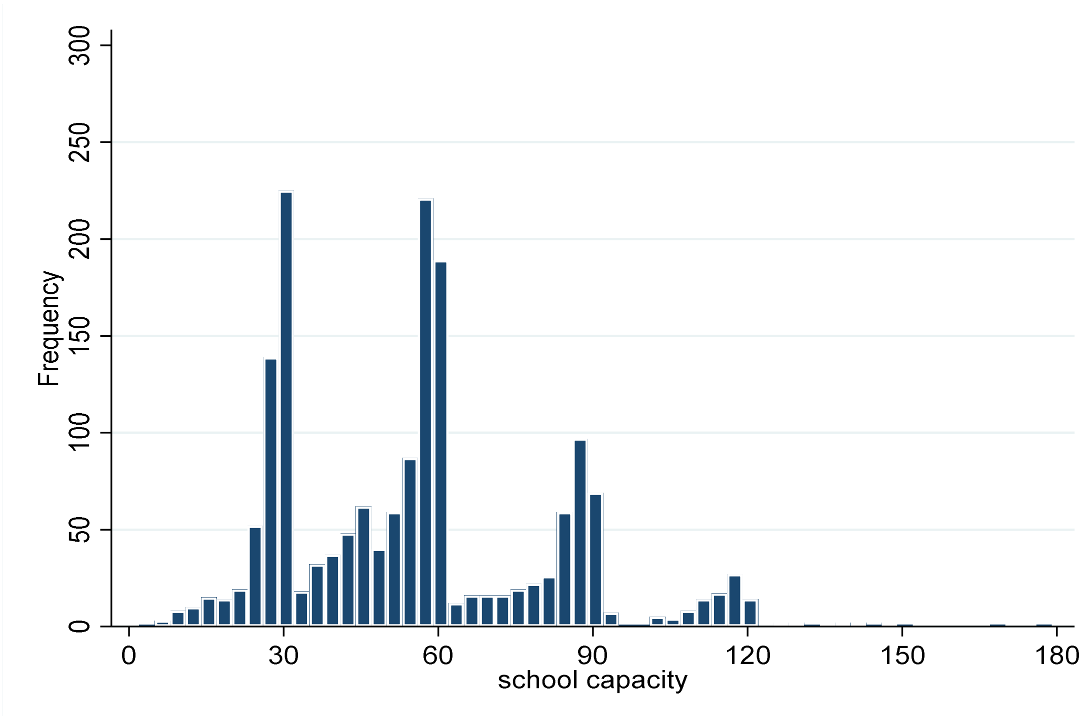
Panel B shows a similar, fuzzier, pattern for school enrolment, as expected.[9] Consistent with the idea that catchment boundary is initially overestimated, Panel C of Figure C.2 shows that median distance threshold decreases roughly monotonically as the algorithm is iterated. When assignment replication is concluded, the estimated median catchment boundary at oversubscribed schools is approximately 470 metres.[10]

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 C.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, marking the benchmark of 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.

[9]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.

[10]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.

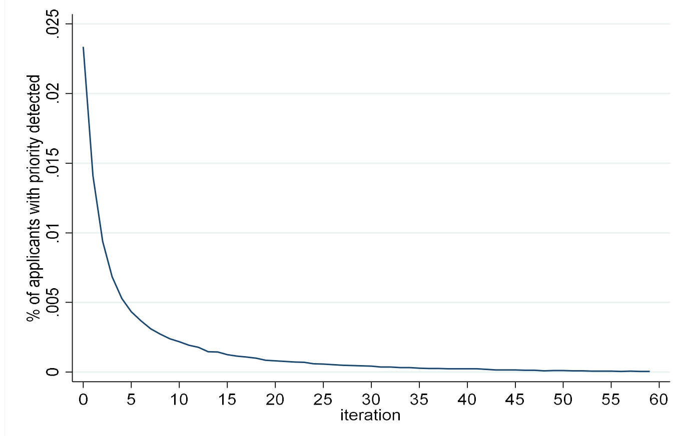
Figure C.1: School capacity



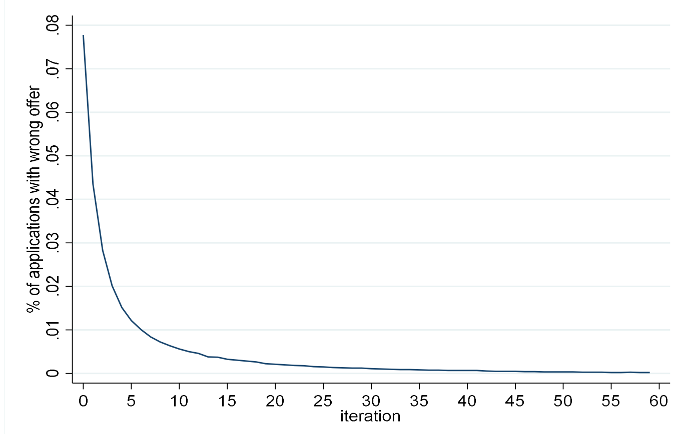
Note. The figure plots the distribution of school capacity in London primary schools. Capacity is approximated by the number of offers issued. Bars represent frequency counts in three-units-wide bins, computed using one observation per school-year. See Section 3 and Appendix C for details.

Figure C.2: Replication of school assignment

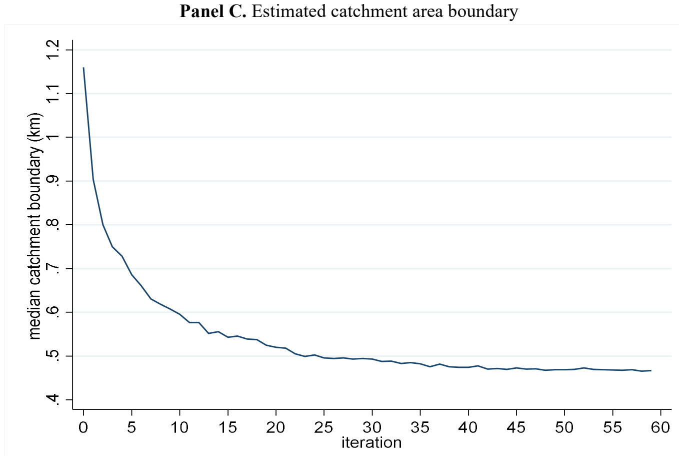
Panel A. Unobserved priorities



Panel B. Error in school offer replication

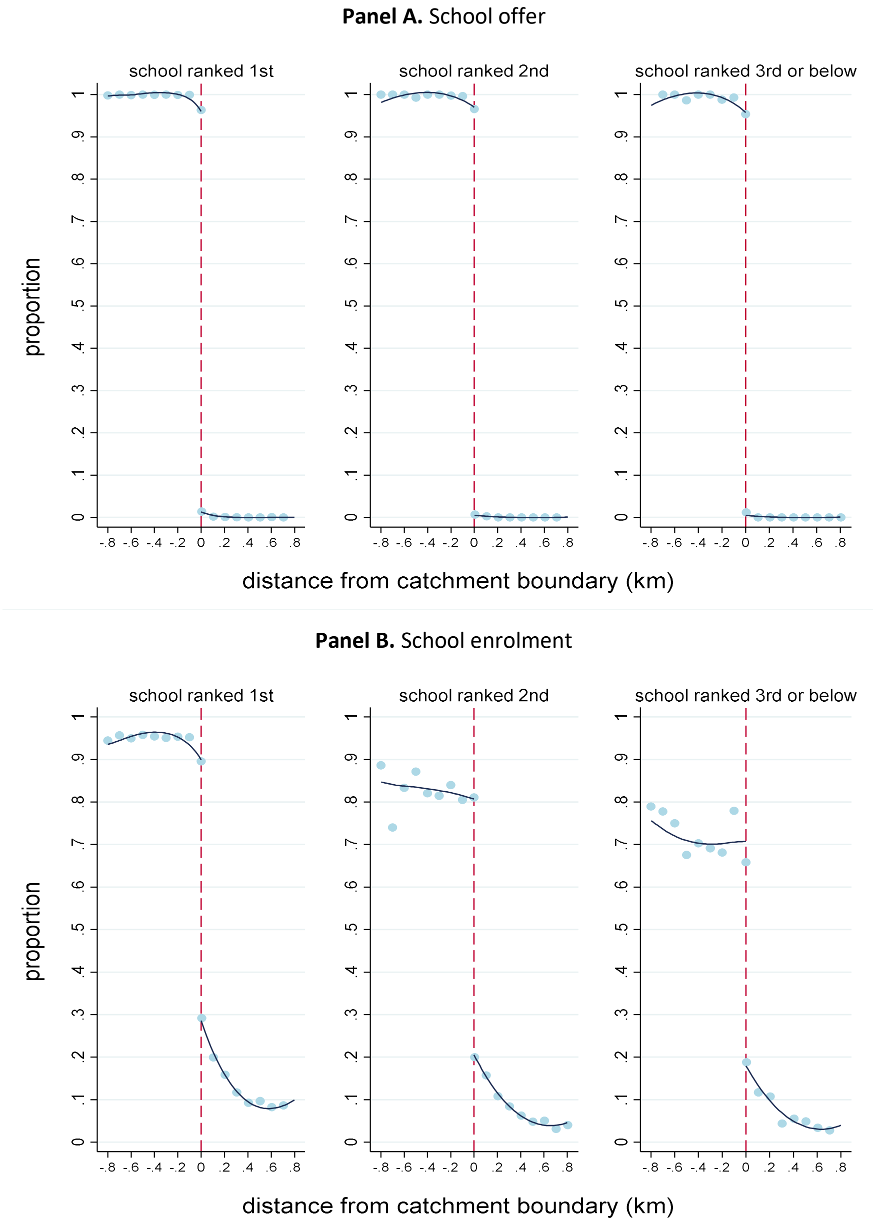


Panel C. Estimated catchment area boundary



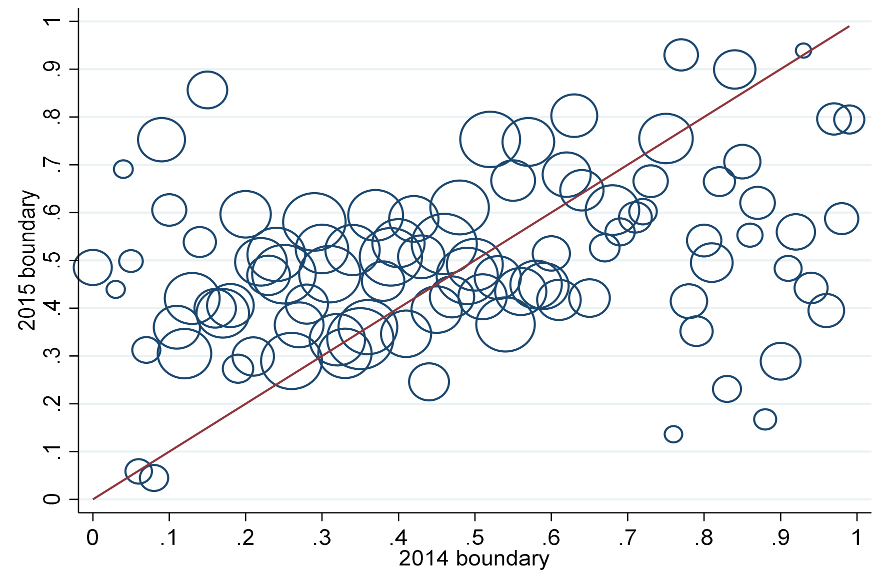
Note. The figure plots the fraction of applicants with admission priority detected (Panel A), the fraction of applicants with wrong predicted offer (Panel B), and median estimated catchment boundary (Panel C) by iteration of the school assignment replication algorithm. The sample includes all applicants 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 C for details.

Figure C.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 C for details.

Figure C.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 100-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 C for details.

D 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 enroled 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, and special education needs, and the deprivation index in the local area of residence (LSOA), and δ_t are year dummies.[11] I additionally control for lagged achievement using EYFSP assessments.[12] 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 ($\hat{\epsilon}_{ijt}$):

$$\hat{\alpha}_s = \frac{\sum\limits_t \sum\limits_i \hat{\epsilon}_{ist}}{\sum\limits_t N_{st}},$$

where N_{st} is the number of students enroled at school s in year t.

Under heterogeneous school effects, I need additional assumptions for 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 $\bar{\mu}_s = 1/N \sum\limits_i \mu_{is}$. VA estimates, therefore, are comparable across institutions only if the average match effect is constant across schools.[13] Under the hypothesis that match effects

[11] 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). All scores are standardised to have zero mean and unit variance by year.

[12] 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. All scores are standardised to have zero mean and unit variance by year. See section 3 on the potential limitations of using EYFSP as baseline achievement measures.

[13] 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 enroled students to be constant across schools. Variation in the latter quantity is likely limited in my empirical analysis since 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 comparing students at the same school 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 increases KS1 score by 0.19σ, 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 D.1. The correlation between estimated school quality and absolute achievement is 37%, suggesting that 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.

I use VA estimates to identify schools that are improving over time. I compute school-year level residuals from equation (11) to estimate year-specific VA, and compute the difference with respect to the previous year. The time series of average VA across all schools is plotted in Figure D.2, showing that estimated effectiveness is roughly constant over my period of study. I measure VA growth at the school as the average absolute annual growth in the estimated VA in the three most recent cohorts available, from 2014 to 2016.

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.

sorting on gains".

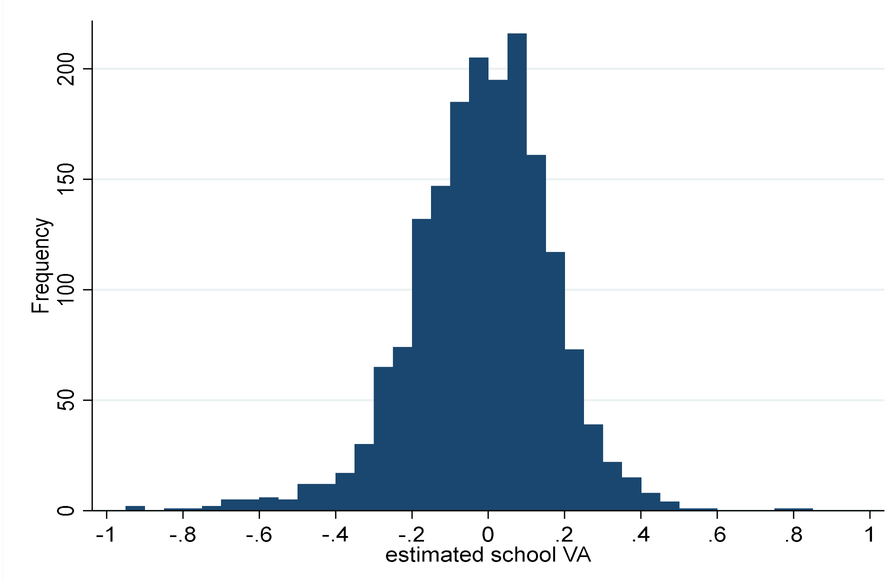
I show in table D.2 that these characteristics are highly predictive of academic achievement, and that further including other student characteristics such as ethnicity or language does not substantially increase the explained variability. As shown in Table D.1, the distribution of students across the eight covariate cells is fairly homogeneous.

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 (cell 1 in Table D.1), and include interactions of school indicators with all other covariate cells. I compute bootstrapped standard errors of match effects using 300 replications. Third, I use bootstrapped 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 D.3 shows mean and standard deviation of estimated match effects.

Estimated school VA strongly predicts student achievement. Table D.4 reports estimates from student-level regression of KS1 score on student's school VA, where both KS1 and school VA are averaged across subjects. Column (1) shows that a 1σ increase in school VA is associated with 0.32σ larger KS1 scores across all students in my estimation sample. Column (4) restricts the sample to students at risk of admission, and column (7) only considers students enroling at the same school from reception to Year 2. In the latter group, the forecast coefficient increases to 0.47σ. These estimates are strongly significant, but also significantly lower than 1, indicating that school VA does not fully predict student achievement. This may reflect both bias in OLS VA estimates, which has been found to be small (Angrist et al., 2016; Chetty et al., 2016; Angrist et al., 2017, 2021), or heterogeneity in school effects. The $\mathbb{R}^2$ coefficient of these regressions is about 0.45, and it remains similar when adding school VA growth (columns 2, 5, and 8) or when using estimated school match effect based on student's characteristics (columns 3, 6, and 9).

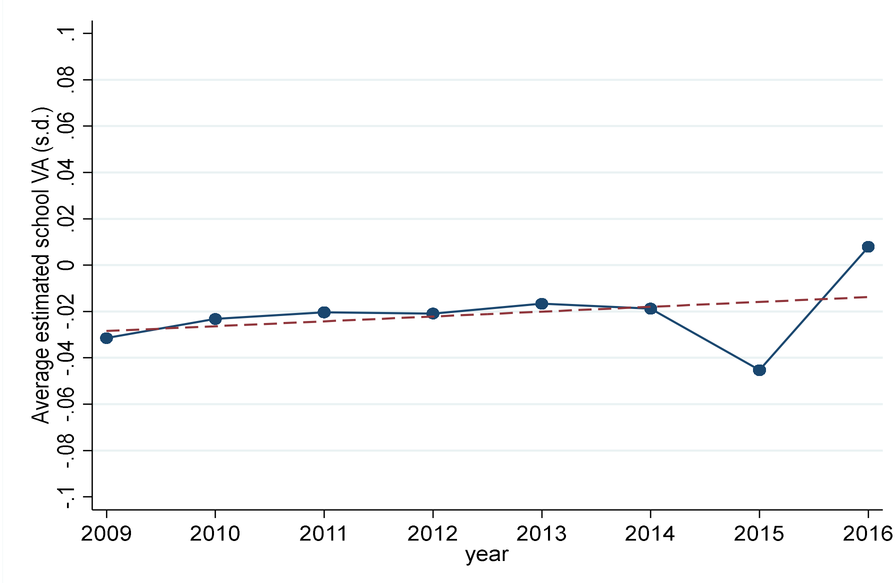
Impacts of enrolment at more-preferred schools on student's school VA are reported in Table D.5. Enroling at a student's first choice increases school VA by about 0.01σ compared to a school ranked second or lower (Panel A), increasing to 0.012σ when compared to schools ranked third or lower (Panel B) and to 0.022σ when compared to schools ranked fourth or lower (Panel C). These results imply a small degree of sorting into more effective schools, which slightly increases when considering school match effects based on student's characteristics (0.016σ, 0.022σ, and 0.031σ, respectively). Finally, I find no evidence that enroling at most preferred schools is related to VA growth.

Figure D.1: Estimated school VA



Note. This figure plots the distribution of estimated school value-added. It reports frequency counts in 0.05-wide bins using one observation per school. Estimates are scaled in standard deviations. Value-added is estimated using KS1 scores in 2009-2016. Student-level regressions control for language, ethnicity, free school meal eligibility, special education needs, gender, deprivation index in the LSOA of residence, and achievement at the reception year (EYFSP). The outcome is standardised to have zero mean and unit variance by year. Value-added is computed as school-level residuals from the regression, separately by subject. Plotted is the average value-added across subjects. See Appendix D for details.

Figure D.2: Estimated school VA growth



Note. This figure plots the time series of the average estimated value-added across schools. Value-added is estimated using KS1 scores in 2009-2016. Student-level regressions control for language, ethnicity, free school meal eligibility, special education needs, gender, deprivation index in the LSOA of residence, and achievement at the reception year (EYFSP). The outcome is standardised to have zero mean and unit variance by year. Value-added is computed as school-year-level residuals from the regression, separately by subject. Plotted is the average value-added across subjects. See Appendix D for details.

Table D.1: Covariate cells

Cell	Socioeconomic disadvantage	Female	High-ability	N	%
1	0	0	0	97,922	14.22
2	0	0	1	100,158	14.54
3	0	1	0	70,518	10.24
4	0	1	1	119,456	17.34
5	1	0	0	99,696	14.47
6	1	0	1	53,046	7.70
7	1	1	0	77,709	11.28
8	1	1	1	70,243	10.20
TOT				**688,748**	**99.99**

Note. This table shows frequency counts by covariate cells used to estimate school match effects based on student's characteristics. 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 define a student high-ability if EYFSP (reception year) achievement is above the median. Zeroes and ones indicates a student belongs or does not belong to a given group, respectively. See Appendix D for details.

Table D.2: 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 D for details. ***p<0.01. ** p<0.05. * p<0.1

Table D.3: 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 D for details.

Table D.4: 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 and Appendix D for details on estimated school VA and match effects. ***p<0.01. ** p<0.05. * p<0.1

Table D.5: Impacts on school VA

	Average school VA		Growth in average school VA		School match effect based on student's characteristics	
	(1)	(2)	(3)	(4)	(5)	(6)
Enroled in 1st choice	0.0101*	0.0100*	0.0033	0.0036	0.0162**	0.0153**
	(0.0056)	(0.0056)	(0.0053)	(0.0053)	(0.0071)	(0.0071)
Enroled in 1st choice	0.0124*	0.0123*	0.0030	0.0033	0.0220***	0.0216***
	(0.0065)	(0.0065)	(0.0064)	(0.0064)	(0.0084)	(0.0083)
Enroled in 2nd choice	0.0043	0.0043	-0.0006	-0.0005	0.0111	0.0120*
	(0.0056)	(0.0056)	(0.0053)	(0.0052)	(0.0070)	(0.0070)
Enroled in 1st choice	0.0220***	0.0218***	0.0064	0.0067	0.0309***	0.0308***
	(0.0078)	(0.0078)	(0.0074)	(0.0074)	(0.0099)	(0.0098)
Enroled in 2nd choice	0.0151**	0.0150**	0.0033	0.0033	0.0212**	0.0237***
	(0.0073)	(0.0073)	(0.0068)	(0.0068)	(0.0090)	(0.0090)
Enroled in 3rd choice	0.0252***	0.0249***	0.0090	0.0090	0.0234**	0.0259**
	(0.0085)	(0.0085)	(0.0079)	(0.0079)	(0.0106)	(0.0105)
N	22,628	22,628	22,628	22,628	22,628	22,628
Tie-breaker controls	Y	Y	Y	Y	Y	Y
Assignment risk (all schools)	Y	Y	Y	Y	Y	Y
Individual characteristics	Y	Y	Y	Y	Y	Y
Baseline achievement		Y		Y		Y

Note. The table shows estimates of the impact of enroling at most-preferred schools on estimated school value-added. Estimates and specifications follow columns (1)-(2) of Table 4. Dependent variables are estimated school value-added in columns (1)-(2), estimated school value-added growth in columns (3)-(4), adn estimated school match effect based on student's characteristics in columns (1)-(2) Robust standard errors are reported in parentheses. See Appendix D and Section 5 for details. ***p<0.01. ** p<0.05. * p<0.1

E 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 (see also Appendix C), students in my sample have marginal priority in 48% of applications.[14]

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.[15] 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 $d_{is} < (\tau_s - \delta)$,[16]
- is *never seated* otherwise.

[14] 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.

[15] I use catchment boundaries obtained in Appendix C 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.

[16] 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:[17]

- $p_{is} = 0$ 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.

[17]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 E.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 E for details.

F 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 enroled at school s (see Appendix D 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 instrumented 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:

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

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 enroling 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):

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

G Robustness checks

I consider two further exercises to test the robustness of my results. First, an important empirical choice in my approach is the bandwidth used to estimate assignment risk (see Appendix E). Estimates of match effects on achievement in mathematics under different bandwidth choices are reported in Table G.1. The optimal data-driven bandwidths chosen for my main analysis are 333 metres and 243 metres in 2015 and 2016, respectively. Columns (1)-(2) consider a unique bandwidth equal to 300 metres across cohorts. This is close to the baseline bandwidth on average, as reflected in the similar sample size. As expected, results in columns (1)-(2) of Table G.1 are similar to estimates in columns (3)-(4) of Table 4. The baseline bandwidth is increased by 25% in columns (3)-(4) of Table G.1, increasing the sample size proportionally. Results are somewhat smaller, but similar to the baseline estimates. Finally, the baseline bandwidth is decreased by 25% in columns (3) and (6), decreasing the sample size proportionally. Results are somewhat larger, but again similar to the baseline estimates.

Second, I test the sensitivity of my results to the exclusion of local authorities where unobserved admission priorities are more likely. I replicate centralised assignment by using school offer to correct for unobserved priorities (see Appendix C). In columns (1)-(2) of Table G.2, I exclude from estimation the four districts granting admission priority to students residing within predefined catchment areas. In columns (3)-(4), I exclude the five districts breaking ties by walking distance rather than straight-line distance.[18] Both these institutional settings may generate larger errors in the estimation of assignment risk. Nonetheless, estimates are similar to baseline findings, suggesting that different admission arrangements do not affect my results.

[18]In summer 2021, the local authorities of Barnet, Brent, Hillingdon, and Redbridge prioritise students using pre-defined catchment areas; the local authorities of Tower Hamlets, Hounslow, Newham, Redbridge and Richmond Upon Thames break ties by walking distance.

Table G.1: Robustness to the choice of bandwidth

	bw = 300m		bw = 1.25Xbaseline		bw = 0.75Xbaseline	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Same-risk design (dep. var.: VA-adjusted KS1score in mathematics)						
Enroled in 1st choice	0.1608***	0.1242**	0.1303**	0.0764*	0.1626**	0.1533***
	(0.0624)	(0.0511)	(0.0555)	(0.0459)	(0.0716)	(0.0588)
Enroled in 2nd choice	0.0675	0.1113***	0.0273	0.0494	0.1231**	0.1627***
	(0.0524)	(0.0426)	(0.0454)	(0.0371)	(0.0620)	(0.0507)
N	23,423	23,423	26,366	26,366	21,702	21,702
Tie-breaker controls	Y	Y	Y	Y	Y	Y
Assignment risk (all schools)	Y	Y	Y	Y	Y	Y
Assignment risk (ranked schools)						
School where enroled FEs						
Individual characteristics	Y	Y	Y	Y	Y	Y
Baseline achievement		Y		Y		Y
Panel B. Within- school design (dep. var. KS1 score in mathematics)						
Enroled in 1st choice	0.1226**	0.0916**	0.1036**	0.0624	0.1373**	0.1023*
	(0.0574)	(0.0464)	(0.0521)	(0.0423)	(0.0671)	(0.0542)
Enroled in 2nd choice	0.0453	0.0660*	0.0296	0.0293	0.0703	0.0832**
	(0.0421)	(0.0341)	(0.0372)	(0.0303)	(0.0504)	(0.0407)
N	23,423	23,423	26,366	26,366	21,702	21,702
Tie-breaker controls	Y	Y	Y	Y	Y	Y
Assignment risk (all schools)						
Assignment risk (ranked schools)	Y	Y	Y	Y	Y	Y
School where enroled FEs	Y	Y	Y	Y	Y	Y
Individual characteristics	Y	Y	Y	Y	Y	Y
Baseline achievement		Y		Y		Y

Note. The table explores robustness of estimated match effects at the most-preferred schools by bandwidth choice. Reported are estimates from specifications analogue to columns (3)-(4) of Table 4. Bandwidth chosen to estimate assignment risk is 300 meters for all cohorts in columns (1)-(2); it is equal to the baseline bandwidth (333 and 243 in 2015 and 2016, respectively) increased by 25% in columns (3)-(4); it is equal to the baseline bandwidth decreased by 25% in columns (5)-(6). Robust standard errors are reported in parentheses. See Appendix G for details. ***p<0.01. ** p<0.05. * p<0.1

Table G.2: Robustness to unobserved admission priorities

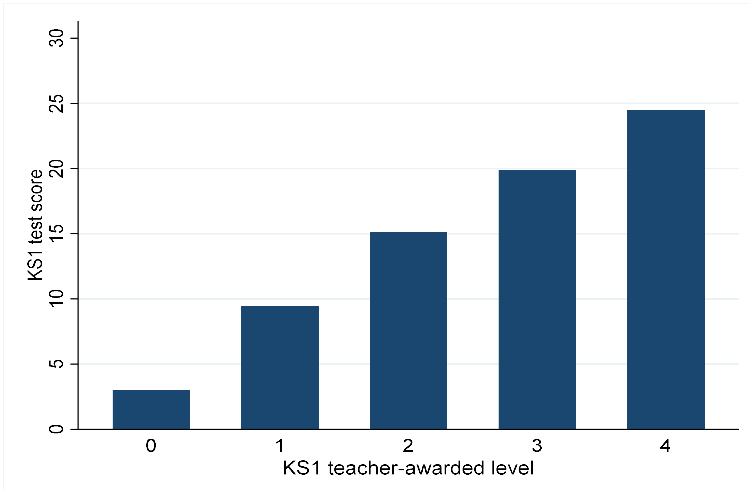
	LAs not using pre-defined catchment		LAs using straight-line distance	
	(1)	(2)	(3)	(4)
	Panel A. Same-risk design (dep. var.: VA-adjusted KS1score in			
Enroled in 1st choice	0.1384**	0.1027**	0.1696***	0.1463***
	(0.0637)	(0.0523)	(0.0643)	(0.0527)
Enroled in 2nd choice	0.0712	0.0924**	0.0708	0.0889**
	(0.0549)	(0.0447)	(0.0548)	(0.0446)
N	19,549	19,549	18,878	18,878
Tie-breaker controls	Y	Y	Y	Y
Assignment risk (all schools)	Y	Y	Y	Y
Assignment risk (ranked schools)				
School where enroled FEs				
Individual characteristics	Y	Y	Y	Y
Baseline achievement		Y		Y
	Panel B. Within- school design (dep. var. KS1 score in mathematics)			
Enroled in 1st choice	0.1451**	0.0815*	0.1492**	0.1075**
	(0.0593)	(0.0477)	(0.0610)	(0.0488)
Enroled in 2nd choice	0.0749*	0.0554	0.0608	0.0558
	(0.0445)	(0.0361)	(0.0450)	(0.0365)
N	19,549	19,549	18,878	18,878
Tie-breaker controls	Y	Y	Y	Y
Assignment risk (all schools)				
Assignment risk (ranked schools)	Y	Y	Y	Y
School where enroled FEs	Y	Y	Y	Y
Individual characteristics	Y	Y	Y	Y
Baseline achievement		Y		Y

Note. The table explores robustness of estimated match effects at the most preferred schools by admission priorities. Reported are estimates from specifications analogue to columns (3)-(4) of Table 4 (Panel A) or columns (5)-(6) of Table 4 (Panel B). Columns (1)-(2) exclude from estimation the four local authorities prioritising students living in pre-defined catchment areas. Columns (3)-(4) exclude from estimation the five local authorities breaking ties by walking, rather than straight-line, distance. Robust standard errors are reported in parentheses. See Appendix G for details. ***p<0.01. ** p<0.05. * p<0.1

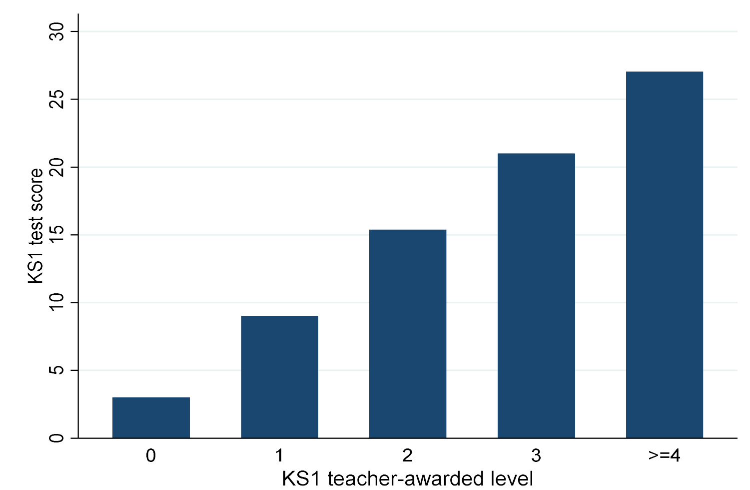
H Additional tables and figures

Figure H.1: KS1 test scores and teacher assessments

Panel A. English



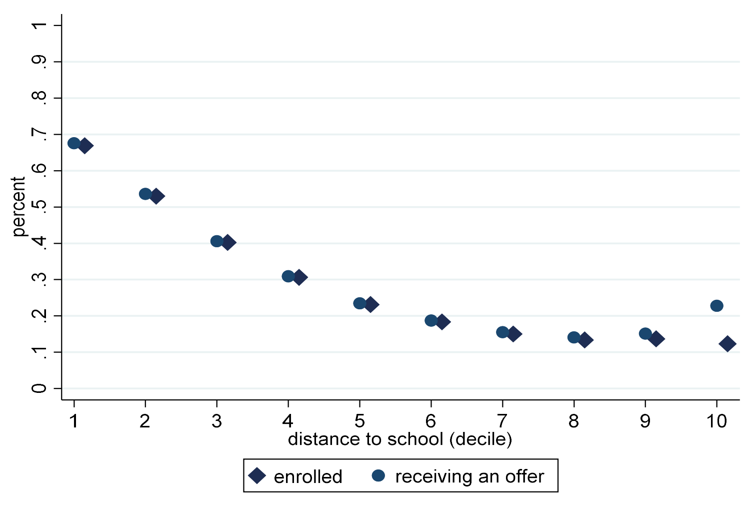
Panel B. Mathematics



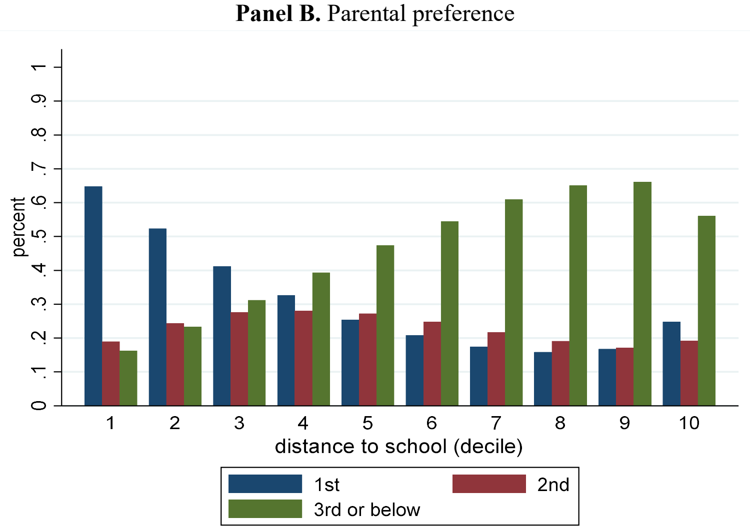
Note. This figure shows correlation between KS1 test scores and teacher-awarded achievement levels. The sample includes achievement data from the 2009-2015 cohorts (KS1 test scores are no longer disclosed from 2016). Level "0" codes the "working towards level 1" outcome. The maximum level awarded is 5 in mathematics and 4 in English. See Section 2 for details

Figure H.2: School admission and distance to school

Panel A. School admission



Panel B. Parental preference



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.

Table H.1: Descriptive statistics: full sample

	Mean (1)	SD (2)
N. of schools listed	3.2069	1.8733
List one school	0.2684	0.4431
List six schools	0.2132	0.4096
List three schools or more	0.5759	0.4942
Offered1st choice	0.8280	0.3773
Offered 2nd choice	0.0819	0.2743
Offered 3rd choice	0.0320	0.1759
Offered one of the top three choices	0.9419	0.2339
Offered one of the listed schools	0.9681	0.1757
Enrol at the offered school at reception	0.8644	0.3423
Enrol in a state-funded school at reception	0.9598	0.1964
Enrol in a state-funded school at Year 2	0.9307	0.2540
Enrol in a different school at Year 2	0.1715	0.3769
N	199,180	

Note. This table shows descriptive statistics about applicants to at least one mainstream state-funded primary school in Greater London in 2014 and 2015. Columns (1) and (2) report averages and standard deviations, respectively. Variables involving school offers are conditional on non-missing observations (N=197,308). See Section 2 for details.

Table H.2: Balance tests (single covariates)

	Uncontrolled			Score specificaton			Within-school		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Demographics									
Free school meal eligible									
Offered 1st choice	-0.0349***	-0.0404***	-0.0412***	-0.0113	-0.0117	-0.0191	-0.0058	-0.0107	-0.0177
	(0.0025)	(0.0032)	(0.0039)	(0.0117)	(0.0131)	(0.0142)	(0.0100)	(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)	198.77 (0.000)	85.33 (0.000)	53.71 (0.000)	0.92 (0.336)	0.46 (0.630)	0.99 (0.397)	0.33 (0.565)	0.63 (0.531)	1.08 (0.356)
Special Education Needs	-0.0056***	-0.0089***	-0.0121***	-0.0025	-0.0027	-0.0040	0.0126	0.0015	-0.0040
Offered 1st choice	(0.0007)	(0.0010)	(0.0013)	(0.0025)	(0.0028)	(0.0032)	(0.0110)	(0.0124)	(0.0137)
		-0.0081***	-0.0115***		-0.0005	-0.0021		-0.0231**	-0.0288**
Offered 2nd choice		(0.0013)	(0.0015)		(0.0033)	(0.0037)		(0.0114)	(0.0128)
			-0.0112***			-0.0041			-0.0152
Offered 3rd choice			(0.0018)			(0.0038)			(0.0152)
Joint significance (p-value)	58.49 (0.000)	36.67 (0.000)	28.48 (0.000)	0.98 (0.323)	0.52 (0.592)	0.67 (0.569)	2.61 (0.106)	1.42 (0.241)	1.31 (0.271)
White	0.0666***	0.0720***	0.0748***	0.0022	-0.0131	-0.0208	-0.0034	-0.0040	-0.0055*
Offered 1st choice	(0.0032)	(0.0040)	(0.0046)	(0.0131)	(0.0147)	(0.0161)	(0.0021)	(0.0025)	(0.0029)
		0.0277***	0.0307***		-0.0347**	-0.0439***		-0.0012	-0.0028
Offered 2nd choice		(0.0056)	(0.0061)		(0.0136)	(0.0154)		(0.0027)	(0.0032)
			0.0203**			-0.0248			
Offered 3rd choice			(0.0082)			(0.0186)			-0.0041
									(0.0034)
Joint significance (p-value)	439.99 (0.000)	188.48 (0.000)	119.91 (0.000)	0.03 (0.864)	3.33 (0.036)	2.75 (0.041)	1.31 (0.253)	2.74 (0.064)	2.16 (0.091)
Black									
Offered 1st choice	-0.0371***	-0.0428***	-0.0472***	-0.0112	-0.0186*	-0.0194	0.0110	0.0107	0.0095
	(0.0025)	(0.0033)	(0.0039)	(0.0097)	(0.0108)	(0.0119)	(0.0083)	(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)	216.06 (0.000)	93.98 (0.000)	59.14 (0.000)	1.34 (0.248)	1.96 (0.140)	1.31 (0.269)	1.77 (0.183)	0.89 (0.412)	0.62 (0.605)
Deprivation in local area									
Offered 1st choice	-0.0221***	-0.0207***	-0.0166***	-0.0000	-0.0011	-0.0026	0.0098***	0.0103***	0.0102**
	(0.0009)	(0.0011)	(0.0013)	(0.0035)	(0.0040)	(0.0043)	(0.0033)	(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)	651.48 (0.000)	244.71 (0.000)	157.74 (0.000)	0.00 (0.990)	0.19 (0.824)	0.45 (0.720)	8.76 (0.003)	4.40 (0.012)	2.94 (0.032)
% with higher education in local area									
Offered 1st choice	0.0184***	0.0206***	0.0195***	-0.0006	-0.0000	0.0004	-0.0044**	-0.0038*	-0.0049**
	(0.0007)	(0.0009)	(0.0010)	(0.0020)	(0.0022)	(0.0024)	(0.0019)	(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)	784.95 (0.000)	325.96 (0.000)	201.39 (0.000)	0.10 (0.746)	0.31 (0.732)	0.29 (0.834)	8.76 (0.003)	5.54 (0.018)	2.45 (0.061)
Baseline achievement									
All subjects									
Offered 1st choice	0.1286***	0.1508***	0.1769***	0.0610*	0.0463	0.0631	0.0292	0.0503	0.0848**
	(0.0071)	(0.0091)	(0.0106)	(0.0336)	(0.0369)	(0.0400)	(0.0283)	(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)	331.65 (0.003)	150.58 (0.000)	103.17 (0.000)	3.29 (0.070)	2.07 (0.126)	1.84 (0.137)	1.06 (0.303)	1.74 (0.175)	3.25 (0.021)
Mathematics									
Offered 1st choice	0.1055***	0.1189***	0.1401***	0.0496	0.0237	0.0406	0.0404	0.0525	0.0709**
	(0.0071)	(0.0090)	(0.0105)	(0.0351)	(0.0383)	(0.0417)	(0.0299)	(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)	223.83 (0.003)	97.02 (0.003)	67.07 (0.000)	2.00 (0.156)	2.22 (0.109)	1.94 (0.121)	1.83 (0.176)	1.29 (0.276)	1.42 (0.235)
N	172,099	172,099	172,099	22,628	22,628	22,628	22,628	22,628	22,628
Ranked school FEs									
Offered school FEs									
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

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 H.3: 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	1739

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 can have different oversubscription statuses in each year. 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 2 for details. ***p<0.01. ** p<0.05. * p<0.1

Working Paper del Dipartimento di Economia e Finanza

1. L. Colombo, H. Dawid, *Strategic Location Choice under Dynamic Oligopolistic Competition and Spillovers*, novembre 2013.
2. M. Bordignon, M. Gamalerio, G. Turati, *Decentralization, Vertical Fiscal Imbalance, and Political Selection*, novembre 2013.
3. M. Guerini, *Is the Friedman Rule Stabilizing? Some Unpleasant Results in a Heterogeneous Expectations Framework*, novembre 2013.
4. E. Brenna, C. Di Novi, *Is caring for elderly parents detrimental to women's mental health? The influence of the European North-South gradient*, novembre 2013.
5. F. Sobbrio, *Citizen-Editors' Endogenous Information Acquisition and News Accuracy*, novembre 2013.
6. P. Bingley, L. Cappellari, *Correlation of Brothers Earnings and Intergenerational Transmission*, novembre 2013.
7. T. Assenza, W. A. Brock, C. H. Hommes, *Animal Spirits, Heterogeneous Expectations and the Emergence of Booms and Busts*, dicembre 2013.
8. D. Parisi, *Is There Room for 'Fear' as a Human Passion in the Work by Adam Smith?*, gennaio 2014.
9. E. Brenna, F. Spandonaro, *Does federalism induce patients' mobility across regions? Evidence from the Italian experience*, febbraio 2014.
10. A. Monticini, F. Ravazzolo, *Forecasting the intraday market price of money*, febbraio 2014.
11. Tiziana Assenza, Jakob Grazzini, Cars Hommes, Domenico Massaro, *PQ Strategies in Monopolistic Competition: Some Insights from the Lab*, marzo 2014.
12. R. Davidson, A. Monticini, *Heteroskedasticity-and-Autocorrelation-Consistent Bootstrapping*, marzo 2014.
13. C. Lucifora, S. Moriconi, *Policy Myopia and Labour Market Institutions*, giugno 2014.
14. N. Pecora, A. Spelta, *Shareholding Network in the Euro Area Banking Market*, giugno 2014.
15. G. Mazzolini, *The economic consequences of accidents at work*, giugno 2014.
16. M. Ambrosanio, P. Balduzzi, M. Bordignon, *Economic crisis and fiscal federalism in Italy*, settembre 2014.
17. P. Bingley, L. Cappellari, K. Tatsiramos, *Family, Community and Long-Term Earnings Inequality*, ottobre 2014.
18. S. Frazzoni, M. L. Mancusi, Z. Rotondi, M. Sobrero, A. Vezzulli, *Innovation and export in SMEs: the role of relationship banking*, novembre 2014.
19. H. Gnutzmann, *Price Discrimination in Asymmetric Industries: Implications for Competition and Welfare*, novembre 2014.
20. A. Baglioni, A. Boitani, M. Bordignon, *Labor mobility and fiscal policy in a currency union*, novembre 2014.
21. C. Nielsen, *Rational Overconfidence and Social Security*, dicembre 2014.
22. M. Kurz, M. Motolese, G. Piccillo, H. Wu, *Monetary Policy with Diverse Private Expectations*, febbraio 2015.
23. S. Piccolo, P. Tedeschi, G. Ursino, *How Limiting Deceptive Practices Harms Consumers*, maggio 2015.
24. A.K.S. Chand, S. Currarini, G. Ursino, *Cheap Talk with Correlated Signals*, maggio 2015.
25. S. Piccolo, P. Tedeschi, G. Ursino, *Deceptive Advertising with Rational Buyers*, giugno 2015.

26. S. Piccolo, E. Tarantino, G. Ursino, *The Value of Transparency in Multidivisional Firms*, giugno 2015.
27. G. Ursino, *Supply Chain Control: a Theory of Vertical Integration*, giugno 2015.
28. I. Aldasoro, D. Delli Gatti, E. Faia, *Bank Networks: Contagion, Systemic Risk and Prudential Policy*, luglio 2015.
29. S. Moriconi, G. Peri, *Country-Specific Preferences and Employment Rates in Europe*, settembre 2015.
30. R. Crinò, L. Ogliari, *Financial Frictions, Product Quality, and International Trade*, settembre 2015.
31. J. Grazzini, A. Spelta, *An empirical analysis of the global input-output network and its evolution*, ottobre 2015.
32. L. Cappellari, A. Di Paolo, *Bilingual Schooling and Earnings: Evidence from a Language-in-Education Reform*, novembre 2015.
33. A. Litina, S. Moriconi, S. Zanaj, *The Cultural Transmission of Environmental Preferences: Evidence from International Migration*, novembre 2015.
34. S. Moriconi, P. M. Picard, S. Zanaj, *Commodity Taxation and Regulatory Competition*, novembre 2015.
35. M. Bordignon, V. Grembi, S. Piazza, *Who do you blame in local finance? An analysis of municipal financing in Italy*, dicembre 2015.
36. A. Spelta, *A unified view of systemic risk: detecting SIFIs and forecasting the financial cycle via EWSs*, gennaio 2016.
37. N. Pecora, A. Spelta, *Discovering SIFIs in interbank communities*, febbraio 2016.
38. M. Botta, L. Colombo, *Macroeconomic and Institutional Determinants of Capital Structure Decisions*, aprile 2016.
39. A. Gamba, G. Immordino, S. Piccolo, *Organized Crime and the Bright Side of Subversion of Law*, maggio 2016.
40. L. Corno, N. Hildebrandt, A. Voena, *Weather Shocks, Age of Marriage and the Direction of Marriage Payments*, maggio 2016.
41. A. Spelta, *Stock prices prediction via tensor decomposition and links forecast*, maggio 2016.
42. T. Assenza, D. Delli Gatti, J. Grazzini, G. Ricchiuti, *Heterogeneous Firms and International Trade: The role of productivity and financial fragility*, giugno 2016.
43. S. Moriconi, *Taxation, industry integration and production efficiency*, giugno 2016.
44. L. Fiorito, C. Orsi, *Survival Value and a Robust, Practical, Joyless Individualism: Thomas Nixon Carver, Social Justice, and Eugenics*, luglio 2016.
45. E. Cottini, P. Ghinetti, *Employment insecurity and employees' health in Denmark*, settembre 2016.
46. G. Cecere, N. Corrocher, M. L. Mancusi, *Financial constraints and public funding for eco-innovation: Empirical evidence on European SMEs*, settembre 2016.
47. E. Brenna, L. Gitto, *Financing elderly care in Italy and Europe. Is there a common vision?*, settembre 2016.
48. D. G. C. Britto, *Unemployment Insurance and the Duration of Employment: Theory and Evidence from a Regression Kink Design*, settembre 2016.
49. E. Caroli, C.Lucifora, D. Vigani, *Is there a Retirement-Health Care utilization puzzle? Evidence from SHARE data in Europe*, ottobre 2016.
50. G. Femminis, *From simple growth to numerical simulations: A primer in dynamic programming*, ottobre 2016.
51. C. Lucifora, M. Tonello, *Monitoring and sanctioning cheating at school: What works? Evidence from a national evaluation program*, ottobre 2016.

52. A. Baglioni, M. Esposito, *Modigliani-Miller Doesn't Hold in a "Bailinable" World: A New Capital Structure to Reduce the Banks' Funding Cost*, novembre 2016.
53. L. Cappellari, P. Castelnovo, D. Checchi, M. Leonardi, *Skilled or educated? Educational reforms, human capital and earnings,* novembre 2016.
54. D. Britto, S. Fiorin, *Corruption and Legislature Size: Evidence from Brazil*, dicembre 2016.
55. F. Andreoli, E. Peluso, *So close yet so unequal: Reconsidering spatial inequality in U.S. cities*, febbraio 2017.
56. E. Cottini, P. Ghinetti, *Is it the way you live or the job you have? Health effects of lifestyles and working conditions,* marzo 2017.
57. A. Albanese, L. Cappellari, M. Leonardi, *The Effects of Youth Labor Market Reforms: Evidence from Italian Apprenticeships*; maggio 2017.
58. S. Perdichizzi, *Estimating Fiscal multipliers in the Eurozone. A Nonlinear Panel Data Approach*, maggio 2017.
59. S. Perdichizzi, *The impact of ECBs conventional and unconventional monetary policies on European banking indexes returns*, maggio 2017.
60. E. Brenna, *Healthcare tax credits: financial help to taxpayers or support to higher income and better educated patients? Evidence from Italy*, giugno 2017.
61. G. Gokmen, T. Nannicini, M. G. Onorato, C. Papageorgiou, *Policies in Hard Times: Assessing the Impact of Financial Crises on Structural Reforms*, settembre 2017.
62. M. Tettamanzi, *E Many Pluribus Unum: A Behavioural Macro-Economic Agent Based Model*, novembre 2017.
63. A. Boitani, C. Punzo, *Banks' leverage behaviour in a two-agent New Keynesian model*, gennaio 2018.
64. M. Bertoni, G. Brunello, L. Cappellari, *Parents, Siblings and Schoolmates. The Effects of Family-School Interactions on Educational Achievement and Long-term Labor Market Outcomes*, gennaio 2018.
65. G. P. Barbetta, G. Sorrenti, G. Turati, *Multigrading and Child Achievement*, gennaio 2018.
66. S. Gagliarducci, M. G. Onorato, F. Sobbrio, G. Tabellini, *War of the Waves: Radio and Resistance During World War II,* febbraio 2018.
67. P. Bingley, L. Cappellari, *Workers, Firms and Life-Cycle Wage Dynamics*, marzo 2018.
68. A. Boitani, S. Perdichizzi, *Public Expenditure Multipliers in recessions. Evidence from the Eurozone*, marzo 2018.
69. M. Le Moglie, G. Turati, *Electoral Cycle Bias in the Media Coverage of Corruption News*, aprile 2018.
70. R. Davidson, A. Monticini, *Improvements in Bootstrap Inference*, aprile 2018.
71. R. Crinò, G. Immordino, S. Piccolo, *Fighting Mobile Crime*, giugno 2018.
72. R. Caminal, L. Cappellari, A. Di Paolo, *Linguistic skills and the intergenerational transmission of language*, agosto 2018.
73. E. Brenna, L. Gitto, *Adult education, the use of Information and Communication Technologies and the impact on quality of life: a case study*, settembre 2018.
74. M. Bordignon, Y. Deng, J. Huang, J. Yang, *Plunging into the Sea: Ideological Change, Institutional Environments and Private Entrepreneurship in China*, settembre 2018.
75. M. Bordignon, D. Xiang, L. Zhan, *Predicting the Effects of a Sugar Sweetened Beverage Tax in a Household Production Model*, settembre 2018.
76. C. Punzo, L. Rossi, *The Redistributive Effects of a Money-Financed Fiscal Stimulus*, gennaio 2019.
77. A. Baglioni, L. Colombo, P. Rossi, *Debt restructuring with multiple bank relationships*, gennaio 2019.

78. E. Cottini, P. Ghinetti, S. Moriconi, *Higher Education Supply, Neighbourhood effects and Economic Welfare*, febbraio 2019.
79. S. Della Lena, F. Panebianco, *Cultural Transmission with Incomplete Information: Parental Perceived Efficacy and Group Misrepresentation,* marzo 2019.
80. T. Colussi, Ingo E. Isphording, Nico Pestel, *Minority Salience and Political Extremism*, marzo 2019.
81. G. P. Barbetta, P. Canino, S. Cima, *Let's tweet again? The impact of social networks on literature achievement in high school students: Evidence from a randomized controlled trial*, maggio 2019.
82. Y. Brilli, C. Lucifora, A. Russo, M. Tonello, *Vaccination take-up and health: evidence from a flu vaccination program for the elderly*, giugno 2019.
83. C. Di Novi, M. Piacenza, S. Robone, G. Turati, *Does fiscal decentralization affect regional disparities in health? Quasi-experimental evidence from Italy*, luglio 2019.
84. L. Abrardi, L. Colombo, P. Tedeschi, *The Gains of Ignoring Risk: Insurance with Better Informed Principals*, luglio 2019.
85. A. Garnero, C. Lucifora, *Turning a Blind Eye? Compliance to Minimum Wages and Employment*, gennaio 2020.
86. M. Bordignon, M. Gamalerio, E. Slerca, G. Turati, *Stop invasion! The electoral tipping point in anti-immigrant voting*, marzo 2020.
87. D. Vigani, C. Lucifora, *Losing control? Unions' Representativeness, "Pirate" Collective Agreements and Wages*, marzo 2020.
88. S. L. Comi, E. Cottini, C. Lucifora, *The effect of retirement on social relationships: new evidence from SHARE*, maggio 2020.
89. A. Boitani, S. Perdichizzi, C. Punzo, *Nonlinearities and expenditure multipliers in the Eurozone*, giugno 2020.
90. R. A. Ramos, F. Bassi, D. Lang, *Bet against the trend and cash in profits*, ottobre 2020.
91. F. Bassi, *Chronic Excess Capacity and Unemployment Hysteresis in EU Countries. A Structural Approach*, ottobre 2020.
92. M. Bordignon, T. Colussi, *Dancing with the Populist. New Parties, Electoral Rules and Italian Municipal Elections*, ottobre 2020.
93. E. Cottini, C. Lucifora, G. Turati, D. Vigani, *Children Use of Emergency Care: Differences Between Natives and Migrants in Italy*, ottobre 2020.
94. B. Fanfani, *Tastes for Discrimination in Monopsonistic Labour Markets*, ottobre 2020.
95. B. Fanfani, *The Employment Effects of Collective Bargaining*, ottobre 2020.
96. O. Giuntella, J. Lonsky, F. Mazzonna, L. Stella, *Immigration Policy and Immigrants' Sleep. Evidence from DACA*, dicembre 2020.
97. E. Cottini, P. Ghinetti, E. Iossa, P. Sacco, *Stress and Incentives at Work*, gennaio 2021.
98. L. Pieroni, M. R. Roig, L. Salmasi, *Italy: immigration and the evolution of populism*, gennaio 2021.
99. L. Corno, E. La Ferrara, A. Voena, *Female Genital Cutting and the Slave Trade*, febbraio 2021.
100. O. Giuntella, L. Rotunno, L. Stella, *Trade Shocks, Fertility, and Marital Behavior* , marzo 2021.
101. P. Bingley, L. Cappellari, K. Tatsiramos, *Parental Assortative Mating and the Intergenerational Transmission of Human Capital*, aprile 2021.
102. F. Devicienti, B. Fanfani, *Firms' Margins of Adjustment to Wage Growth. The Case of Italian Collective Bargaining*; aprile 2021.
103. C. Lucifora, A. Russo, D. Vigani, *Does prescribing appropriateness reduce health expenditures? Main e_ects and unintended outcomes*, maggio 2021.

104. T. Colussi, *The Political Effects of Threats to the Nation: Evidence from the Cuban Missile Crisis*, giugno 2021.
105. M. Bordignon, N. Gatti, M. G. Onorato, *Getting closer or falling apart? Euro countries after the Euro crisis*, giugno 2021.
106. E. Battistin, M. Ovidi, *Rising Stars*, giugno 2021.
107. D. Checchi, A. Fenizia, C. Lucifora, PUBLIC SECTOR JOBS: *Working in the public sector in Europe and the US,* giugno 2021.
108. K. Aktas, G. Argentin, G. P. Barbetta, G. Barbieri, L. V. A. Colombo, *High School Choices by Immigrant Students in Italy: Evidence from Administrative Data,* luglio 2021.
109. B. Fanfani, C. Lucifora, D. Vigani, *Employer Association in Italy. Trends and Economic Outcomes*, luglio 2021.
110. F. Bassi, A. Boitani, *Monetary and macroprudential policy: The multiplier effects of cooperation,* settembre 2021.
111. S. Basiglio, A. Foresta, G. Turati, *Impatience and crime. Evidence from the NLSY97,* settembre 2021.
112. A. Baglioni, A. Monticini, D. Peel, *The Impact of the ECB Banking Supervision Announcements on the EU Stock Market*, novembre 2021.
113. E. Facchetti, L. Neri, M. Ovidi, *Should you Meet The Parents? The impact of information on non-test score attributes on school choice*, dicembre 2021.
114. M. Bratti, E. Cottini, P. Ghinetti, *Education, health and health-related behaviors: Evidence from higher education expansion*, febbraio 2022.
115. A. Boitani, C. Dragomirescu-Gaina, *News and narratives: A cointegration analysis of Russian economic policy uncertainty*, aprile 2022.
116. D. Delli Gatti, J. Grazzini, D. Massaro, F. Panebianco, *The Impact of Growth on the Transmission of Patience,* luglio 2022.
117. I. Torrini, C. Lucifora, A. Russo*, The Long-Term Effects of Hospitalization on Health Care Expenditures: An Empirical Analysis for the Young-Old Population*, luglio 2022.
118. T. Colussi, M. Romagnoli, E. Villar, *The Intended and Unintended Consequences of Taxing Waste*, settembre 2022.
119. D. Delli Gatti, G. Iannotta, *Behavioural credit cycles,* settembre 2022.
120. C. Punzo, G. Rivolta, *Money versus debt financed regime: Evidence from an estimated DSGE model*, novembre 2022.
121. Ovidi M., *Parents Know Better: Sorting on Match Effects in Primary School*, novembre 2022.