

Ovidi, Marco

Working Paper

Parents know better: Primary school choice and student achievement in London

Working Paper, No. 919

Provided in Cooperation with:

School of Economics and Finance, Queen Mary University of London

Suggested Citation: Ovidi, Marco (2021) : Parents know better: Primary school choice and student achievement in London, Working Paper, No. 919, Queen Mary University of London, School of Economics and Finance, London

This Version is available at:

<https://hdl.handle.net/10419/247188>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Parents know better: primary school choice and student achievement in London

Marco Ovidi

ISSN 1473-0278

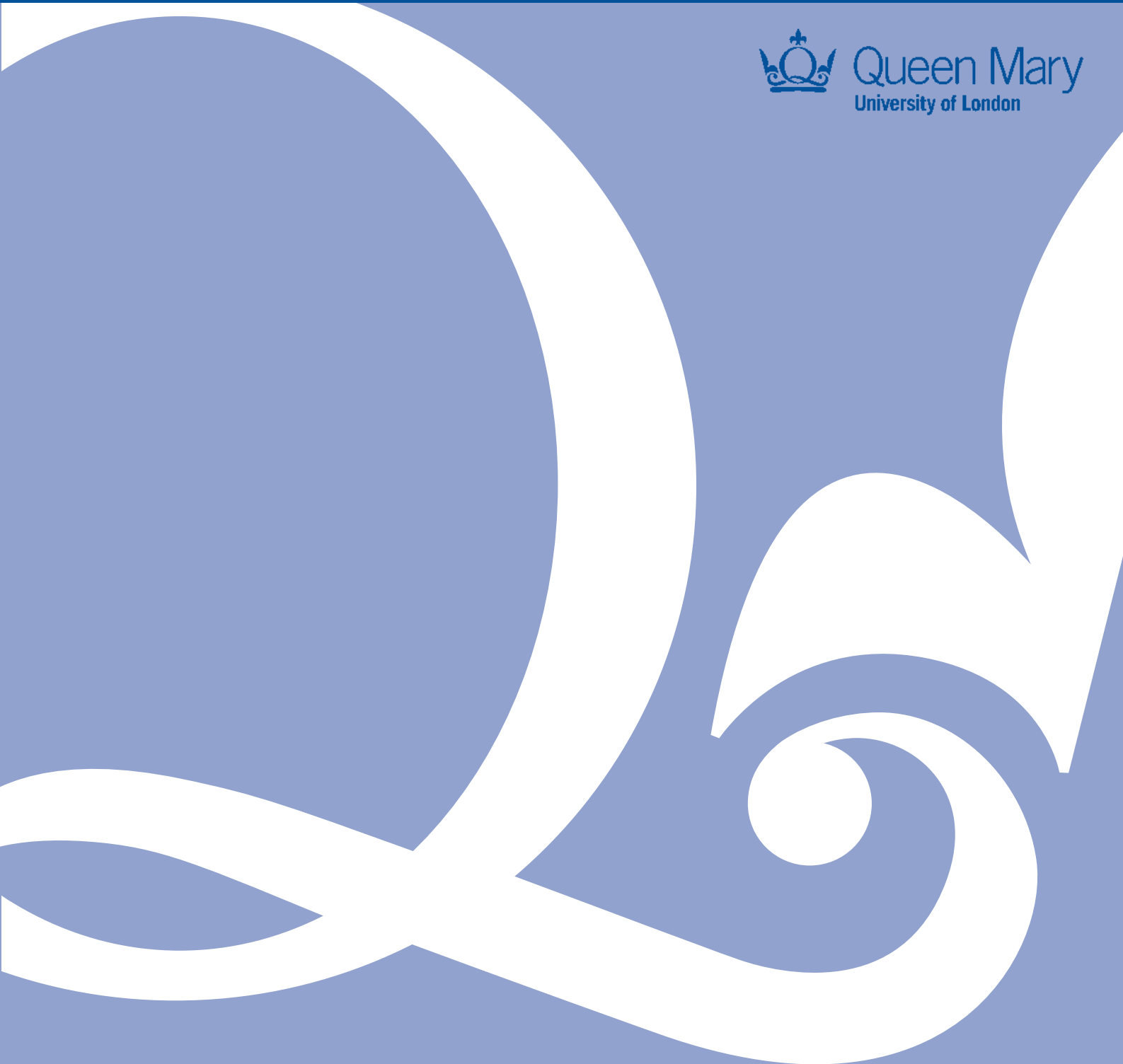
Working Paper No. 919

January 2021

Ù&@ [|Á ÀÒ&[} [{ æ• Áæ ãÁŌā æ &^



Queen Mary
University of London



Parents know better: primary school choice and student achievement in London *

Marco Ovidi[†]

Queen Mary University of London

6th January 2021

Abstract

Expanding parental choice in education may increase system-wide productivity if parents select schools that form a specifically good match with their children. I investigate the effect of attending a preferred school on student achievement in London primary schools. I exploit as good as random variation in admission to preferred schools arising from centralised assignment which awards school offer based on admission priority, residential distance, and parental preference. I compare students around year-specific catchment boundaries that cannot be exactly anticipated by parents. I find that attending the school of choice increases student achievement compared to an institution with lower parental preference but same value-added. Results suggest that parents select schools that are specifically effective in increasing their children's achievement, improving the efficiency of school seats allocation. I show that parents of low-ability males select schools with lower peer quality and these likely better suit their learning needs.

JEL Codes: H75, I21, I24, I28

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

*I am especially grateful to Erich Battistin for careful support and guidance. My thanks to Francesca Cornaglia and Anna Raute for frequent discussion and useful advice. I also thank Marco Bertoni, Massimiliano Bratti, Damon Clark, Lorenzo Cappellari, Antonio Dalla Zuanna, Francesco Fasani, Emla Fitzsimons, Francois Gerard, Soledad Giardili, Ellen Greaves, Peter Hull, Marco Manacorda, José Montalbán Castilla, Lorenzo Neri, Barbara Petrongolo, Santiago Perez Vincent, Camille Terrier, Olmo Silva, Laura van der Erve, Larissa Zierow and seminar participants at Queen Mary University, the Institute of Fiscal Studies, the 6th SIdE WEEE Workshop, the 17th Brucchi Luchino workshop, the 22nd IZA Summer School in Labour Economics, the 10th IWAE workshop, the 72nd European Meeting of the Econometric Society, the 4th IZA workshop on Economics of Education, the II Padova Applied Economics Workshop, the 35th AIEL National Conference of Labour Economics, and Education work in progress seminars at LSE for helpful discussions and comments. Any responsibility for the views expressed in the article rests solely with the author.

[†]School of Economics and Finance, Queen Mary University of London, Mile End Road, London E1 4NS, UK. Contact: m.ovid@qmul.ac.uk.

1 Introduction

Educational policy-makers around the world are increasingly expanding parental choice in an effort to increase school productivity. Open enrolment programmes, in particular, elicit parental preferences about desired institutions and assign students to the highest school in this list with available seats.¹ Parental choice can increase system-wide productivity by sparking demand-side pressure on schools to improve (Hoxby, 2003). This channel requires parents to reward the most effective schools, in the sense of their causal impact on student achievement (Rothstein, 2006). It is strongly debated, however, whether parents seek effective schools or merely respond to indicators driven by neighbourhood composition, such as test scores. Empirical evidence to date suggests that parents may not value effective schools but rather reward geographical proximity and peer quality (Mizala and Urquiola, 2013; Imberman and Lovenheim, 2016; MacLeod and Urquiola, 2019; Abdulkadiroglu et al., 2020).

I study whether parents sort into schools that are specifically effective in raising achievement of their children. Even if peer preferences weaken the incentives for schools to improve their effectiveness, sorting based on student-school match may strengthen the case for parental choice (Barseghyan et al., 2019). Selection on student-school match implies that the impact of attending the school of choice on achievement exceeds the average school value-added. In previous studies estimating the impact of attending a preferred school, these two effects are not separately assessed and compared.²

Using administrative records from centralised assignment of primary school students in London, I show substantial heterogeneity in parental rankings of a given school, even conditional on highly-valued attributes such as peer quality and distance to residence. This possibly reflects selection of schools by parents based on the specific suitability for their children’s learning needs. Leveraging quasi-experimental variation in school assignment, I investigate whether attending the school of choice affects student learning trajectories over

¹Beyond England, studied in this paper, school open enrolment policies have been implemented in many of the largest U.S. districts, serving about 8 million students (Whitehurst, 2017) and in other urban areas around the world such as Amsterdam (De Haan et al., 2015), Barcelona (Calsamiglia and Guell, 2018), Paris Fack et al., 2019) and Beijing (He, 2017).

²For example, Deming et al. (2014) find positive effects of attending the preferred school on postsecondary educational outcomes only for applicants with higher gains in school value-added, and conclude that parental choice does not improve school productivity but just redistributes seats in effective institutions. Their investigation, however, does not distinguish the impact of attending a school with higher parental preference from the average effectiveness of the attended school.

and beyond school value-added.

The ideal experiment would compare the learning outcomes of students randomly enrolled in otherwise identical schools, except for the preference rank assigned by parents. I approximate random variation in school assignment by exploiting tie-breaking embedded in centralised assignment (Abdulkadiroglu et al., 2019). School offers in London are generated by local districts through a deferred acceptance algorithm (DA, Gale and Shapley, 1962) based on parental preferences and school priorities. School seats are typically rationed, with the large majority of institutions being oversubscribed. In case of excess demand, distance to the school is used as a tie-breaker between applicants with equal priority, generating catchment areas that vary year by year depending on the equilibrium allocation of school seats.

I consider applicants living at the boundary of preferred schools' catchment who, based on their preferences, have the same chance of admission. I compare students enrolled, for example, in the most preferred school to students who did not get a seat at that school because their place of residence is marginally beyond the catchment boundary. I address selection from application choices by conditioning on preference variables that are relevant for admission risk (Abdulkadiroglu et al., 2014, 2017, 2019). The identifying assumption is that parents are not able to perfectly anticipate the catchment boundary, which depends on choices of all other candidates at the time of application. Consistent with this expectation, I show that applicants on either side of the boundary have similar observable characteristics, and that parental preference for the school is continuous at the distance cut-off.³

I compare students enrolled in schools with similar average quality. The simple comparison of outcomes at the admission cut-off conflates the match effect from attending a preferred school and the gain in average school value-added. I control for school quality by using baseline estimates of school value-added, obtained following Deming et al. (2014). I show

³Previous research has demonstrated how distance-based admissions resulted in fierce competition for residential housing in the vicinity of popular schools, with parents willing to pay a substantial premium (e.g., Machin, 2011; Gibbons et al., 2013 and Battistin and Neri, 2017). This makes identification more challenging in London compared to other urban districts with lottery assignment, such as Boston (Abdulkadiroglu et al., 2011), Charlotte (Deming et al., 2014), and Denver (Abdulkadiroglu et al., 2017). Others have considered distance-based eligibility for policy interventions (Masi, 2018) or school admission (Gorman and Walker, 2020) to investigate school choice of low-income families or impacts of missing out on the most preferred school. These studies, however, have not exploited the quasi-experimental variation arising from centralised assignment.

that these estimates closely predict causal effects of school on achievement, relaxing concerns about potentially imprecise measurement of school quality. I also show that my conclusions hold when using a complementary research design that compares students enrolling in the same school, ranked by their parents with different preferences.

London offers a unique context to study parental choice. Primary schools are small and enrol just about 50 students on average. High population and school density imply that the typical family faces several alternatives within short commuting distances. Parental choice is well-established and data on school performance are made widely available to the public to ensure comparability of institutions. With high competition for seats, parents must target extremely narrow areas around a school: catchment boundaries average at about 600 meters and can be as close as 300 meters from the school. As a result, the fraction of applicants missing out on preferred schools is systematically the highest in the country.

I use administrative records on all primary school applicants in 2014-15, with about 200,000 students involved in the centralized assignment process. Data are linked to the census of all students in the state education system including socioeconomic characteristics, educational achievement, and granular information on residential location. I trace catchment area of all schools, and the relative position of each applicant with respect to the boundary, by replicating the algorithm used for school seats allocation. The main limitation in my empirical analysis is that applicants enjoying admission priorities, mostly siblings of current students, are imperfectly observed. I show that postcode of residence can be used to build a proxy for siblings at the school based on location of students in older cohorts. Results are unchanged when using this proxy to control for admission priority.

I begin by documenting that parents rank schools by proximity and peer quality. Under DA, records on rank-order lists of schools reveal parental preference among listed institutions (Fack et al., 2019).⁴ Consistently with other studies (e.g., Abdulkadiroglu et al., 2020), I show that, conditional on distance and peer quality, preferences are unrelated to school value-added. Similar conclusions are drawn when considering feasible schools that parents did not rank. I show that parents are close to maximise available peer quality but leave substantial value-added “on the table”, in line with findings by Ainsworth et al. (2020).

⁴As they likely omit to rank unfeasible schools, Fack et al. (2019) show parental rankings constitute a partial order of their true preferences. For example, first-choice school is not necessarily the most preferred institution, but it is preferred to the second choice as long as parents act rationally.

However, I show substantial heterogeneity in parental preferences. Distance, peer quality, and value-added explain only 40% of the variation in preferences, increasing to just 50% when accounting for unobserved school traits. Nevertheless, I show that parental rankings at application reflect robust and solid preference for schools. First, parents avoid lower-ranked schools than the one assigned even three years after application. Second, the decision of moving their children to another school responds to distance and peer quality, similarly to patterns observed at application. Heterogeneity and consistency of parental preferences motivate the investigation of returns to sorting into specific schools.

I find positive effects on students learning from attending the school of choice over and beyond school quality. Enrolment in a preferred school increases achievement in mathematics by 0.09 s.d. compared to a similarly effective institution with lower parental preference. I show that my conclusions are robust to alternative specifications and parametric choices of running variable controls.

One potential explanation is that parents of low-ability male students choose schools with lower peer quality and these institutions better suit their learning needs. I show that match effects are more pronounced for boys with relatively low achievement at entrance, and that peer quality of the first-choice school follows similar patterns. Institutions with lower absolute performance may form a better match with low-ability males as these are found to react less productively than girls to new environments with higher-quality peers (Hastings et al., 2006; Kling et al., 2007; Deming et al., 2014; Murphy and Weinhardt, 2020). Consistently with this hypothesis, estimated effects are largest in schools in the bottom quartile of the peer quality distribution.

My results have important implications for educational policies. First, returns to sorting into specific schools implies that parental choice may benefit school productivity by improving the quality of the student-school match. Second, my results imply that returns to school inputs are heterogeneous across students. This has important consequences for school accountability systems, often based on value-added models implicitly assuming homogeneous school impacts.

To my knowledge, this paper presents unique evidence of achievement returns to sorting into the school of choice. Sorting of different workers into different firms is found to account for an important part of earnings inequality (Bonhomme et al., 2019). In contrast, Kline

and Walters (2016), Cornelissen et al. (2018), Walters (2018), and Abdulkadiroglu et al. (2020) find no or negative selection on achievement gains in pre-school programmes and high schools.⁵ My study also presents the first investigation of parental choice effects in primary schools. As educational decisions at early stages are crucial for student development (Chetty et al., 2011; Heckman et al., 2013), this fills an important gap in the literature.

A large array of studies has investigated impacts of attending high schools and colleges chosen by parents, finding at most moderate achievement effects (e.g., Pop-Eleches and Urquiola, 2013; Abdulkadiroglu et al., 2014; Hoekstra et al., 2018; Gorman and Walker, 2020). In a meta-analysis, Beuermann and Jackson (2020) find a small and statistically insignificant effect across studies in the literature. This paper uniquely isolates parental choice effect from returns driven by school effectiveness.

Finally, many studies leverage data on submitted rankings to investigate parental preference for school attributes (e.g. Hastings et al., 2009; Burgess et al., 2015; Glazerman and Dotter, 2017; Burgess et al., 2019; Ainsworth et al., 2020). I describe parental preferences accounting for the set of accessible schools, addressing recent concerns on truthfulness of reported rankings under DA (Fack et al., 2019). In a further step, I document substantial heterogeneity in parental rankings of observationally similar schools, and consistency of reported preferences with choice behaviour over time.

The paper unfolds as follows. Section 2 presents the institutional background of parental choice in London. Section 3 describes the data and the replication of centralised school assignment. Parental preference for school attributes is explored in Section 4, while Section 5 presents the empirical strategy. Results are discussed in Section 6, and Section 7 concludes.

2 Institutional context

Primary education and school choice in London

Primary education in England spans seven grades, from age 5 to 11 and is organised in three phases. Primary school starts with a reception year, when students turn 5, which concludes the Early Years Foundation Stage (EYFS). At the end of the reception year, students are

⁵Kirkeboen et al. (2016) document sorting on comparative advantage in the choice of field of study at university.

assessed against several learning goals to inform teachers and parents on their readiness for Year 1. The second phase is Key Stage 1 (KS1), spanning Year 1 and Year 2. At the end of KS1 students receive teacher assessments evaluating their achievement in mathematics, science, and English, separately for reading and writing. The final phase is Key Stage 2 (KS2), from Year 3 to Year 6, at the end of which students sit externally-marked standardised exams in mathematics and English. For all these phases, the National Curriculum sets core knowledge and achievement objectives.

I consider state-funded schools, the main provider of primary education. In England, in fact, only less than 5% of students opt for private primary institutions.⁶ Within state-funded schools, different types of institutions exhibit different degrees of autonomy from Local Authorities (LAs), the English school districts. Most frequent are Community schools, fully controlled and funded by the LA. Voluntary controlled and voluntary aided schools are established by private, mainly religious, organisations but are mostly funded by the LA and have limited autonomy. Finally, foundation schools and academies are the most independent state schools from the LA. Academies, similarly to US charter schools, are not bound by the National Curriculum and enjoy considerable autonomy in management.

Parental choice among state-funded schools is well-established in England. Since the 1980s, the open enrolment policy guarantees parents the right of choosing a school for their children, as long as demand does not exceed capacity. Parents are required to rank up to six schools at application, in order of preference. LAs have the statutory requirement to provide a school place to local children and assign applicants to the highest preference school available. Schools receive funding from LAs mainly based on enrolment count and are therefore incentivised to attract parental demand to fill capacity⁷.

The wide availability of data on school performance ensures comparability across institutions and sparks competition to enter schools with high absolute achievement. Parental choice is informed by school performance tables, published annually since 1996. They collect information on academic performance, both standardised test scores and value-added measures, and on intake composition of the schools. Institutions with excellent test scores are typically sought-after by parents and they easily become oversubscribed (see Burgess et al.,

⁶ Author's own calculation from official 2019 data on students count by school phase and sector, available at <https://www.gov.uk/government/statistics/schools-pupils-and-their-characteristics-january-2019>.

⁷ Primary schools have a statutory class size cap of 30 students.

2015, Burgess et al., 2019, and Section 4 below).

Admission criteria to oversubscribed schools have had important impact on gentrification and urban development. When demand exceeds capacity, applicants are mostly admitted in order of proximity, generating fierce competition in the housing market to secure residence close to preferred institutions. 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 width of catchment areas, however, varies year by year according to supply and demand for school seats. Therefore, parents are hardly able to precisely anticipate the location of catchment area boundary.

London is an ideal context to study school choice, with a dense supply of schools and high competition for popular institutions. The 33 LAs in Greater London form the most populated urban area in Europe. Primary schools are typically small, enrolling about 50 students per cohort, implying the average family has potential access to several schools. Absolute achievement at KS2 exams is higher than the national average and this difference is driven by a dense right tail of institutions serving exceptionally performing students, as almost 20% of London primary schools fall in the top decile nationally. About 70% of schools are oversubscribed and parents must target a narrow area to obtain admission into preferred institutions, the average catchment area among oversubscribed schools is just 600 meters wide. The fraction of students missing out on their top choices is systematically the highest around the country.⁸

School assignment

Assignment to school is centrally regulated by the School Admissions Code. Applicants are admitted to the school listed by parents as 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 typically prioritised and, within

⁸Aggregate statistics on school admission are publicly available at <https://www.gov.uk/government/collections/statistics-school-applications>.

priority groups, distance to school is used as tie-breaker to assign school offer. First, schools are required to give precedence to children with particularly disadvantaged backgrounds, a situation concerning a very low share of students.⁹ Second, applicants with siblings currently enrolled at the school are usually prioritised.¹⁰ Finally, exceptional admission criteria are permitted to religious schools, which typically set requirements based on faith. All applicants outside these categories have equal priority in admission.

School districts across England assign seats through a deferred acceptance mechanism (DA, Gale and Shapley, 1962) matching students to the highest preference school with available seats. Since 2007, DA is adopted nationwide for centralised school assignment after the previously popular Boston mechanism was banned. The latter, prioritising applicants who rank the school as first choice, has been proven more vulnerable to strategic preference reporting (Pathak and Sonmez, 2013). Intuitively, parents may rank a ‘safe’ school as first choice even if they would prefer a school where admission is less likely in order not to miss out on both institutions. DA algorithms do not suffer from this problem as school priority does not depend on parental preference. As long as parents act rationally, their ranking of schools reflects the true order of preference among listed institutions (Fack et al., 2019).

In particular, preferences, priorities and school capacities are mapped into offers through the student-proposing DA algorithm. Each student initially applies to the most preferred school. Applicants are ranked by priority and tie-breaker value and provisionally admitted up to capacity. In subsequent rounds, students who are rejected apply to the next-best school in their application form and are ranked jointly with applicants provisionally admitted up to this point. School retains applicants up to capacity and rejects the rest, who in turn apply to the next-best school. The algorithm stops when no rejection takes place. Some applicants may be left unassigned.¹¹

⁹Highest priority in school admission is given to children looked after by the LA, corresponding to the 0.5% of children under 18 years of age in London in 2019 (official counts are available at <https://www.gov.uk/government/statistics/children-looked-after-in-england-including-adoption-2018-to-2019>). In addition, priority is usually granted to children with a statement of special education needs, 0.8% in my working sample. The two groups are not mutually exclusive.

¹⁰Burgess et al. (2019) reports that in the Millennium Cohort Study, a British longitudinal study including a detailed parental survey, the 43% of children has a school-age sibling at the time of admission to secondary school. This reportedly varies substantially with family income, from 33% to 67% in bottom and top income decile, respectively.

¹¹Students disqualified from all preferred schools are assigned by the LA to an institution with spare capacity. This happens to about 4% of applicants in my sample.

Parents across the country receive a single school offer in mid-April, deemed National Offer Day. Parents who are unsatisfied with the assignment can join waiting list at preferred schools with the same priority, and may obtain admission if applicants with offer give up their place. Finally, parents have the right to appeal the offer decision in case of irregularities, though admission outcome is rarely overturned¹².

3 Data

I exploit administrative data on applicants to state-funded primary schools in London in 2014 and 2015. Individual-level records include rank-order lists of schools submitted by parents to LAs and the school offered to each applicant as a result of the assignment mechanism. Data on applications 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 student postcode of residence, a granular information on residential location spanning an average of 15 properties and often corresponding to one single building in London. To measure proximity to school, I compute the linear distance from each applicant’s postcode to all primary schools around.¹³

Assessments at the end of KS1 are the outcomes considered in my empirical investigations. Students are assessed by teachers at age 7, after three years of primary school. Results are grouped in three categories, depending on student achieving below, at, or above the expected standard, corresponding to Level 2 in the National Curriculum. Three different subjects are assessed – English, separately for reading and writing, and mathematics.¹⁴ Though 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).¹⁵ Importantly, students sit national tests in

¹²Among the 688 London primary schools with appeal data in 2015 (about 40% of the total), the 95% recorded no appeal resolved in parents’ favour.

¹³I compute distance using centroids coordinates for English postcodes obtained from www.doogal.co.uk. For applicants with missing postcode (about 3%), I impute distance 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.

¹⁴Students are also tested in science, but this assessment is not very informative since the 83% of students in my sample are judged as “working at the expected standard”. I do not consider this subject in my analysis.

¹⁵Moderation is monitored by the Standards and Testing Agency (STA) and involves a visit from an external moderator. The moderator reviews a sample of students’ classwork on which the assessment was

mathematics and reading and the end of KS1, with an optional writing test, which scores are not disclosed but are meant to inform teacher assessments. Burgess and Greaves (2013) find almost 80% of students are awarded the same achievement level in teacher assessments and standardised tests at the end of primary school (KS2), suggesting that teacher judgement is broadly in line with test scores.¹⁶ Overall, institutional details and empirical evidence suggest that KS1 assessments provide a reliable and comparable measure of student achievement. To control for academic ability at entrance, I consider Early Years Foundation Stage Profile (EYFSP) assessments. They test students against 17 learning goals and are completed during the reception year (the year a pupil turns 5), when students have just entered compulsory education. Similarly to KS1 assessments, EYFSP results are grouped in three categories, depending on student achieving below, at, or above expected standards in each learning goal.

I observe detailed baseline characteristics of students that serve as control variables in my analysis. Individual-level records include gender, free lunch eligibility, special education needs, language, and ethnicity group. Moreover, indexes computed at the local area (LSOA) level enrich the range of socioeconomic traits observed. First, the income deprivation index (IDACI) measures the proportion of children in income deprived families in the local area and is included in administrative records. In addition, I merge NPD data with socioeconomic local area characteristics from the 2011 population Census, such as the proportion of adult residents achieving qualifications at the higher education level.

Sample description

I consider students entering the reception year between 2014 and 2015 who ranked at least one London primary school at application. The working sample consists of 199,220 applicants and 638,756 student-preference observations, with the average applicant ranking between 3 and 4 schools. As presented in Appendix Table A.1, primary schools in London serve a population of students whose social background is strikingly mixed – about 41% are white and a similar fraction does not speak English at home, compared to 78% and 12% in the rest of England,

based. Moderation can result in changes to teacher assessments or, in case of systematic lack of evidence on teacher judgements, in school being reported for maladministration.

¹⁶Burgess and Greaves (2013) also find evidence of bias in teacher assessments based on ethnicity. I observe student ethnic group and control for this variable in my analysis.

respectively. London students are more likely to have a disadvantaged background, with higher proportion of students eligible for free lunch or with special education needs than the rest of the country. Despite this difficult context, primary schools in the capital outperform the average national institution in terms of academic achievement, with higher proportion of students achieving above expectations at Year 2 in all subjects assessed. Notably, the gap in favour of students in London is wider in Year 2 than at primary school entrance.

Primary school admission is substantially more competitive in London than in the rest of England, as reported in the bottom part of Appendix Table A.1. Parents in the capital exercise choice more actively, being 15 percentage points more likely to rank three schools or more at application¹⁷. The proportion of applicants admitted to the first choice is about 82% in London, around 7 percentage points lower than in the rest of the country. Parents closely comply with centralised school assignment. Take-up rate is very high, with 87% of students enrolled in the offered institution at the reception year (see Appendix Table A.1, column 1). This fraction is 3 percentage points lower than the national average (column 2), partly reflecting higher propensity to enrol at private schools among families in the capital (4% of parents choose private school against 2% in the rest of England).¹⁸

Replication of centralised school assignment

Centralised assignment breaking ties by distance implies that, if a school is oversubscribed, no offer is granted to applicants located further than a specific threshold. Such a threshold, however, is not observed as administrative data do not track the admission process. School assignment depends also on parental preference and school priorities, which interact with distance to determine school offer (see Section 2 above). First, I show that school offer rate is not entirely driven by distance and that admission cut-off cannot be directly inferred from available data. I then replicate the assignment mechanism to trace catchment boundaries and identify applicants at the margin for admission.

Appendix Figure A.1 shows that school admission rates are not deterministic conditional on distance. The probability of receiving an offer markedly decreases with distance to school,

¹⁷The comparison on propensity to rank six schools is not presented as in most English districts application form is restricted to three schools.

¹⁸I consider a student as enrolled at private school if not tracked into any state-funded school after application.

and the figure is very similar for enrolment, represented by diamonds in Appendix Figure A.1. However, rather than dropping to zero, admission rates gradually diminish, and this happens for two reasons. First, there is variability in parental ranking of the school. Regardless of distance, applicants ranking the school lower than first choice are offered a place only if they miss out on all institutions ranked with higher preference. For example, this explains why offer rate of applicants in the bottom decile of distance is far from deterministic, at about 0.7 in Appendix Figure A.1.¹⁹ Second, particular categories of applicants, as detailed in Section 2 above, are admitted with priority independently from their location. For example, this partly explains why offer rate of applicants in the top decile of distance to school is non-negligible, at about 0.2 in Appendix Figure A.1.

Replication of school assignment is complicated by data availability, as I have no information on demographics that are relevant to define school priorities over applicants. The catchment boundary, defined as the distance to school of the last student admitted²⁰, is not identified when replicating assignment based solely on parental preference and distance.²¹ Students with priority, however, are partially detectable in the data. Intuitively, if an applicant with offer lives beyond the distance threshold estimated without considering school priorities, she must have precedence in admission. I show in the Appendix how I achieve replication of school offer based on this idea.

The main limitation in my empirical analysis is that priority status – mostly involving siblings of current students²² – remains partially unobserved.²³ I infer priority only when binding for admission, implying that students with precedence are undetected, first, at schools other

¹⁹Consistent with this expectation, Panel B of Appendix Figure A.1 shows that about 35% of parents located next to the school have ranked it less than first choice.

²⁰Admission cut-off is defined only for oversubscribed schools.

²¹I replicate centralised assignment by running DA for all schools at the same time. In practice, however, 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.

²²In addition, I do not observe parental faith, often used to grant priority at religious schools. As the error in catchment boundary measurement is likely more serious in this case, I do not consider religious schools in estimation.

²³A minor limitation is that distance to the catchment boundary, which I use as running variable, is measured with error if the last admitted applicant, which defines the distance cut-off, enjoys admission priority. In that case, the correct cut-off is distance of the applicant located immediately closer to the school. As the median school has 55 seats and a catchment boundary of about 650 meters, density of applicants ensures the measurement error, when present, is likely small. Moreover, measurement error in the catchment boundary is constant for all applicants and cancels out when comparing students around the cut-off.

than the offered one and, most importantly, when located within the catchment boundary. Unobserved priority is unlikely to constitute a major concern for my analysis for two reasons. First, only a minority of students at the right-hand side of the catchment boundary (about 25%) are flagged with priority. Second, in order to bias my results, applicants with unobserved priority would need to display differential potential outcomes conditional on parental preference and distance to the school. Siblings of current students, who obtained a place at the school within the past six years, are likely comparable to peers with similar school ranking and residential sorting. The fact that a number of student characteristics associated with potential outcomes, including lagged achievement, are balanced around the catchment boundary (see Section 5 below) supports the validity of my design. Nevertheless, I discuss in the Appendix that my measure of priority is broadly in line with a proxy for siblings at the school based on location of students in older cohorts and that, when excluding applicants within catchment boundary who are likely to enjoy priority, results are substantially unaffected.

Distribution of parental rankings around the catchment boundary suggests applicants are not able to exactly anticipate the admission cut-off. Figure 1 shows parental preference for the school (Panel A, where a value of 1 indicates first choice) and the share of parents ranking the school first (Panel B) are continuous around the catchment boundary. In line with evidence on parental preferences presented in the next Section, the figure displays decreasing parental preference with distance to school. The decrease accelerates, on average, starting at a slightly shorter distance than the catchment boundary, suggesting that parents adjust their application behaviour based on their expectation about the cut-off realisation. However, the graph shows no discontinuity, implying that the exact location of the catchment boundary, as expected, is likely unpredictable by parents.²⁴

²⁴Appendix Figure A.5, comparing catchment area boundary for the same school across the two years considered, further reinforces this expectation. While catchment areas are fairly stable over time (median change is zero, and within 180 meters for half of institutions), variation in the distance cut-off has no predictable direction and is arguably hard to anticipate.

4 Parental preference for schools

Ranked schools

I begin by describing parental preferences for geographical proximity, peer quality and school effectiveness. Under DA, submitted rankings reflect the true preference order among listed schools as long as parents act rationally (Fack et al., 2019). I plot average attributes of listed institutions by parental rank, conditional on feasibility and number of preferences expressed. I also explore differential preference for school attributes by socioeconomic status.²⁵ Specifically, I estimate the following regression:

$$A_{is} = \gamma_1 + \sum_{s=2}^6 \gamma_p \mathbb{1}(s = p) + X'_{is} \delta + u_{is}, \quad (1)$$

where A_{is} is the attribute of school ranked s -th by student i . The vector of controls X'_{is} includes number of preferences fixed effects, an indicator for ex-post feasibility of school ranked s -th by student i , and school attributes other than A_{is} (e.g., school value-added and distance when considering peer quality). In this formulation, parameters γ_1 to γ_6 estimate average attributes of schools ranked first to sixth, conditional on controls.

Parents are surrounded by several schools at short distance from residence and they rank them in order of proximity.²⁶ Panel A of Figure 2 plots distance to school by parental rank using predicted values from equation (1), separately for applicants with local area deprivation above or below the median. First choice of parents in better-off areas is on average around 800 meters from residence, and all ranked schools are located within 1.2 km. The corresponding figure in more deprived contexts is very similar, with slightly shorter distances on average. This difference could reflect either higher utility cost of travel for disadvantaged families or supply-side differences such as higher population density in worse-off neighbourhoods.

Though all parents rank schools by peer quality, those in relatively deprived areas demand institutions with lower absolute achievement. The left-hand graph in Panel B of Figure 2 plots standardised final test scores at the school, measured at the time of application.²⁷ Peer

²⁵Hastings et al. (2009) find parents from disadvantaged contexts exhibit weaker preference for academic performance.

²⁶Here and below, I consider applicants within 2 kilometers from ranked school, corresponding to the 90th percentile of the distance to school distribution.

²⁷I measure peer quality by school average standardised test scores at final year, which I compute pooling

quality of the first choice ranked by relatively advantaged parents is one standard deviation (hereafter, σ) above the average, and it markedly decreases with parental rank, indicating that parents value absolute achievement. This result likely reflects the dissemination of school performance tables, published annually to inform parental choice, in which test scores are headline measures. A similar pattern is observed in worse-off areas, but peer quality of first choice is substantially lower, by about 0.8σ , and similar to the score of the sixth choice in relatively affluent neighbourhoods. This stark difference likely reflects segregated access to top-scoring schools through residential sorting.

Conditional on distance and peer quality, parents do not respond to school value-added. Panel C of Figure 2 plots school value-added. I follow Deming et al. (2014) and estimate a baseline measure of school value-added as regression-adjusted test scores growth, averaged at the school level.²⁸ Estimated school value-added has about 17% correlation with absolute achievement, suggesting high-scoring schools are not necessarily highly effective. After controlling for peer quality, parental rankings are roughly orthogonal to school value-added, in line with findings in other contexts (see MacLeod and Urquiola, 2019 for a review).

Parental preferences result in excess demand for schools with high peer quality. Attributes of oversubscribed schools are described in Appendix Table A.3. I define a school as oversubscribed if the number of applicants missing out on any higher-preference institution exceeds capacity. In fact, these are the only candidates who would receive an offer if school capacity marginally increased. School seats are often rationed, with the 60% of institutions experiencing excess demand.²⁹ Oversubscribed schools have about one σ higher peer quality, suggesting strong reaction of parental demand to absolute achievement. In line with evidence from preference data, oversubscription is substantially less responsive to school effectiveness.

9 cohorts of data up to 2014. Scores are averaged across mathematics and reading.

²⁸Specifically, I compute school average residuals from an individual-level regression of KS1 assessments on student socioeconomic characteristics and baseline achievement. One σ higher school value-added improves the probability of scoring above standards at Year 2 assessments by about 7 percentage points (28% of the sample average). The value-added distribution is plotted in Appendix Figure A.6 (see footnote to the figure for estimation details).

²⁹I consider here schools oversubscribed by at least 5 seats (results are robust to this choice). Institutions oversubscribed by one seat or more are the 69%. I also note that a significant fraction of schools has an impressive degree of excess demand, with the 37% of institutions oversubscribed by 20 seats or more (versus an average enrolment count of about 50).

Feasible schools

Following [Ainsworth et al. \(2020\)](#), I compare attributes of the school where applicants enrol (one of the top two choices for most applicants) with those at other feasible institutions. This complements the description of parental preferences by outlining the supply of schools. As argued in [Fack et al. \(2019\)](#), parents foreseeing admission chances can ‘skip the impossible’ and give up application to preferred schools located too far away. Therefore, understanding parental preference requires to account for the distribution of ex-post feasible schools faced by each applicant. I define the individual feasible school set as the collection of schools, to which the student may or may not have applied, which would have been accessible based on distance (see the Appendix for details). Feasible set of the average applicant includes 6 schools within 2 kilometers from residence. The large majority of parents has some degree of choice, with 75% potentially accessing at least 3 schools.

Parents from different socioeconomic contexts travel very similar distances to primary school, about 600 meters. Panel A of [Figure 3](#) depicts average distance to student’s school and to the closest feasible institution by decile of local area deprivation, alongside the average in the feasible set.³⁰ Parents choose schools at relatively short distance from residence, with 59% of applicants enrolled in the closest accessible institution. On average, parents give up schools closer to residence by about 200 meters, likely trading off distance with other valued attributes. Interestingly, this difference is lower for applicants in wealthiest areas, likely choosing residence close to desired institutions.

Most of the difference in peer quality across local deprivation is accounted for by residential sorting. As shown in Panel B of [Figure 3](#), the average feasible school in wealthiest areas has about 1.2σ higher peer quality than in most deprived neighbourhoods, and schools where students enrol exhibit a comparable difference. Affluent parents leave very little in terms of peer quality, being systematically enrolled in schools with absolute achievement close to the highest available. Applicants at the top of deprivation distribution, instead, would be able to access institutions with about 0.5σ higher peer quality. This probably reflects steeper trade-offs between school test scores and distance for relatively disadvantaged parents, as they enrol at school with similar distance from residence than better-off peers.

Regardless of socioeconomic background, parents miss out on accessible schools with

³⁰Graphical representation here follows [Figure 2](#) in [Ainsworth et al. \(2020\)](#).

highest value-added. Students could potentially access institutions with 0.6 to 1σ higher effectiveness (see Panel C of Figure 3). The figure suggests that parental preferences are remarkably more consistent with maximisation of peer quality rather than value-added, in line with findings by Ainsworth et al. (2020). If returns to school were homogeneous, this would correspond to 0.11-0.19 σ foregone achievement by Year 2. However, I argue in what follows that there is substantial heterogeneity in both returns to school and parental preferences.

Heterogeneity and consistency over time

Average patterns mask substantial disagreement in parental rankings. As shown in Figure 3, distance and peer quality explain just about 40% of the variation in parental preferences. The figure is substantially unchanged when adding school value-added, school type, and peer composition. Parental rankings vary substantially even conditional on unobserved school traits, as adding school fixed effects explains just 50% of the variation.

Nevertheless, analysis of compliance with assignment suggests that parents behave coherently with preferences expressed at application. The share of parents enrolled into the assigned institution at the reception year is very high on average (87%), and it strongly increases with preference for the school offered (see Appendix Figure A.7). Blue bars in Panel A show that compliance rate exceeds 90% for students offered their most preferred school, and it drops to 50% for applicants assigned to their sixth choice. Non-complying students enrol mostly at state schools not ranked at application (6%) or private institutions (4%), as shown in Panel B.³¹ The 2% of students enrolls at a school with higher parental preference than the one assigned, most likely through waiting lists. Interestingly, almost no applicant (less than 1%) enrolls at a school with lower parental preference than the offered institution.

A similar conclusion is reached when considering student mobility in the first years of primary school. First, the likelihood of moving children to a different institution by Year 2 decreases with preference for the offered school, as shown in Panel A of Appendix Figure A.7 (red bars). Residential mobility, in contrast, is roughly orthogonal to school assignment (green bars).³² Second, a negligible share of students move to a school ranked with lower

³¹Likelihood of enrolment to private schools discontinuously increases beyond the catchment boundary (see Appendix Figure A.9). I argue in the Appendix that selective attrition does not invalidate my results.

³²I define residential mobility as an indicator variable equal to one if a student's home postcode changes. Given the evidence on residential sorting discussed in Section 4 above, this result suggests that parents who

parental preference. Virtually all students who move to another institution enrol either at a state school not listed at application (12%) or to a private school (4%), as shown in Panel C. Very few students move to a school with higher parental preference, suggesting that centralised assignment is well enforced. Third, the decision to move to a different school responds to peer quality and distance, similarly to preferences stated at application (see Section 4 above). I show in the Appendix that student mobility is more likely when the assigned institution falls short of the school of choice in terms of absolute performance and proximity to residence, but it does not depend on school value-added.³³

Overall, though parents strongly respond to peer quality and distance, school rankings are highly heterogeneous and yet represent solid preferences, guiding parental choice behaviour even after several years. The question arises whether such idiosyncratic choices have an impact on students learning, reflecting sorting based on specific student-school match.

5 Empirical strategy

Research design

Under DA assignment, school offers depend solely on parental preferences, school priorities, and distance to school (see Section 2 above). Controlling for these variables is therefore sufficient to eliminate selection bias from residential sorting and application choice. With oversubscription, the distance between schools and place of residence is used as tie-breaker among applicants with equal priority. I exploit the idea that, depending on preferences and priorities, a subset of applicants is as good as randomly assigned near the distance cut-off. My identification strategy follows [Abdulkadiroglu et al. \(2014\)](#) and builds on two steps. First, I isolate the sample of applicants for which the tie-breaker is binding for admission. Second, I compare students located around the school catchment boundary.

Consider the case of parents ranking three schools, labeled A, B and C. To fix ideas, assume that these are the only schools which are potentially accessible to parents in a certain neighbourhood. I begin by considering all applicants for which school A is the most preferred.

are willing to move their residence secure location close to desired schools before the assignment takes place rather than moving in response to the assignment.

³³I refer to the Appendix for discussion on the interpretation of my estimates in presence of school mobility induced by school offer.

These applicants can be grouped into those ranking B second and C third (ABC), and those ranking C second and B third (ACB). I consider applicants to school A with equal admission priority and living around the catchment boundary, regardless of how they ordered less-preferred schools at application. At the catchment boundary, the chance of receiving an offer from school A is independent of the application choice, as DA does not consider parental preferences other than those at A. Comparing students at the boundary, therefore, is sufficient to eliminate selection bias from application choice.

Similar reasoning can be applied to schools ranked by parents lower than the first choice. Consider now the case of parents ranking school A as second choice. These applicants are of two possible types: those ranking B first and C third (BAC), and those ranking C first and B third (CAB). I consider applicants living around the catchment boundary of school A who have equal admission priority and live outside the catchment of their first choice, regardless of how they ordered other institutions at application. In general, following [Abdulkadiroglu et al. \(2014\)](#), I consider sequential samples of applicants to school A: those ranking A first; those ranking A second and excluded from their first choice; those ranking A third and excluded from their first and second choice, and so on.

Within these samples, students are sharply assigned by distance, as visualised in Panel A of Figure 5.³⁴ The figure depicts offer rates in 100-meters-wide bins of distance to the catchment boundary, which I employ as running variable in my analysis, and a local linear polynomial fitted to underlying observations. The left-hand graph represents equal-priority applicants to the most preferred institution, pooling cut-offs at all first-choice schools. Admission rate sharply drops at the cut-off as school capacity is reached, generating exogenous variation in assignment. A similar design is displayed for lower-ranked schools. Central and right-hand graphs in Panel A of Figure 5 plot offer rate of equal-priority applicants demanding schools with lower preference, second and third to sixth respectively, conditional on being excluded from any more-preferred school. Note that the same applicant may be located around cut-off of more than one school if excluded from the first choice. School assignment closely corresponds to enrolment at the reception year, as shown in Panel B of Figure 5. The

³⁴I consider in estimation applicants to oversubscribed (See Section 4 above for definition) and non-faith schools. Appendix Table A.2 reports descriptive statistics for the estimation sample in columns (3)-(4). Comparison with columns (1)-(2) shows students around the catchment boundary of oversubscribed schools have similar characteristics to the population of applicants in London, have moderately better achievement and are slightly less likely to be eligible for free lunch.

assignment mechanism implies that students excluded from the school of choice enrol to an institution with lower parental preference, as can be seen in Panel C of Figure 5 – where a rank of 1 indicates first choice. For example, about 70% of applicants denied the first choice are offered a seat in the second or third most preferred schools.

My research design builds on recent methodological contributions on how to leverage centralised assignment for empirical research. Abdulkadiroglu et al. (2017, 2019) argue that DA assignment embeds as good as random variation in school offer conditional on parental preferences and school priorities, labeled parental “type” to indicate that these are likely correlated with potential outcomes. However, full-type conditioning is often not feasible: in the sample I consider, for example, there are almost as many types as the number of applicants. They show, first, that conditioning on the ex-ante probability of receiving an offer is sufficient to control for parental type and, second, that the risk of admission is much coarser than type and depends on few key assignment variables. In my empirical analysis, I eliminate selection bias by conditioning on the components of parental type that are relevant for assignment risk.³⁵

Empirical framework

Following Abdulkadiroglu et al. (2020), potential outcome of student i at school ranked s -th can be written as:

$$Y_{is} = \nu_i + \alpha_s + \mu_{is}, \quad (2)$$

where ν_i is the student’s general ability, α_s is the school value-added, i.e. its average causal impact on achievement, and μ_{is} is the idiosyncratic match between student i and the school ranked as s -th choice. In a model where parents sort on their children’s comparative advantage in achievement production (Roy, 1951), μ_{is} is expected to be positive.

Let D_{is} be a dummy variable indicating enrolment at the school ranked s -th. The outcome I observe for student i can be written as:³⁶

³⁵Computation of the admission propensity score derived in Abdulkadiroglu et al. (2019) is complicated by imperfect observation of admission priorities (see Section 3 above). Nevertheless, I discuss in the Appendix how to adapt their formula to my institutional context, and I show that main conclusions hold when using the propensity score specification as a sensitivity check.

³⁶I assume here for simplicity that all students rank six schools and are admitted to one of the listed institutions. I also assume that all offers are accepted. Notation could be extended to deviate from this setting but would get more cumbersome.

$$Y_i = Y_{i1} + \sum_{s=2}^6 D_{is}(Y_{is} - Y_{i1}).$$

Substituting into equation (2), we obtain:

$$Y_i = \nu_i + \alpha_1 + \sum_{s=2}^6 D_{is}(\alpha_s - \alpha_1) + \sum_{s=2}^6 D_{is}(\mu_{is} - \mu_{i1}) + \mu_{i1}. \quad (3)$$

Let \mathcal{B}_{i1} denote distance of applicant i from the the most preferred school's boundary (a similar reasoning applies to lower-ranked choices). I pool applicants at all first choices and consider the following parameter:

$$E[Y_i|\mathcal{B}_{i1} = 0^-] - E[Y_i|\mathcal{B}_{i1} = 0^+]. \quad (4)$$

This quantity represents the comparison of outcomes of students living marginally within the catchment boundary ($\mathcal{B}_{i1} = 0^-$), who are admitted at their first choice, with those located marginally outside, who are excluded ($\mathcal{B}_{i1} = 0^+$). The causal parameter identified by this comparison is the effect on student outcomes of missing out an offer from the most preferred school.

The comparison of students around cut-off for admission rests on the assumption that the catchment area boundary cannot be perfectly anticipated by parents. The admission cut-off changes over time depending on the number and parental rank of applications to the school, priorities at the school and changes in density of school-age children in the neighbourhood (e.g., because of newcomers, see Appendix Figure A.5), rendering exact sorting with respect to catchment boundary extremely unlikely (see Figure 1). This motivates the following continuity condition at the catchment boundary:

$$E[\nu_i|\mathcal{B}_{i1} = 0^-] = E[\nu_i|\mathcal{B}_{i1} = 0^+],$$

which implies that $\mathcal{B}_{i1} = 0^-$ and $\mathcal{B}_{i1} = 0^+$ students have the same ability, on average. Furthermore, we have that:

$$E[\alpha_1|\mathcal{B}_{i1} = 0^-] = E[\alpha_1|\mathcal{B}_{i1} = 0^+].$$

Finally, I assume a third continuity condition:

$$E[\mu_{i1}|\mathcal{B}_{i1} = 0^-] = E[\mu_{i1}|\mathcal{B}_{i1} = 0^+].$$

If parents rank preferences based on expected returns to achievement, I do not expect any

systematic difference in the match component at the boundary considering that applicants have ranked the same first choice.

Substituting the definition of Y_i from (3), and applying the continuity conditions, we obtain the following expression for the comparison in (4):

$$\sum_{s=2}^6 E[D_{is}(\alpha_1 - \alpha_s)|\mathcal{B}_{i1} = 0] + \sum_{s=2}^6 E[D_{is}(\mu_{i1} - \mu_{is})|\mathcal{B}_{i1} = 0].$$

For example, if all students missing out on the first choice were offered the second choice, the comparison at the boundary would be equal to:

$$E[\alpha_1 - \alpha_2|\mathcal{B}_{i1} = 0] + E[\mu_{i1} - \mu_{i2}|\mathcal{B}_{i1} = 0]. \quad (5)$$

I am interested here in the second element of the this equation, capturing whether parents rank schools based on specific match with their children.

To isolate the parameter of interest, I rely on baseline estimates of school value-added (see Section 4 above) and implement the comparison in equation (4) conditional on effectiveness of ranked institutions. Ideally, I would compare students enrolling at schools of exactly the same quality:

$$E[Y_i|\mathcal{B}_{i1} = 0^-, \alpha_1 = \alpha_2] - E[Y_i|\mathcal{B}_{i1} = 0^+, \alpha_1 = \alpha_2]. \quad (6)$$

In practice, I use value-added estimates as proxy for α_s and consider the sample of students for whom the school of choice and the next-best school in their preferences have similar value-added.³⁷ As applicants who are denied their school of choice most likely enrolls at their next-best preference, such difference represents the expected value-added gain in attending the desired institution. This comparison rests on the assumption that school quality is precisely measured by value-added estimates. In the next Section, I provide support for this assumption by showing that value-added estimates predict causal effects on achievement.³⁸

As a sensitivity check, I use an alternative research design exploiting heterogeneity in parental rankings. I compare students assigned to the same school, ranked by their parents

³⁷In my preferred specification, I consider students with the same value-added decile. Columns (5)-(6) of Appendix Table A.2, reporting characteristics of students in this sub-sample, show they are observationally similar to the full estimation sample (columns 3-4), apart for moderately higher achievement.

³⁸Furthermore, heterogeneous returns to school imply that value-added estimates are not necessarily comparable across institutions as they reflect achievement growth of a potentially selected sample of students. Comparison of schools with similar value-added is valid under the assumption that selection on gains is not different across the institution of choice and the next-best school in student preferences.

with different preferences. Operationally, I estimate discontinuities at the catchment boundary with school fixed effects. This specification compares students just admitted to the school of choice to applicants whose next-best preference is the same school and are just excluded from their institution of choice. The benefit of this design is that it fully controls for school quality by holding the institution constant. However, as I compare students of different type based on their preferences, this comes at the cost of a stronger identifying assumption. Estimates can be interpreted as causal effects as long as, conditional on distance to their schools of choice, students with different preferences have the same potential outcomes.

Estimation and covariate balance

Consider the sample of students: (a) for which school j is the s -th listed in the application ($s = 1, \dots, 6$), (b) without priority at school j , and (c) not admitted to any of the schools preferred to school j . Students in this sample are indexed to i , and the school ranked as s -th choice is indexed by j . Let Z_{is} be the indicator for receiving an offer from school ranked s -th. I start by testing covariate balance around the catchment boundary in the sample defined using (a), (b), and (c) above and depending on the value of s . I consider the following specification:

$$W_{is} = \pi_{0j(i,s)} + \pi_1 Z_{is} + f(\mathcal{B}_{is}) + u_{is}, \quad (7)$$

where W_{is} is a baseline characteristic of student i applying to the s -th choice, and $\pi_{0j(i,s)}$ is a full set of school of choice fixed effects. I control non-parametrically for \mathcal{B}_{is} , denoting distance to the catchment boundary of the s -th choice, by including a linear trend estimated separately on each side of the cut-off and by considering kernel-weighted estimates of equation (7).³⁹ To increase precision, I add a full set of number of schools listed fixed effects and individual socioeconomic characteristics other than W_{is} . As this specification stacks applications with different preferences, I control for parental rank fixed effects and interact parental rank with running variable controls. Standard errors are clustered at the student level.

Considering applicants at risk of admission, socioeconomic characteristics of students at the two sides of the catchment boundary are balanced. Table 1 compares uncontrolled

³⁹I employ a triangular kernel centered at the boundary and select optimal data-driven bandwidth following Calonico et al. (2014), separately for each outcome variable. Local linear polynomials are recommended by the econometric literature on RD designs as opposite to higher-order polynomials of the running variable (Gelman and Imbens, 2019).

differences in covariate means by school admission, reported in column 1, with estimates of π_1 from equation (7), in columns 2 to 4.⁴⁰ Applicants receiving an offer from the school of choice are disproportionately less likely to be eligible for free lunch, are more likely white, live in neighbourhoods with lower deprivation and score higher at reception year assessments. Once distance to the catchment boundary is controlled for, differences are substantially smaller and generally not statistically significant. The results are in line with the idea that, conditional on offer risk, admission to school is as good as randomly assigned and provide evidence in support of the continuity conditions imposed above.

The causal effect of attending the school of choice is estimated via 2SLS by instrumenting D_{is} , a dummy indicating school enrolment at the s -th preference, with Z_{is} . The main outcome of interest, denoted by Y_{is} , is an indicator for scoring above expected standards at Year 2 assessments. I consider the following specification:

$$Y_{is} = \beta_{0j(i,s)} + \beta_1 D_{is} + f(\mathcal{B}_{is}) + \epsilon_{is}, \quad (8)$$

where notation and control variables follow equation (7). β_1 in equation (8) corresponds to the comparison in equation (4), and estimates the local average treatment effect (LATE) of attending the school of choice vis-à-vis an institution ranked with lower parental preference (see Panel C of Figure 5). It measures achievement gains from attending the school of choice for compliers, i.e. students who would enrol at the school only if offered a seat. LATE is a policy-relevant parameter in my context as it represents the expected impact of a marginal increase in school capacity. The corresponding first stage of equation (8) is:

$$D_{is} = \alpha_{0j(i,s)} + \alpha_1 Z_{is} + f(\mathcal{B}_{is}) + \eta_{is}, \quad (9)$$

where the parameter α_1 provides an estimate of the average discontinuity in school enrolment around the catchment boundary, visualised in Panel B of Figure 5. Finally, estimation relies on the assumption that receiving an offer can only shift students into the school of choice, regarded as monotonicity (Imbens and Angrist, 1994). In my context, the presence of defiers, i.e. applicants who would enrol at the school of choice only if denied an offer, is unlikely and the existence of this group can be reasonably ruled out.

⁴⁰Columns 3 and 4 add fixed effects of next-best school in student preferences and area of residence, respectively.

6 Effects of attending the school of choice

Returns to sorting

The school offer instrument generates a strong first stage for enrolment at the school of choice. Estimates of α_1 in equation (9) are reported in Panel A of Table 2. The first stage is about 63%, corresponding to the average discontinuity in school enrolment across boundaries of preferred schools (see Panel B of Figure 5). The figure shows that the largest part of non-compliance arises beyond the catchment boundary, as nearly excluded applicants find a seat in the school of choice despite not obtaining an offer in the first place. The pattern is consistent with the possibility of joining waiting lists.⁴¹

Attending the school of choice, on average, leads to a moderate achievement increase in mathematics with respect to institutions with lower parental preference. Instrumental variable estimates of β_1 in equation (8) are presented in Panel C of Table 2. Estimated LATE on assessments in mathematics is about 0.09σ in column (1), and remains similar when adding controls for next-best choice fixed effects (i.e., the second choice when most preferred school is considered, and so on, in column 2) and neighbourhood of residence (column 3). Estimated effects on reading and writing are small and not statistically different from zero, and the same result is obtained when stacking achievement outcomes across all subjects.⁴²

I find positive effects of attending the school of choice above and beyond the impact of school quality. Table 3 report estimates from a specification similar to equation (8) augmented with an interaction between school enrolment dummy and the difference in value-added between the school of choice and the next-best institution in parental preferences. Results in columns (1)-(3) in Panel B show that attending the school of choice increases student achievement in mathematics by about 0.09σ with respect to an institution of similar value-added but with lower parental preference. Results are unchanged when controlling for expected gains in school value-added quintiles, deciles or ventiles (see columns 1, 2, and 3,

⁴¹These applicants are “always takers” in the heterogeneous treatment effects jargon. Conversely, a small fraction of students does not enrol at the school of choice even if offered a seat. These are “never takers”, who prefer seeking a different state school or a private institution.

⁴²Results are in line with findings in the literature, summarised by [Beuermann and Jackson \(2020\)](#), documenting a small average effect of attending the school of choice, and not statistically significant. These estimates, however, combine average school value-added and student-school match effect at the institution of choice (see Equation 5).

respectively).

Results suggest that school quality is credibly accounted for in this comparison. First, value-added estimates predict causal effects on achievement. Estimates of interaction coefficients in Table 3 are positive and generally strongly statistically significant. Moreover, they decrease proportionally with the width of value-added bin considered. Averaging across subjects, one quintile of expected gain in school value-added is associated with 0.02σ causal achievement gain, while this estimate shrinks to 0.01σ and 0.006σ when considering deciles and ventiles, respectively (see columns 10, 11, and 12, Panel B). Second, comparing students with similar expected value-added gains does not capture the effect of school quality. Columns (13)-(15) of Table 3 show the effects of entering the school of choice on value-added of the institution where student enrolls. Average value-added gain is positive among students with next-best preference in the same value-added bin of the school of choice, but the gain is substantially lower than the causal effects on achievement estimated for this sub-group. Moreover, value-added gains estimated by interaction coefficients are substantially higher, and yet are associated with lower causal achievement gains than the one estimated holding value-added constant.⁴³

Nevertheless, I show that my conclusions hold when using a complementary research design that compares students within the same school. I exploit heterogeneity in parental preferences and estimate β_1 in equation (8) including fixed effects for the school where student enrolls. I find positive effects for students just admitted to the school of choice with respect to students enrolling at the same institution, ranked with lower preference. Instrumental variable estimates, reported in Columns (4)-(6) of Table 2, show a statistically significant effect of about 0.16σ in mathematics. Difference with results in Table 3 are driven by lower first stage, as reduced form results are very similar (see Panel A and Panel B, respectively).

Estimation results are robust to different empirical specifications. Appendix Table A.5 presents estimates of match effects of attending the school of choice obtained when opting for different parametric choices of running variable controls.⁴⁴ Sample is restricted to applicants located within 0.5, 0.8 or 1 kilometre from the catchment boundary, and controlling for

⁴³These results suggest that school value-added estimates strongly but not perfectly predict causal achievement gains, in line with the literature on validation of OLS value-added models (Chetty et al., 2016; Angrist et al., 2016, 2017).

⁴⁴Parametric and non-parametric estimates provide mutually reinforcing specification checks, as recommended by the literature on RD design (Lee and Lemieux, 2010).

quadratic (columns 1-3) or cubic (columns 4-6) polynomial functions of distance to the catchment boundary. Estimates holding school value-added constant, from specifications similar to column (2) of Table 2, are presented in Panel A. Estimates from the within-school design, from specifications similar to column (5) of Table 2, are presented in Panel B. Overall, both research designs, and consistently across different specifications, suggest that parents select schools that are specifically effective in raising achievement of their children.

Heterogeneous effects and discussion

I next explore heterogeneity in match effects of attending the school of choice. Panel A of Table 4 reports estimates of the match effect from specifications analogue to column (2) in Panel B of Table 3. I report here the uninteracted effect of school enrolment, estimating achievement impacts of attending a preferred school conditional on value-added. Full-sample estimate is reported in column (1) for convenience, and sub-samples of students based on socioeconomic characteristics are considered in subsequent columns. Positive and statistically significant estimates are only found for male (column 3) and low-ability students (column 5), defined as those with baseline achievement below median. I combine these two sub-groups in column (7) and find that positive match effects are solely detected for boys entering primary education with a relatively low achievement level. Similar results are found in Panel B when comparing students who enrol in the same school, analogously to column 5 of Table 2.

Based on findings in the literature, male students with relatively low achievement at baseline may have specific learning needs. Females are often found to respond more productively to new environments with higher-quality peers. [Hastings et al. \(2006\)](#) and [\(Deming et al., 2014\)](#) show that positive effects of school choice on achievement and postsecondary education are concentrated among girls, who are found to demand higher-performing institutions. [Bailey and Dynarski \(2011\)](#) argue that females may have better interactions in classrooms and families.⁴⁵ On the other hand, recent findings in education have established that, conditional on absolute academic performance, student achievement increases with ability rank in the classroom, and this result is stronger among males (e.g., [Murphy and Weinhardt, 2020](#)).

A potential explanation behind match effects I uncover is that parents of low-ability males

⁴⁵Relatedly, females and males exhibit opposite impacts of moving to lower-poverty neighbourhoods, as the former better interact with the new environment ([Kling et al., 2007](#)).

select schools with lower absolute achievement and these are particularly suitable for them. First, I find that parents of boys with relatively lower achievement select schools with lower peer quality with respect to similar-ability girls, as can be seen in Table 5. Reported are estimates of coefficients on female indicator from regressions of peer quality of first-choice school, including an interaction term between gender and an indicator for achievement at entrance below median. Results on the uninteracted coefficient represent gender differences for low-ability students. Peer quality of first-choice schools follow the same patterns of heterogeneous effects described above, with low-ability girls applying to schools with 0.009σ higher absolute performance (column 1) with respect to boys with similar ability. Interestingly, the gap disappears when considering high-ability students, as shown by the interaction term. Moreover, column 2 shows that parents of low-ability males choose to give up more in terms of peer quality, selecting schools with relatively lower academic performance within their feasible set.⁴⁶

Second, different application choices by gender and ability are reflected in different peer quality gains from attending the school of choice. Appendix Table A.6 reports estimates from specifications similar to Panel A of Table 4 where the dependent variable is peer quality. Gender difference in peer quality gains is substantially more pronounced among low-ability students, with an estimated 0.3σ for males against 0.75σ for females (see columns 6 and 7).

Third, match effects are more pronounced for students enrolling in schools with relatively low peer quality. Table 6 reports estimates of heterogeneous effects by peer quality at the institution where the student enrolls. Presented are estimates of specifications controlling for school value-added similarly to Panel A of Table 4. Largest and statistically significant effect of attending the school of choice on student achievement is estimated when considering students in the bottom peer quality quartile (see column 2).

However, I note that this mechanism potentially explains only a part of the heterogeneous sorting effects. In particular, it is consistent with Panel A of Table 4, showing that achievement of low-ability males increases by attending the school of choice with respect to a less-preferred institution of the same quality. On the contrary, estimates reported in Panel B cannot be explained by observable differences between preferred schools. Indeed, these

⁴⁶As it could be expected, differences are larger for students facing higher variability in peer quality of feasible schools (columns 3-4).

capture heterogeneity in parental preference for the same school within students belonging to the same sub-group. Latter estimates suggest that a substantial part of the return to sorting I find is attributable to unobservable school characteristics.

Finally, I note that match effects uncovered here are potentially consistent with broader mechanisms than specific complementarities between students and the learning environment at the school selected by their parents. In particular, achievement of students missing out on the school of choice may suffer from negative motivational effects. Furthermore, parents and students themselves may adjust their behaviour depending on school assignment (Pop-Eleches and Urquiola, 2013). The overall direction of these effects, however, is ambiguous. While students may feel discouraged when denied a seat in the preferred school, parents may increase their own effort, for example by providing help with children’s homework as found by Pop-Eleches and Urquiola (2013). Interestingly, they also find students entering a school with higher-quality peers are more likely to feel marginalised, consistent with the mechanism proposed above. Unfortunately, I cannot directly test motivational or behavioural channels with the data at hand.

7 Summary and conclusion

One key argument supporting the recent and rapid expansion of school choice programmes is that competition among schools to attract enrolment would enhance school productivity (Hoxby, 2003). A growing literature suggests, however, that parents reward schools with high peer quality rather than those effectively improving achievement, implying scarce incentives for schools to improve (Barseghyan et al., 2019; MacLeod and Urquiola, 2019). This paper provides evidence on a relatively under-explored mechanism through which school choice can improve system-wide productivity. I use administrative records on the universe of applicants to London primary schools to study whether parents target institutions that are a better-than-average match for their children.

I begin by documenting that parental rankings are substantially heterogeneous even conditional on observed and unobserved characteristics and yet represent solid preferences consistent with parental choice over time. I then leverage centralised school assignment to identify the effect of attending the school of choice on student achievement over and beyond the

impact of average school quality.

I find that attending a preferred school increases academic achievement in mathematics with respect to an institution with the same average quality and ranked lower by parents. Match effects are concentrated on male students with relatively low achievement at entrance, a sub-group of students that may have particularly specific learning needs (e.g., [Murphy and Weinhardt, 2020](#)). My results represent novel evidence of returns to sorting into specific schools, with important implication for education policy makers. As findings suggest that, at least in large urban setting with high density of schools, parents input private information on the specific suitability of different institutions for their children, parental choice may increase productivity of the educational system and provide a reliable basis for school accountability. The specific sources of complementarities of students and schools in achievement production are a potentially promising direction for future research.

References

- Abadie, A. and Cattaneo, M. D. (2018). Econometric methods for program evaluation. *Annual Review of Economics*.
- Abdulkadiroglu, A., Angrist, J. D., Dynarski, S. M., Kane, T. J., and Pathak, P. A. (2011). Accountability and flexibility in public schools: Evidence from Boston’s charters and pilots. *The Quarterly Journal of Economics*, 126:699–748.
- Abdulkadiroglu, A., Angrist, J. D., Narita, Y., and Pathak, P. A. (2017). Research design meets market design: using centralized assignment for impact evaluation. *Econometrica*, 85(5):1373–1432.
- Abdulkadiroglu, A., Angrist, J. D., Narita, Y., and Pathak, P. A. (2019). Breaking ties: Regression discontinuity design meets market design. IZA DP No. 12205.
- 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.
- Abdulkadiroglu, A., Pathak, P. A., and Walters, C. R. (2018). Free to choose: Can school choice reduce student achievement? *American Economic Journal: Applied Economics*, 10(1):175–206.
- Ainsworth, R., Dehejia, R., Pop-Eleches, C., and Urquiola, M. (2020). Information and preferences in household demand for school value added. *NBER Working Papers n.28267*.
- Angrist, J. D., Hull, P. D., Pathak, P. A., and Walters, C. R. (2016). Interpreting tests of school value added. *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.

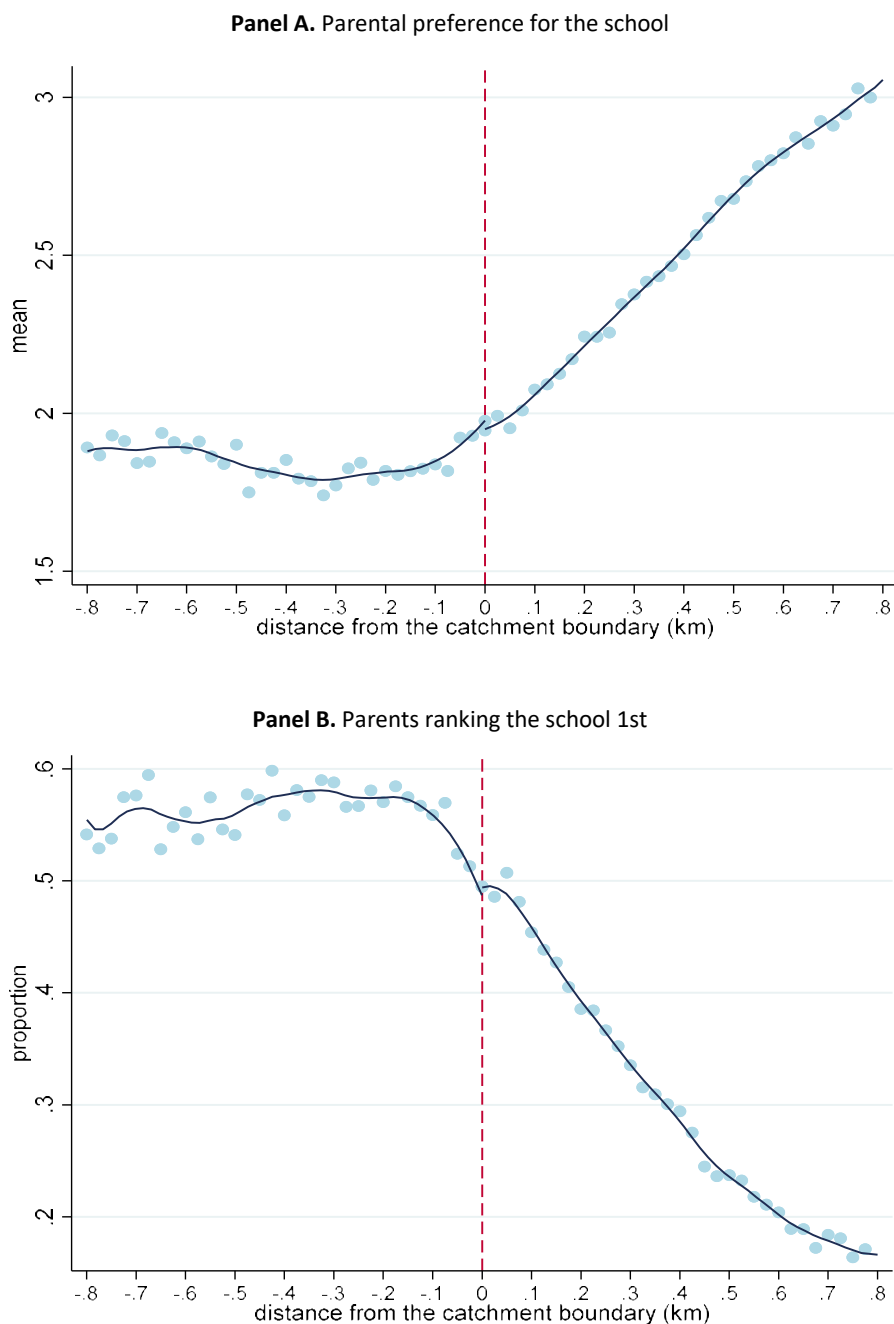
- Bailey, M. J. and Dynarski, S. M. (2011). Gains and gaps: Changing inequality in U.S. college entry and completion. *NBER Working Paper No. 17633*.
- Barseghyan, L., Clark, D., and Coate, S. (2019). Peer preferences, school competition, and the effects of public school choice. *American Economic Journal: Economic Policy*, 11(4):124–158.
- Battistin, E. and Neri, L. (2017). School performance, score inflation and economic geography. *IZA Discussion Paper No. 11161*.
- 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*. Forthcoming.
- Bonhomme, S., Lamadon, T., and Manresa, E. (2019). A distributional framework for matched employer employee data. *Econometrica*, 87(3):699–739.
- 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. (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.
- Carter, G., Pathak, P., and Terrier, C. (2020). Matching practices for secondary public schools—London, England. matching-in-practice.eu.
- 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. (2015). The performance of school assignment mechanisms in practice. Working paper.
- Deming, D. J., Hastings, J. S., Kane, T. J., and Staiger, D. O. (2014). School choice, school quality and postsecondary attainment. *American Economic Review*, 104(3):991–1013.
- Department For Education (2017). Key Stage 1: 2017 Assessment and Reporting Arrangements.
- Fack, G., Grenet, J., and He, Y. (2019). Beyond truth-telling: Preference estimation with centralized school choice and college admissions. *American Economic Review*, 109(4):1486–1529.
- Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):9–15.
- Gelman, A. and Imbens, G. (2019). Why high-order polynomials should not be used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, 37(3):447–456.
- 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. (2020). Heterogeneous effects of missing out on a place at a preferred secondary school in England. *IZA Discussion Paper No. 13167*.

- Hastings, J. S., Kane, T. J., and Staiger, D. O. (2006). Gender and performance: Evidence from school assignment by randomized lottery. *American Economic Review: papers and proceedings*, 96(2):232–236.
- 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 benefits to attending better schools. *Journal of Labour Economics*, 36(4):841–884.
- Hoxby, C. M. (2003). *The Economics of School Choice*, chapter Could school choice be a tide that lifts all boats?, pages 287–341. Chicago: Univ. Chicago Press.
- Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.
- Imberman, S. A. and Lovenheim, M. F. (2016). Does the market value value-added? Evidence from housing prices after a public release of school and teacher value-added. *Journal of Urban Economics*, 91:104–121.
- 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.
- Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.
- Lee, D. S. and Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48:281–355.

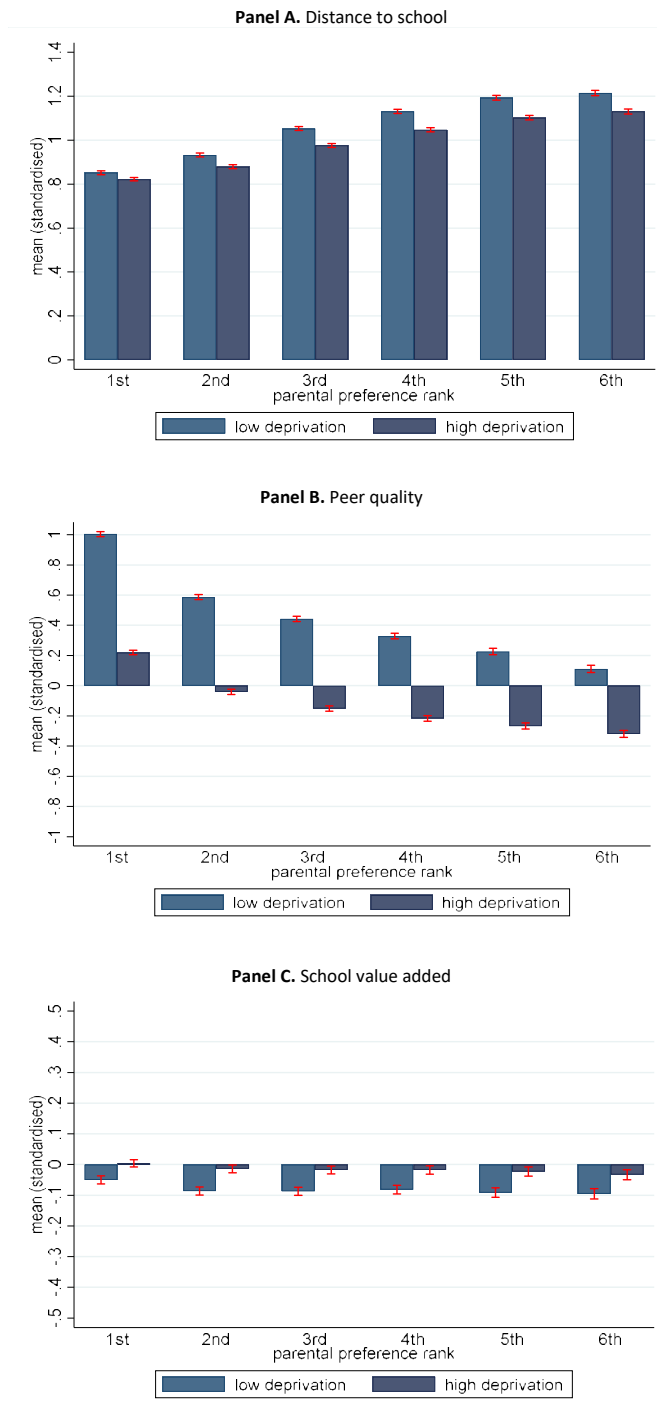
- Machin, S. (2011). Houses and schools: Valuation of school quality through the housing market. *Labour Economics*, 18(6):723–729.
- MacLeod, W. B. and Urquiola, M. (2019). Is education consumption or investment? Implications for school competition. *Annual Review of Economics*, 11:563–589.
- Masi, B. (2018). A ticket to ride: The unintended consequences of school transport subsidies. *Economics of Education Review*, 63:100–115.
- Mizala, A. and Urquiola, M. (2013). School markets: The impact of information approximating schools’ effectiveness. *Journal of Development Economics*, 103:313–335.
- Murphy, R. and Weinhardt, F. (2020). Top of the class: The importance of ordinal rank. *Review of Economic Studies*, 0:1–50.
- 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.
- 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).

Figure 1: Parental preference around the catchment boundary



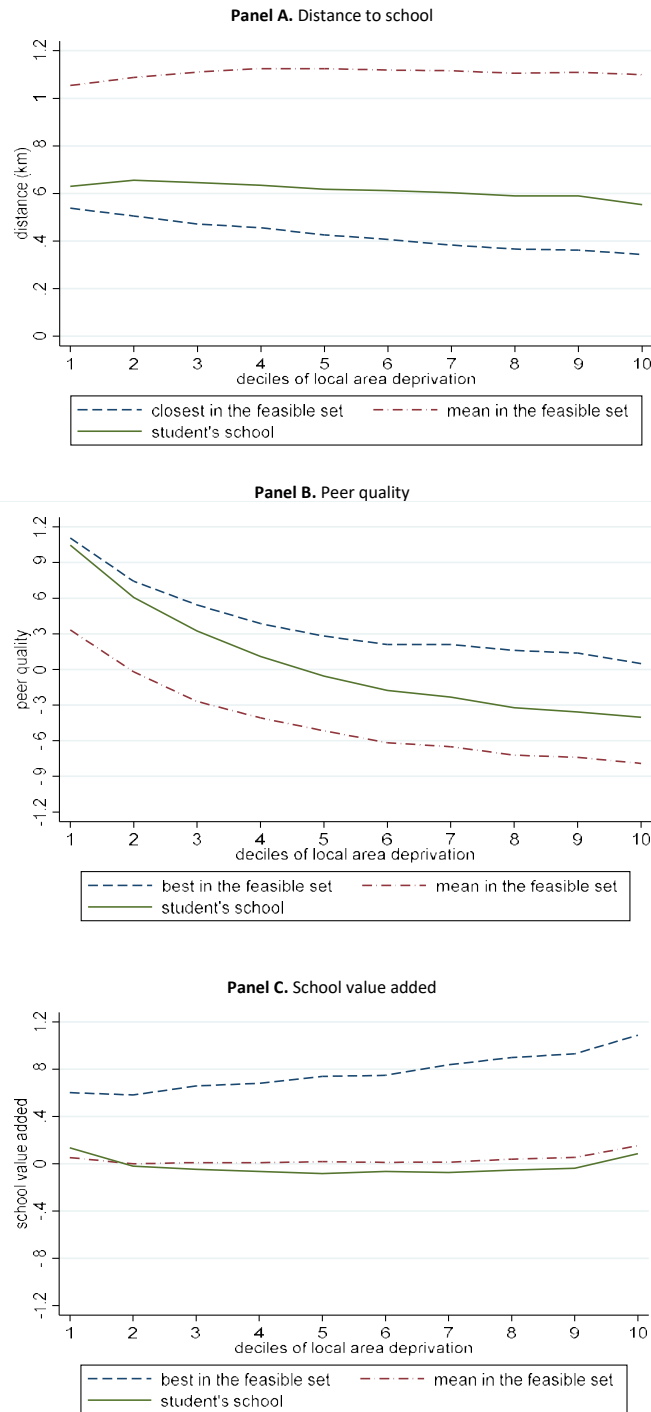
Note. The figure depicts parental rank (Panel A), and the share of parents ranking the school first (Panel B) around the catchment boundary. Preference for the school varies from 1 to 6 indicating first and sixth choice, respectively. Distance to school catchment boundary is represented on the horizontal axis and defined subtracting distance of the last admitted candidate to an applicant's distance to school. Negative values indicate residence within catchment. Markers represent average values in 25-meters-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 cut-off. Catchment boundary is defined for oversubscribed schools and not admitting by religion. The sample is restricted to applicants within 800 meters from the catchment boundary, and excludes last admitted applicants which are used to define school catchment. See Section 3 for details.

Figure 2: Parental rankings and school attributes



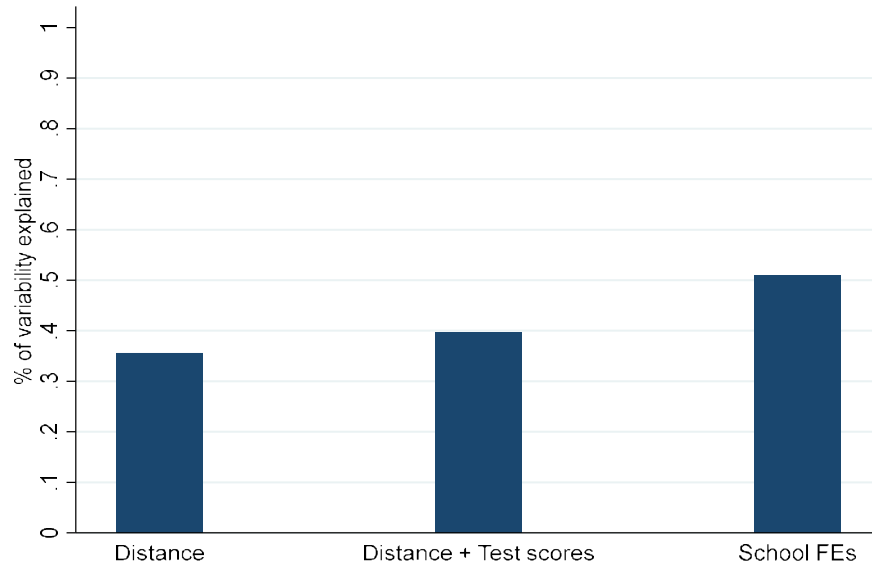
Note. The figure depicts average school attributes by parental rank estimated from equation (1.1). Bars plot predicted values from OLS regressions controlling for school feasibility and n. of schools listed, separately for students with local deprivation above or below the median. Controls also include dummies for quintile groups of school attributes other than the one considered, e.g. distance and peer quality when considering value added. Superimposed in red are 95% confidence intervals of predicted values. Panel A plots distance to school in kilometers computed as linear distance between student postcode and school postcode centroids. Panel B plots peer quality measured by school-level final year test scores averaged across subjects. Panel C plots school value added, estimated by regression-adjusted test scores growth at the school and averaged across subjects. Peer quality and value added are standardised among primary schools in London. Deprivation index is based on average income in the LSOA of residence. See Section 4 for details.

Figure 3: Attributes of feasible schools



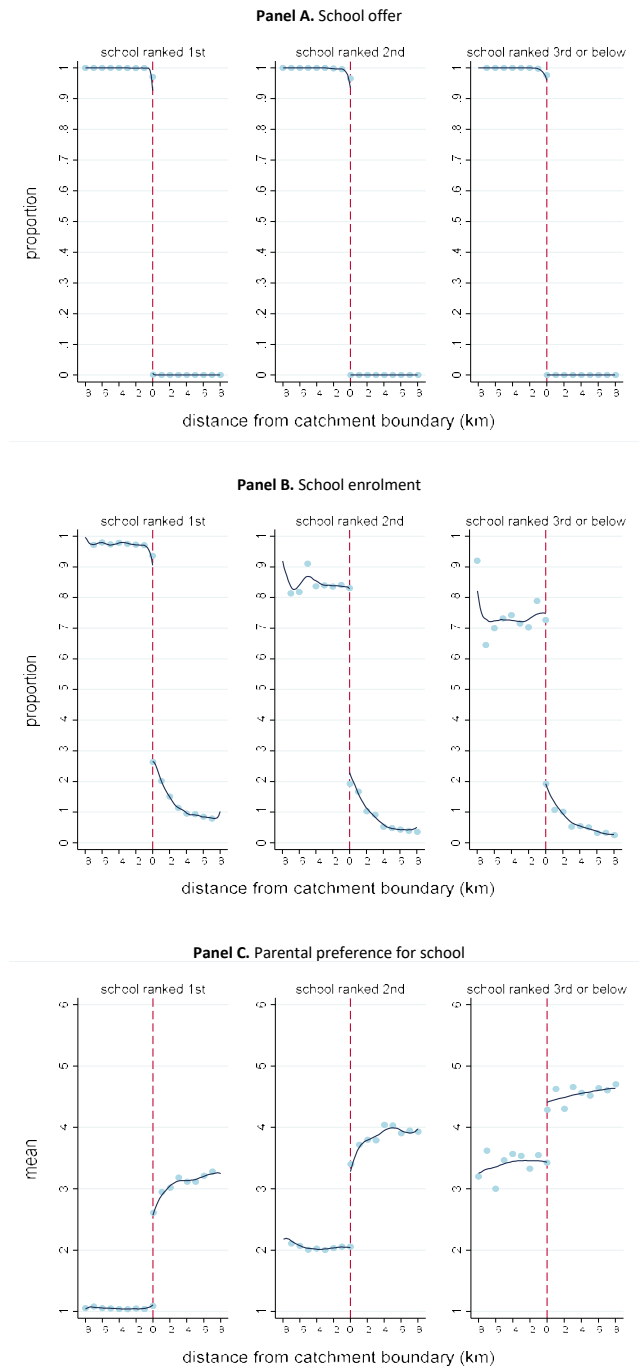
Note. The figure depicts average attributes of feasible schools by decile of deprivation index. Solid lines represent the school where an applicant enrolls, dashed lines represent the best feasible school based on the attribute considered, and dashed and dotted lines represent the average attribute among feasible schools. Panel A plots distance to school, in kilometers. Distance to school is computed as linear distance between student postcode and school postcode centroids. Panel B and C plot peer quality and school value added, respectively, standardised among primary schools in London. Peer quality is measured by school-level final year test scores, while value added is estimated by regression-adjusted test scores growth at the school and averaged across subjects. Deprivation index is based on average income in the LSOA of residence. See Section 4 for details.

Figure 4: Variability in parental rankings



Note. The figure depicts explained variability in parental preferences by school characteristics. Plotted is the adjusted R-squared index from OLS regressions of parental preference for the school on different set of controls. The left bar is from a regression controlling for distance decile indicators, the middle bar adds peer quality decile indicators, and the right bar adds school fixed effects. I consider all schools ranked by parents together with other institutions in individual choice set (these are ex-post feasible schools where parents did not apply). Parental rank of non-listed schools is coded to 7. Regressions further controls for ex-post feasibility and n. of preferences expressed. See Section 4 for details.

Figure 5: Research design



Note. The figure depicts school offer (Panel A), enrolment (Panel B), and parental rank (Panel C) around catchment boundary for schools ranked first, second and third or below at application. Enrolment is measured at the reception year. Preference for the school varies from 1 to 6 indicating first and sixth choice, respectively. Where an applicant is enrolled in none of listed schools, parental rank is coded to 7. Distance to school catchment boundary is represented on the horizontal axis and defined subtracting distance of the last admitted candidate to an applicant's distance to school. Negative values indicate residence within catchment. Markers represent average values in 100-meters-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 cut-off. 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 no admission priority and not eligible at any school ranked higher. See Section 5 for details.

Table 1: Covariate discontinuities

	Uncontrolled	Discontinuity at the boundary		
	(1)	(2)	(3)	(4)
Free school meal eligible	-0.0133*** (0.0014)	-0.0149* (0.0084)	-0.0139 (0.0085)	-0.0140 (0.0086)
Female	0.0037* (0.0021)	0.0051 (0.0137)	0.0106 (0.0139)	0.0121 (0.0142)
Special Education Needs	-0.0017*** (0.0004)	0.0018 (0.0019)	0.0016 (0.0019)	0.0016 (0.0021)
White	0.0193*** (0.0020)	0.0120 (0.0118)	0.0114 (0.0119)	0.0045 (0.0121)
Black	-0.0162*** (0.0015)	0.0177** (0.0076)	0.0146* (0.0077)	0.0132* (0.0078)
Asian	0.0116*** (0.0016)	0.0018 (0.0079)	0.0063 (0.0082)	0.0058 (0.0083)
English as additional language	0.0057*** (0.0020)	0.0093 (0.0125)	0.0181 (0.0125)	0.0238* (0.0126)
Deprivation in area of residence (LSOA)	-0.0160*** (0.0006)	-0.0029 (0.0027)	-0.0010 (0.0027)	-0.0001 (0.0023)
% of population with higher education (LSOA)	0.0094*** (0.0006)	-0.0016 (0.0015)	-0.0019 (0.0015)	-0.0016 (0.0012)
Achievement at Year 0	0.0287*** (0.0043)	0.0203 (0.0237)	0.0182 (0.0241)	0.0273 (0.0246)
N (Free school meal eligible)	361,880	42,127	41,702	41,593
Parental rank and n. of preferences FEs	Y	Y	Y	Y
Running variable LLP controls		Y	Y	Y
School of application FEs		Y	Y	Y
Next-best school FEs			Y	Y
Area of residence (MSOA) FEs				Y

Note. This table shows estimates of covariate balance around the catchment boundary. Column (1) reports OLS estimates of mean difference in baseline characteristics by school offer status, conditional on parental rank and n. of schools listed. Columns (2) to (4) restrict the sample to applicants with no admission priority and who cannot enter any institution listed with higher preference. Reported are estimates of offer balance from equation (1.7), where controls include a local linear polynomial of distance to the catchment boundary, estimated separately on each side of the cut-off. Observations are weighted by a triangular kernel with optimal data-driven bandwidth estimated following Calonico et al. (2014), separately for each outcome variable. Number of observations reported refer to regressions of free school meal eligibility. Specifications in columns (2) to (4) include school of application fixed effects, column (3) adds fixed effects for the next-best school listed by parents and column (4) add neighbourhood of residence (MSOA) fixed effects. All specifications control for individual characteristics other than the one considered as dependent variable. Standard errors are clustered at the individual level and reported in parenthesis. See Section 1.5 for details. ***p<0.01. ** p<0.05. * p<0.1

Table 2: Effects of attending the school of choice

	Average effects			Within-school		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. First Stage						
Enrolled in the school of choice	0.6299*** (0.0180)	0.6337*** (0.0182)	0.6268*** (0.0188)	0.1607*** (0.0167)	0.1722*** (0.0173)	0.1653*** (0.0168)
N	14,733	14,392	14,319	14,505	14,148	14,065
F-statistics	1227.81	1218.63	1110.14	93.00	98.76	96.79
Panel B. Reduced form estimates						
All subjects	0.0201 (0.0191)	0.0264 (0.0193)	0.0226 (0.0194)	0.0291 (0.0211)	0.0343 (0.0215)	0.0284 (0.0216)
Mathematics	0.0583*** (0.0216)	0.0650*** (0.0221)	0.0567** (0.0224)	0.0658*** (0.0242)	0.0763*** (0.0250)	0.0664*** (0.0253)
Reading	-0.0069 (0.0240)	0.0025 (0.0245)	-0.0004 (0.0250)	-0.0026 (0.0270)	0.0046 (0.0278)	-0.0002 (0.0282)
Writing	0.0101 (0.0250)	0.0136 (0.0256)	0.0174 (0.0260)	0.0210 (0.0280)	0.0214 (0.0290)	0.0224 (0.0293)
Panel C. 2SLS estimates						
All subjects	0.0298 (0.0284)	0.0390 (0.0285)	0.0336 (0.0289)	0.0756 (0.0548)	0.0877 (0.0551)	0.0729 (0.0554)
Mathematics	0.0858*** (0.0318)	0.0951*** (0.0323)	0.0835** (0.0330)	0.1641*** (0.0606)	0.1879*** (0.0617)	0.1635*** (0.0625)
Reading	-0.0103 (0.0360)	0.0038 (0.0366)	-0.0006 (0.0375)	-0.0070 (0.0722)	0.0122 (0.0733)	-0.0005 (0.0744)
Writing	0.0152 (0.0376)	0.0203 (0.0382)	0.0262 (0.0391)	0.0564 (0.0750)	0.0564 (0.0764)	0.0591 (0.0773)
N (All subjects)	122,046	120,963	120,780	122,046	120,963	120,780
Parental rank FEs	Y	Y	Y	Y	Y	Y
School of application FEs	Y	Y	Y	Y	Y	Y
Next-best school FEs		Y	Y		Y	Y
Area of residence (MSOA) FEs			Y			Y
School where enrolled FEs				Y	Y	Y

Note. This table shows estimates of the effect of attending the school of choice on student learning. Sample is restricted to applicants to the first choice and applicants to lower ranked schools conditional on missing out on all more-preferred institutions. Reported in Panel A are first stage coefficients on school offer estimated from equation (1.7). Reported in Panel B are reduced form estimates of school offer coefficient, while 2SLS estimates of the school enrolment coefficient from equation (1.8), instrumented using school offer status, are reported in Panel C. Dependent variable is an indicator for scoring above standards at Year 2 assessments by subject. Reported are also coefficients from a specifications stacking all subjects and controlling for subject fixed effects. Columns (4)-(6) add a set of fixed effects for the school where applicant enrolls at reception year. In all regressions, controls include a local linear polynomial of distance to the catchment boundary, estimated separately on each side of the cut-off. Observations are weighted by a triangular kernel with optimal data-driven bandwidth estimated following Calonico et al. (2014). Specifications in columns (1) to (3) include school of application fixed effects, column (2) adds fixed effects for the next-best school listed by parents and column (3) add neighbourhood of residence (MSOA) fixed effects. Next-best school is defined as the closest non-ranked feasible institution when students do not express preferences. Standard errors are clustered at the individual level and reported in parenthesis. See Section 6 for details. ***p<0.01. ** p<0.05. * p<0.1

Table 3: Effects of attending the school of choice by school value-added

	Mathematics			Reading			Writing			All subjects			School value-added		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Panel A. Reduced form estimates															
School Offer	0.0599*** (0.0223)	0.0600*** (0.0223)	0.0599*** (0.0223)	-0.0012 (0.0247)	-0.0011 (0.0247)	-0.0014 (0.0247)	0.0127 (0.0259)	0.0128 (0.0258)	0.0125 (0.0258)	0.0234 (0.0195)	0.0234 (0.0195)	0.0232 (0.0195)	0.0100** (0.0041)	0.0099** (0.0041)	0.0093** (0.0041)
School Offer X value-added quintile gain	0.0101 (0.0066)			0.0236*** (0.0071)			0.0202*** (0.0074)			0.0188*** (0.0058)			0.0592*** (0.0016)		
School Offer X value-added decile gain		0.0049 (0.0033)			0.0114*** (0.0035)			0.0104*** (0.0036)			0.0093*** (0.0028)			0.0301*** (0.0008)	
School Offer X value-added ventile gain			0.0022 (0.0016)			0.0055*** (0.0018)			0.0052*** (0.0018)			0.0045*** (0.0014)			0.0152*** (0.0004)
Panel B. 2SLS estimates															
School Enrolment	0.0877*** (0.0326)	0.0878*** (0.0326)	0.0876*** (0.0326)	-0.0015 (0.0368)	-0.0014 (0.0368)	-0.0020 (0.0368)	0.0191 (0.0385)	0.0192 (0.0385)	0.0186 (0.0385)	0.0346 (0.0287)	0.0347 (0.0287)	0.0342 (0.0287)	0.0153*** (0.0054)	0.0159*** (0.0052)	0.0143*** (0.0052)
School Enrolment X value-added quintile gain	0.0123 (0.0083)			0.0299*** (0.0090)			0.0255*** (0.0094)			0.0235*** (0.0072)			0.0817*** (0.0017)		
School Enrolment X value-added decile gain		0.0059 (0.0041)			0.0145*** (0.0045)			0.0131*** (0.0046)			0.0116*** (0.0036)			0.0415*** (0.0008)	
School Enrolment X value-added ventile gain			0.0027 (0.0021)			0.0069*** (0.0022)			0.0066*** (0.0023)			0.0056*** (0.0018)			0.0210*** (0.0004)
N	41,805	41,805	41,805	38,849	38,849	38,849	39,276	39,276	39,276	118,449	118,449	118,449	64,068	64,068	64,068
School of application FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Parental rank FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Next-best school FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note. This table shows estimates of the effect of attending the school of choice on student learning by school value-added gains. Reported are estimates from specifications similar to column (2) of Table 1, augmented with an interaction term between the school offer dummy and the difference in value-added between the school of choice and the next-best school in student preferences. Panel A reports reduced form estimates of school offer and interaction coefficients, while Panel B reports 2SLS estimates of the school enrolment and interaction coefficients. Columns (1), (4), (7), (10) and (13) consider the difference in school value-added quintiles. Columns (2), (5), (8), (11) and (14), and columns (3), (6), (9), (12) and (15) consider value-added deciles and ventiles, respectively. Next-best school is defined as the closest non-ranked feasible institution when students do not express preferences. Standard errors are clustered at the individual level and reported in parenthesis. See Section 6 for details. ***p<0.01. ** p<0.05. * p<0.1

Table 4: Heterogeneous match effects by student characteristics

	All (1)	Gender		Score at Entrance		Score at Entrance below median	
		Female (2)	Male (3)	Above median (4)	Below median (5)	Female (6)	Male (7)
Panel A. Same value-added							
Mathematics	0.0877*** (0.0326)	0.0333 (0.0465)	0.1057** (0.0481)	0.0590 (0.0559)	0.0950*** (0.0354)	0.0228 (0.0537)	0.1218** (0.0530)
N	41,805	20,229	21,412	19,831	21,757	9,257	12,137
Panel B. Same school							
Mathematics	0.1879*** (0.0617)	0.0543 (0.1029)	0.2612** (0.1027)	0.0881 (0.1152)	0.2521*** (0.0775)	0.0328 (0.1816)	0.4043*** (0.1384)
N	42,574	20,448	21,672	20,053	22,039	9,199	12,128
Parental rank FEs	Y	Y	Y	Y	Y	Y	Y
School of application FEs	Y	Y	Y	Y	Y	Y	Y
Next-best school FEs	Y	Y	Y	Y	Y	Y	Y

Note. This table shows heterogeneous match effects of attending the school of choice on student achievement. Panel A and Panel B report estimates from similar specifications to Table 3 and Table 4, respectively. Column (1) replicates estimate from column (2) in Panel C of Table 3 (Panel A) and from column (5) in Panel C of Table 2 (Panel B). Columns (2) to (7) report estimates of specifications similar to column (1) where the sample is restricted to students with characteristics indicated in column headers. Reported are estimates on Year 2 student achievement in mathematics, considering a dummy equal to one if the student scores above the expected standards. Score at entrance is measured averaging EYFSP teacher assessments at the reception year across the 17 learning goals considered. See Section 6 for details. ***p<0.01. ** p<0.05. * p<0.1

Table 5: Peer quality of first choice by gender and baseline achievement

	All applicants		Applicants with high variability in feasible peer quality	
	Peer quality	Peer quality left on the table	Peer quality	Peer quality left on the table
	(1)	(2)	(3)	(4)
Female	0.0089* (0.0045)	-0.0106*** (0.0037)	0.0183** (0.0074)	-0.0230*** (0.0060)
Female X Score at entrance above median	-0.0115 (0.0104)	0.0130* (0.0078)	-0.0195 (0.0160)	0.0268** (0.0122)
N	124,853	123,595	59,374	59,732
N. of preferences FEs	Y	Y	Y	Y
Student characteristics	Y	Y	Y	Y
Feasible choice set controls	Y	Y	Y	Y

Note. This table shows peer quality of first-choice school by gender and baseline achievement. Reported are OLS coefficients from regressions of peer quality on gender including an interaction term between gender and a dummy variable indicating score at entrance above median. Sample is restricted to applications to first-choice schools. Dependent variable is peer quality of first-choice school in columns (1) and (3) and the difference between peer quality of first-choice school and the highest-performing institution in the feasible set in columns (2) and (4). Columns (1) and (2) consider all applicants, columns (3) and (4) students with standard deviation of peer quality in the feasible set above median. All regressions control for fixed effects of the number of schools listed, language, ethnicity, special education needs, free lunch eligibility, and average peer quality and school value-added in the feasible school set. Peer quality is measured by school-level final year test scores pooling 9 cohorts of data up to 2014 and averaged across subjects. Score at entrance is measured averaging EYFSP teacher assessments at the reception year in 17 different learning areas. See Section 6 for details. *** $p < 0.01$. ** $p < 0.05$. * $p < 0.1$.

Table 6: Heterogeneous match effects by peer quality

	Peer quality				
	All (1)	Bottom quartile (2)	2nd quartile (3)	3rd quartile (4)	Top quartile (5)
Mathematics	0.0877*** (0.0326)	0.1544** (0.0756)	0.0355 (0.0795)	0.0226 (0.0707)	0.1002 (0.0793)
N	41,805	7,271	7,646	8,873	8,694
Parental rank FEs	Y	Y	Y	Y	Y
School of application FEs	Y	Y	Y	Y	Y
Next-best school FEs	Y	Y	Y	Y	Y

Note. This table shows heterogeneous match effects of attending the school of choice on student achievement based on peer quality. Reported are estimates from specifications similar to Panel A of Table 4. Column 1 reports full-sample estimate, while sample is selected in subsequent columns based on quartile of peer quality at the institution where the student enrolls (see column headers). Peer quality is measured by school-level final year test scores pooling 9 cohorts of data up to 2014 and averaged across subjects. See Section 6 for details. ***p<0.01. ** p<0.05. * p<0.1

Appendix

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.⁴⁷

First, I replicate centralised assignment based solely on distance to school and parental preference. I rank applicants to a school in ascending order of distance and iteratively eliminate candidates who are eligible at schools ranked with higher preference. Without observing priorities, this is not sufficient to replicate school offer. As shown in Appendix Figure A.3, catchment boundary estimated solely based on distance fails to retrieve the discontinuity in school offer embedded in centralised assignment. 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 catchment boundary estimated solely based on distance receives school offer, she must have been admitted with priority. In the first step, catchment boundaries are overestimated as admission priority is ignored. The distance to school of last admitted applicant is an upper bound of the true threshold as some candidates are admitted with priority. Therefore, any school offer granted to applicants located beyond the initially estimated threshold reveals priority in admission. These applicants are flagged and replication of school assignment is re-attempted by admitting them first. The procedure is iterated until no applicant with offer is found beyond the estimated threshold.

In detail, the algorithm I set up 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. As it is unobserved, all students start in the same priority group.
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

⁴⁷I 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), as shown in Appendix Figure A.2.

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).
 - (c) Repeat (a) and (b) analogously for second to fifth choice. In particular, if an applicant is eligible at the r -th choice, drop the applicant from the queue at all schools with parental rank lower than r . Retained applicants are re-ranked.
3. Repeat step 2 until no more applicant is dropped from the admission list. Assignment converges in at most 15 iterations.
 4. Assign priority to applicants who are admitted to school according to administrative records but who are ranked beyond school capacity after steps 1-3.
 5. Repeat steps 1-4 until no more applicants with priority are detected. The algorithm converges in 131 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 admitted because of school priority. At 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.

Convergence is shown in Panel A of Appendix Figure A.4, plotting the fraction of applicants with priority identified in each iteration, and showing this monotonically decreases to zero. Panel B of Appendix Figure A.4, depicting errors in school assignment by iteration, shows my assignment almost perfectly corresponds to actual school offer when the procedure is concluded. Consistent with the idea that catchment boundary is overestimated when ignoring priority, Panel C of Appendix Figure A.4 shows that median distance threshold monotonically decreases as applicants with priority are detected.

In an effort to validate the priority measure produced by my algorithm, I compare it with a proxy for siblings at the school, constituting the main source of unobserved priority in my context. I consider all students enrolled at the school of choice at the time of application

and compute the number of students located in the same postcode of a given applicant.⁴⁸ This is an upper bound of the number of siblings at the school, but it arguably provides an interesting proxy given the granularity of postcodes in London (the 80% of applicants' school-postcode combinations is not matched by any currently enrolled student).⁴⁹ Among students located beyond the catchment boundary of the first choice, 80% of those with a schoolmate in the same postcode receive an offer, suggesting that my proxy effectively captures admission priority. Appendix Figure A.10 shows that the share of students estimated to have a sibling is in line with admission priority inferred through my algorithm. As expected, the former varies smoothly around the catchment boundary, providing additional evidence that the distance cut-off is exogenous. Finally, I use this proxy for siblings to test robustness of my main results to unobserved priorities. Estimates in Appendix Table A.7 show that, when excluding applicants within catchment boundary who likely have priority, results are substantially unchanged.

Construction of individual feasible school set

I define the individual feasible school set exploiting school catchment boundaries I obtained from replication of centralised school assignment (see Section 3 and the Appendix). I compute linear distance between student postcode and all schools around, 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 about 200,000 applicants and 1,750 schools in my sample.

I define a school as ex-post feasible if the student is located within catchment 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. Admission is often loosely related to distance as religious schools are allowed to prioritise applicants based on faith. Non-religious undersubscribed schools are included in the individual choice set if they are located within 2 km from student postcode, corresponding to the 90th percentile of distance to school. The individual feasible school set is defined as the collection of ex-post feasible

⁴⁸I keep school-postcode cells with at most two students (the 90% of observations), as higher counts are more likely to reflect densely populated postcodes rather than potential siblings at the school.

⁴⁹Residential mobility across postcodes is a potential source of error in this proxy. However, it concerns a small fraction of students, as shown in Appendix Figure A.7, which I discuss further below.

schools.

Parental choice and school mobility

Parents move their children to a different school after reception year based on peer quality rather than school value-added, consistently with application behaviour described in Section 4 above. Panel A of Appendix Table A.4 presents estimates of linear regressions of school mobility on school attributes for the sample of students at risk of admission. One σ higher difference in peer quality between offered and desired school is associated with 6-7 percentage point higher likelihood of moving to another school after reception year. This difference persists, substantially unchanged, when controlling for school choice covariates as well as individual socioeconomic characteristics (see columns 2 and 3). On the contrary, the estimated coefficient on school value-added is much lower, about 1 percentage point. Residential mobility, likely involving larger costs, is almost unrelated to relative attributes of school offered and the school of choice (see Panel B of Appendix Table A.4).

School mobility response to centralised assignment implies that 2SLS estimates of β_1 in equation (8) capture a combination of initial enrolment and school mobility induced by school offer. Students located just beyond the catchment boundary are about 10 percentage points more likely to move to another school after reception year, as shown in Panel A of Appendix Figure A.9. As achievement is measured in Year 2, the relationship between school offer, initial enrolment (denoted by D_0), and enrolment at Year 2 (denoted by D_1), is represented by the directed acyclic graph (DAG, Abadie and Cattaneo, 2018) in Appendix Figure A.8. School offer, as good as randomly assigned around the catchment boundary, affects achievement only through initial enrolment. The latter, however, leads to the outcome of interest combining two different channels: the direct impact of school where the student initially enrolls and the indirect impact of increased school mobility based on initial enrolment.

Attrition

A potential concern is that non-random attrition based on school offer may hinder comparability of students around the catchment boundary. Consistently with parental response to school offer described in Section 4 above, parents are more likely to opt for private institutions

when missing out on preferred state schools, preventing the observation of achievement outcome (see Panel B of Appendix Figure A.9). Estimates in columns (1) and (2) of Appendix Table A.8 quantify the difference in follow-up rates, which is 12% on average and increasing in parental preference for the school.⁵⁰

For several reasons, it is unlikely that my results are mechanically generated by selective attrition. First, follow-up rate is high (81%) even among students not assigned to their preferred schools, as shown in column (1) of Appendix Table A.8. Potential outcomes of non-followed students would then need to be extremely different to drive my results. Second, baseline achievement of students around the catchment boundary of ranked schools is balanced conditional on observation of the outcome (see Table 1), suggesting that comparability of applicants holds among the subsample of students followed through to Year 2.⁵¹ Third, if differential attrition was driving my results, one would have expected positive and significant effects of entering the school of choice. In contrast, estimates in Table 2 show modest and not statistically significant average effects across subjects.

Nevertheless, I impute achievement outcomes to non-followed students to bound my estimates and show the results are strongly robust to potentially selective attrition. I exploit the fact that my outcome is a dummy (indicating achievement above expected standards at Year 2), and I estimate my empirical model by sequentially assigning to non-followed students beyond the catchment boundary a value of 1 (columns 3-5 of Appendix Table A.8) and a value of zero (columns 6-8). These estimates should represent, respectively, the lower and upper bound for my achievement effect estimates.⁵² Estimation are obtained from specifications analogous to those in Table 2 and 3. Results are very similar from main findings, suggesting that these are not influenced by differential attrition.

Admission propensity score in London

As a sensitivity check, I compute the admission propensity score proposed in [Abdulkadiroglu et al. \(2019\)](#) by adapting the formula they derived in Theorem 1 to school assignment in

⁵⁰I follow the presentation of columns (1) and (2) of Table 7 in [Abdulkadiroglu et al. \(2018\)](#).

⁵¹Baseline achievement is available only for students completing Year 2 assessments as the two variables are in the same data file.

⁵²In order to replicate the estimation of match effects, I also need to impute a school. I assign not-admitted non-followed students the state school from which they received an offer.

London. The propensity score summarises the probability of receiving an offer from the school of choice based on parental preferences, school priorities, and distance to school, which is used as tie-breaker at oversubscribed schools.

Following [Abdulkadiroglu et al. \(2019\)](#), I begin by partitioning applicants in three groups depending on their admission priority. Applicants with admission risk are those with the same priority of the last admitted applicants (“marginal priority” group). Students with higher or lower priority are “always seated” or “never seated” at the school, respectively.⁵³

For students with marginal priority, the building blocks of admission risk are the distance cut-off at the school and the cut-off at each institution which the applicant ranked with higher preference. Students never seated at each preferred school who are located close to the catchment boundary have, at the limit, a 50% probability of receiving an offer. Propensity score is then computed as 50% for students located within a bandwidth around the cut-off, and then adjusted depending on the chances of admission at each school ranked with higher preference.⁵⁴

The admission propensity score closely corresponds to school offer and balances the characteristics of offered and non-offered students, as expected. A regression of school offer on the propensity score yields a coefficient of 0.97 and a R-squared index of about 80%. Students at risk of admission, i.e. with score values strictly between zero and one, are 37,490 in my sample. When restricting the sample to this sub-group and controlling for propensity score dummies, admitted and not-admitted students have very similar characteristics, as shown in columns 2 and 3 of Appendix Table [A.9](#).⁵⁵

When controlling for the admission propensity score, match effects of attending the school of choice are in line with the main results. Averaging across subjects, estimated achievement gains with respect to a school with similar value-added and ranked lower by parents is 0.03σ and statistically significant, driven by larger effects on mathematics (see columns 2 and

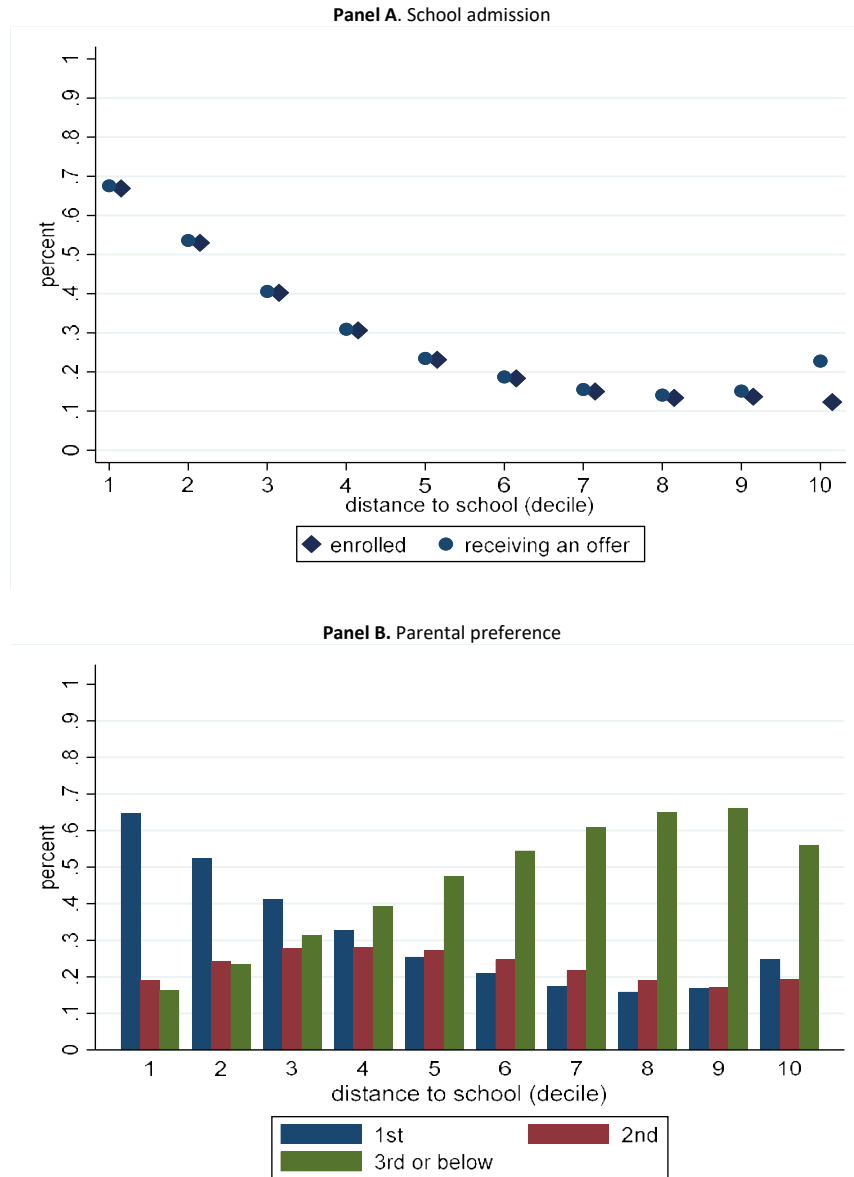
⁵³I use here admission priority as detected by my algorithm replicating centralised assignment (details are discussed above). As the error in observed priority is likely more serious at faith schools, applicants to religious institutions are never assigned to the marginal priority group. Instead, they are considered always or never seated depending on their offer. This implies no exogenous variation in assignment to faith schools, in line with my main empirical analysis.

⁵⁴I use the optimal data-driven bandwidth proposed by [Calonico et al. \(2014\)](#), which in my sample is of 187 meters.

⁵⁵Following [Abdulkadiroglu et al. \(2019\)](#), these specifications further control for linear polynomials of distance to the catchment boundary, interacted with a dummy equal to one for students located within the optimal bandwidth around the cut-off.

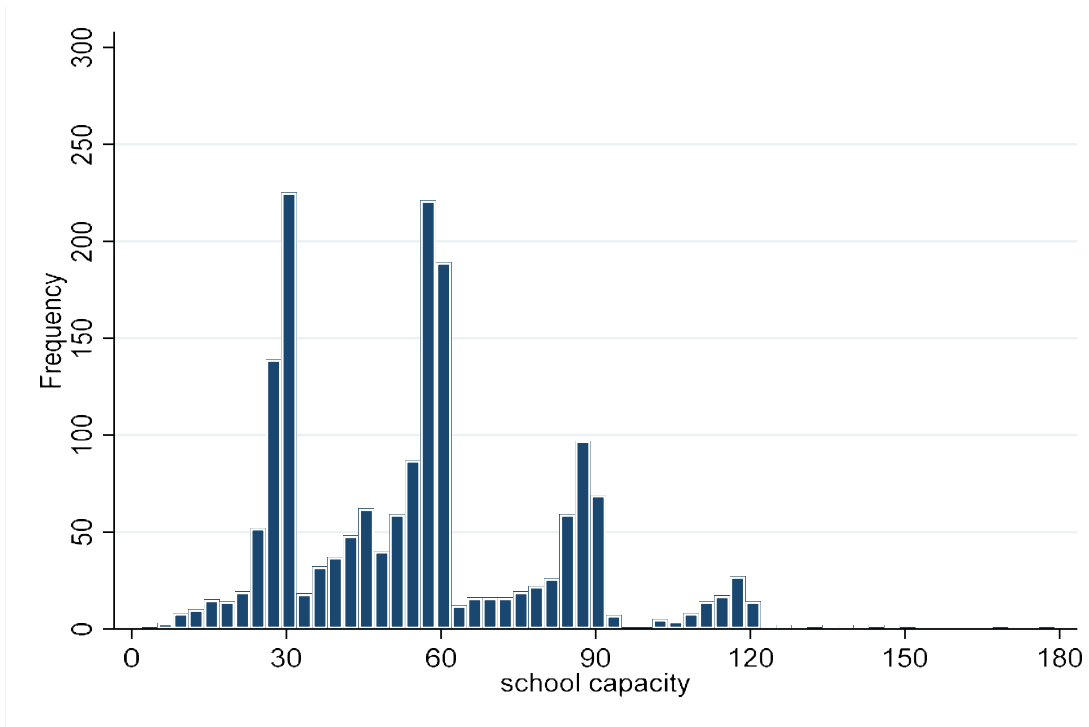
5, Panel A). Somewhat surprisingly, however, within-school estimates are smaller and not statistically different from zero.

Figure A.1: School admission and distance to school



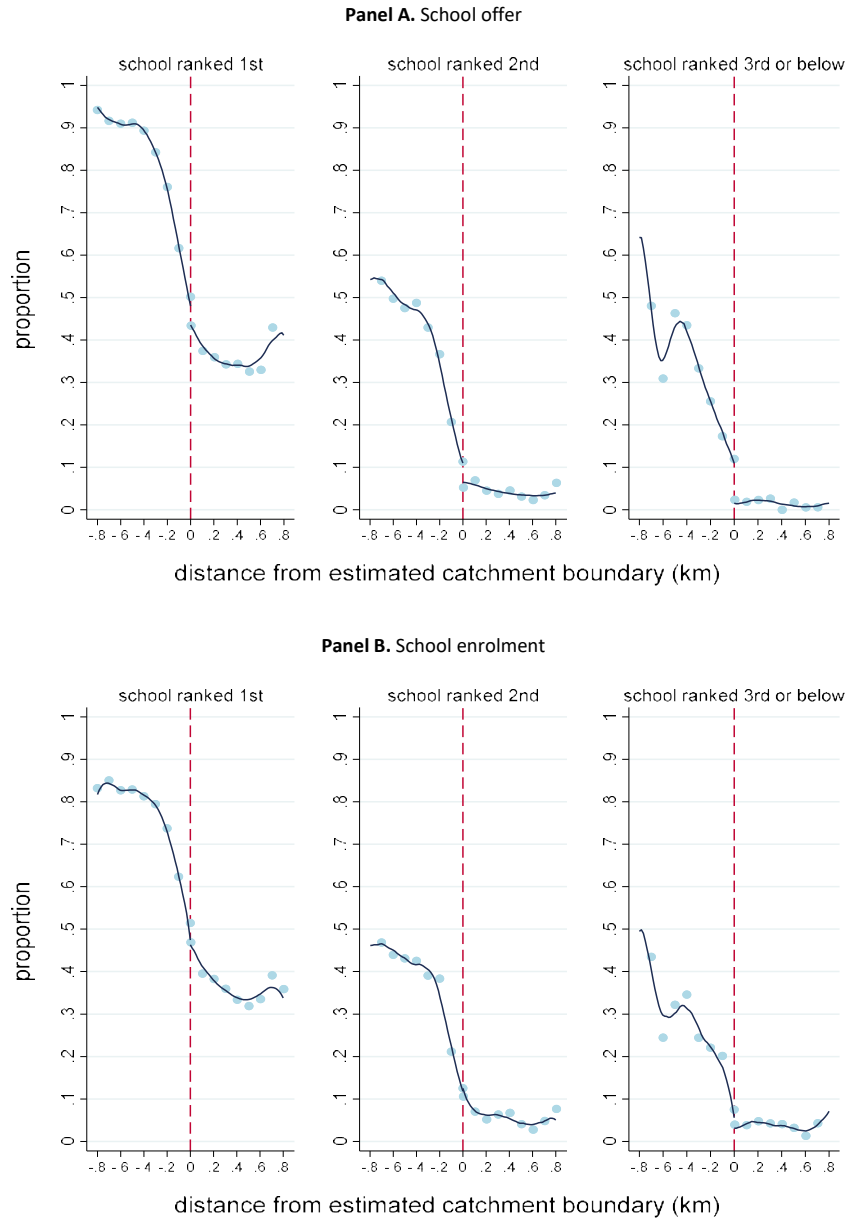
Note. The figure depicts school offer and enrolment rates and parental preference assigned to the school by distance. The sample includes all applicants to at least one London primary school in 2014 or 2015. Offer is reported by markers in Panel A, while diamonds represent enrolment measured at the reception 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.

Figure A.2: School capacity



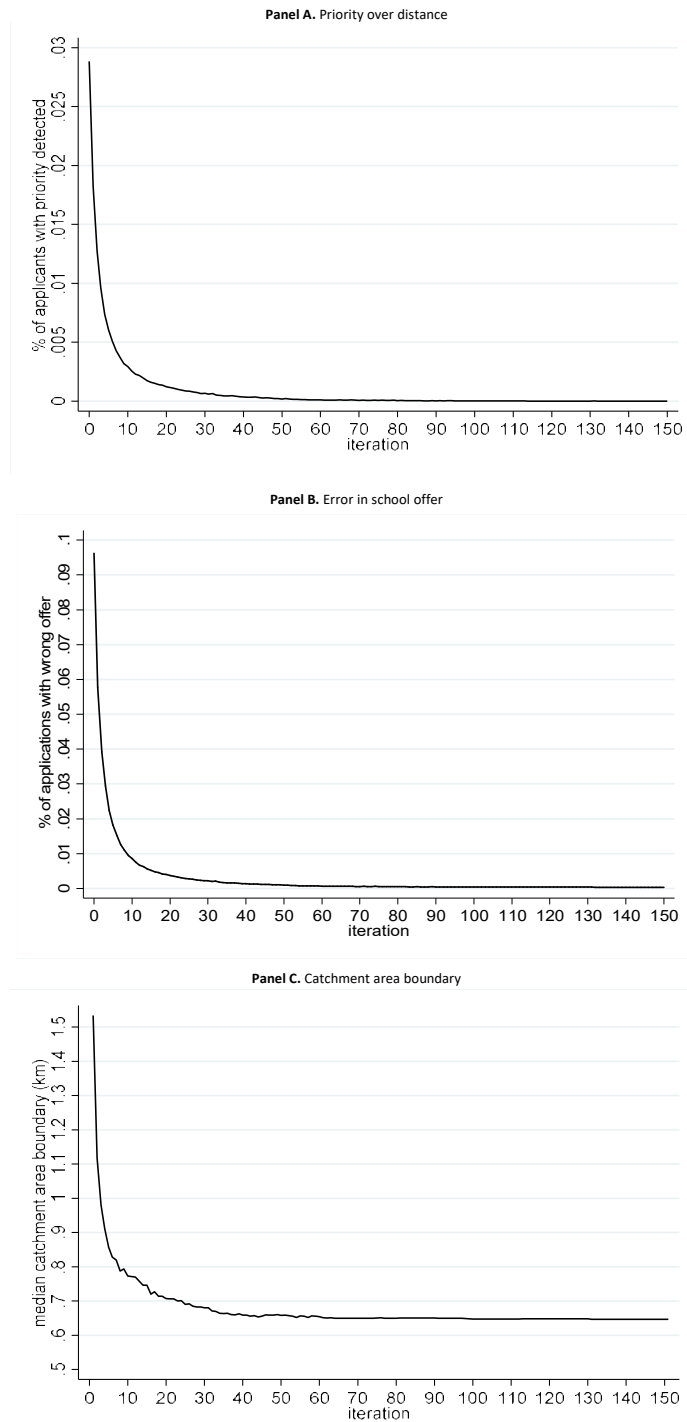
Note. The figure depicts the distribution of school capacity in London primary schools. Capacity is approximated by the number of offers issued. Bars represent frequency counts in three-units-wide bins, computed using one observation per school. See Section 3 for details.

Figure A.3: School assignment solely based on distance



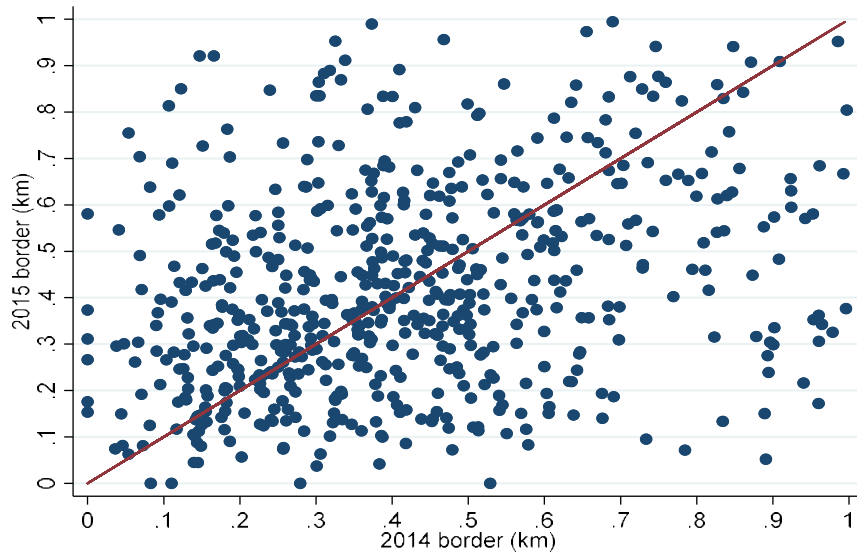
Note. The figure depicts school offer (Panel A) and enrolment (Panel B) around catchment boundary estimated by ranking applicants solely by distance to school. Sub-panel graphs group schools ranked first, second and third or below at application. School enrolment is measured at reception year. Distance to estimated catchment boundary is represented on the horizontal axis and defined subtracting distance of the last admitted candidate to an applicant's distance to school. Negative values indicate residence within estimated catchment. Markers represent average values in 100-meters-wide bins of the running variable and solid line is a local linear fit of underlying observations estimated separately on either side of the cut-off. The sample is restricted to applicants within 800 meters from catchment boundary and to applicants at risk of admission at the school, i.e. those who can not enter any institution listed with higher preference. See Section 5 for details.

Figure A.4: Replication of school assignment



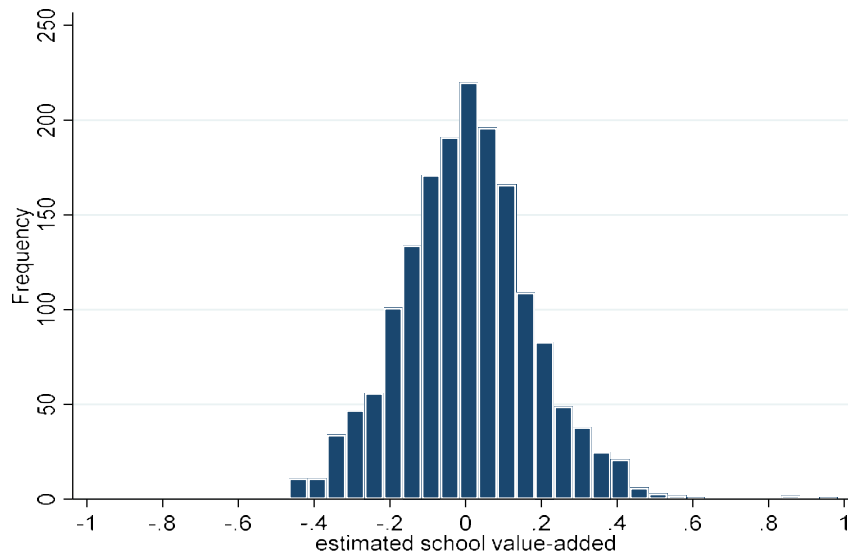
Note. The figure depicts the fraction of applicants with admission priority detected (Panel A), the fraction of applicants with wrong predicted offer (Panel B), and median catchment area boundary (Panel C) by iteration of the school assignment replication. The sample includes all applicants to at least one London primary school in 2014 or 2015. School assignment mechanism is replicated based on school capacity, parental preference and distance to school. Applicants are ranked solely by proximity in iteration 0 and those with offer beyond estimated boundary are flagged as enjoying priority. Subsequent iterations rank pupils by priority as retrieved in the previous round and, conditional on priority, by distance to school. Assignment converges in at most 150 iterations, after which no more applicants are found to enjoy priority. See Section 3 and the Appendix for details.

Figure A.5: Catchment boundaries over time



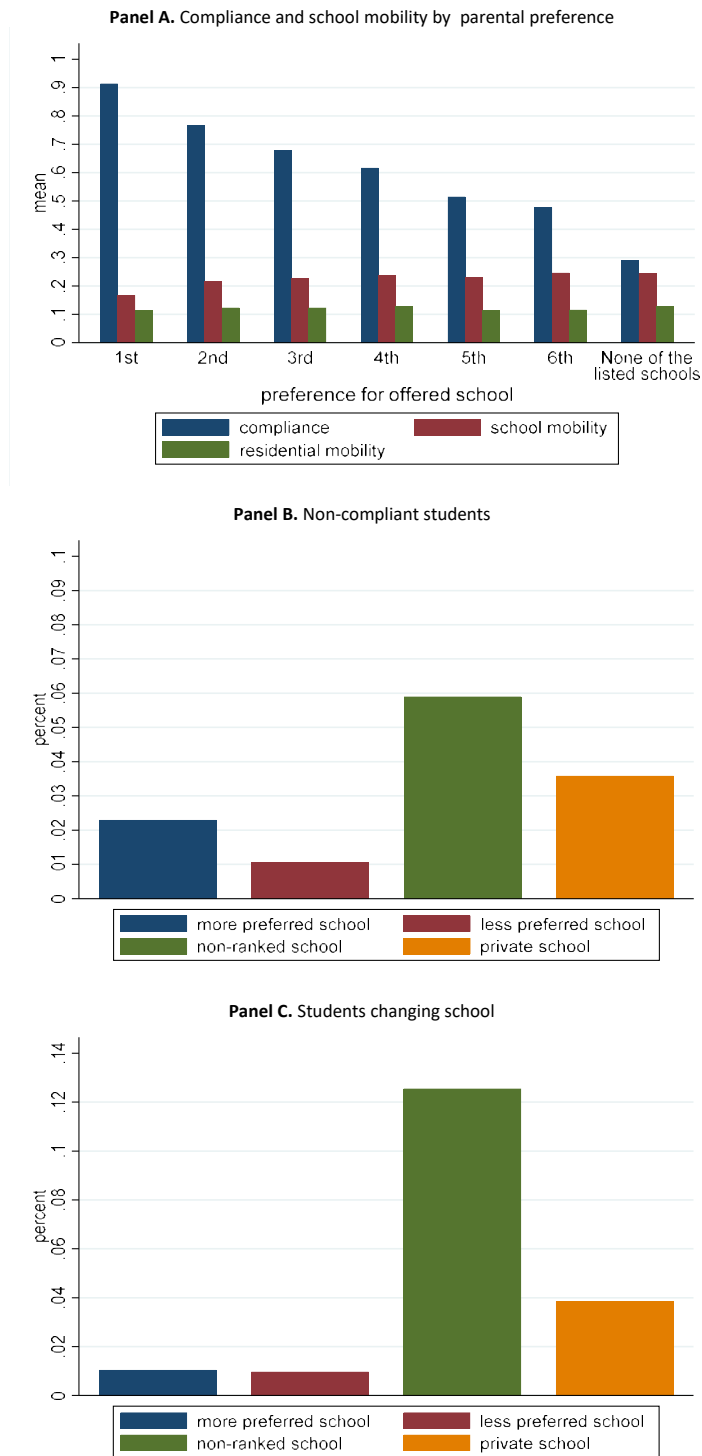
Note. The figure depicts estimated catchment area boundary of oversubscribed, non-faith state schools in 2014 and 2015. Boundaries are traced by replicating the centralised assignment mechanism for all London state primary schools. Markers in blue show one observation per school. The 45-degree line, indicating unchanged catchment boundary, is reported in red. Sample is restricted to schools with catchment boundary within 1km in both years (the 90% of observations). See Section 3 for details.

Figure A.6: Estimated school value-added



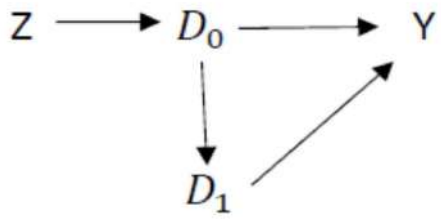
Note. This figure depicts the distribution of estimated school value-added. Reported are frequency counts in 0.05-wide bins using one observation per school. Value-added is estimated at baseline using students completing KS1 assessments in 2014-2016. Regressions consider KS1 outcomes analogous to those used in the main analysis and control for language, ethnicity, free school meal eligibility, special education needs, gender, local area deprivation and achievement at the reception year (EYFSP). 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 average value-added across subjects. See Section 1.4 for details.

Figure A.7: Compliance with assignment and mobility



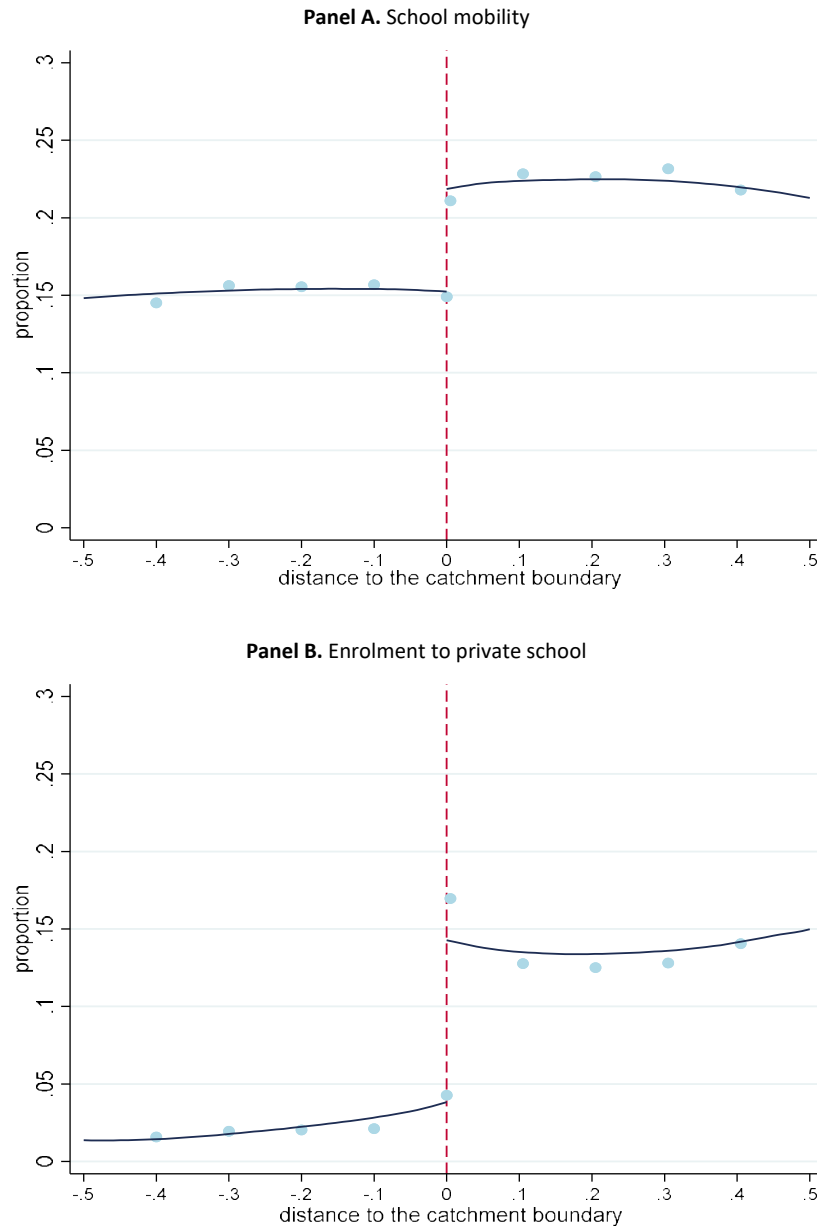
Note. The figure depicts compliance with school offer and school and residential mobility by parental preference. Panel A plots compliance, school mobility and residential mobility rates by parental rank for school offered. Panel B plots the share of students who do not comply with school offer by preference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

Figure A.8: Offer, enrolment and outcome in a DAG



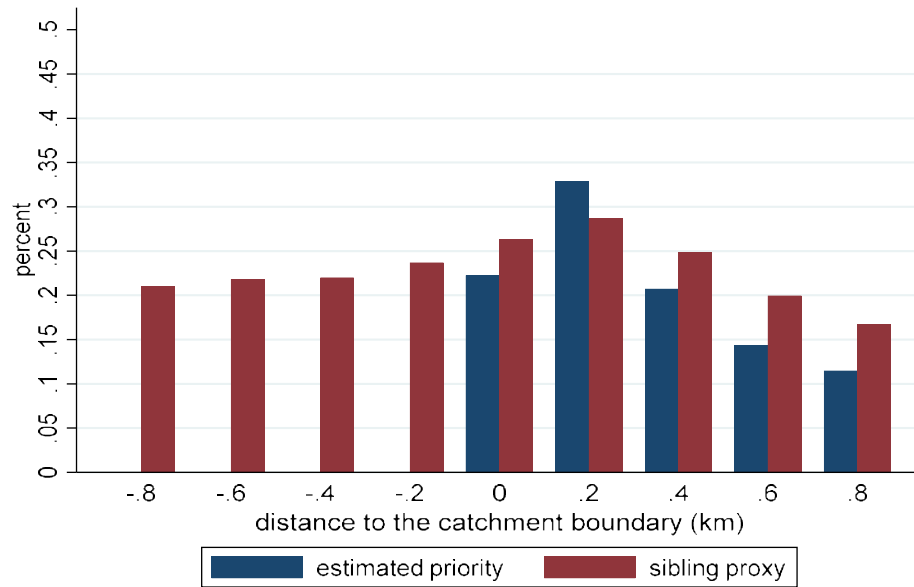
Note. The relationship between instrument, treatment and outcome in a directed acyclic graph. The graph includes initial school enrolment, D_0 ; enrolment at Year 2, D_1 ; and the achievement outcome, Y . See section 6 for details.

Figure A.9: School mobility and private school enrolment around the catchment boundary



Note. The figure depicts school mobility (Panel A) and enrolment to private school (Panel B) around catchment boundary for schools of choice, pooling institution ranked first to sixth at application. School mobility is an indicator variable equal to one if a student enrolls in a different school in Year 2 with respect to the reception year. Enrolment to private school is an indicator variable equal to one if an applicant is not observed in any state school in the reception year. Distance to school catchment boundary is represented on the horizontal axis and defined subtracting distance of the last admitted candidate to an applicant's distance to school. Negative values indicate residence within catchment. Markers represent average values in 100-meters-wide bins of distance from catchment and the solid line is a local linear fit of underlying observations, estimated separately on either side of the cut-off. The sample is restricted to applicants within 500 meters from catchment boundary and to applicants at risk of admission at the school, i.e. those with no priority and not eligible at any school ranked higher. See Section 6 for details.

Figure A.10: Estimated priority and proxy for siblings at the school of choice



Note. The figure depicts the share of applicants estimated to enjoy admission priority (blue bars) or to have a sibling at the school of choice (red bars) as a function of distance to the catchment boundary. Reported are averages in 200-meters-wide bins of distance from catchment boundary for students living within 800 meters from the cut-off. Admission priority is proxied by receiving a school offer while leaving outside the catchment area. Having a sibling at the school is proxied by the number of students at the school of choice living at the same postcode at the time of application. I exclude from the sibling proxy student-postcode combinations with more than 2 potential schoolmates (about 20%). See Appendix A for details.

Table A.1: Descriptive statistics (working sample)

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

Note. This table shows descriptive statistics about applicants to any mainstream state-funded primary school in England (column 1) or to at least one primary school in Greater London (column 2) in 2014 and 2015. Columns (1) and (2) report averages computed using one observation per student, column (3) reports the mean difference between (1) and (2). All statistics are conditional on non-missing observations. See Section 1.3 for details.

*** p<0.01, ** p<0.05, * p<0.1

Table A.2: Descriptive statistics (estimation samples)

	All London		RD sample		Same value-added decile	
	Mean (1)	SD (2)	Mean (3)	SD (4)	Mean (5)	SD (6)
<i>Baseline Characteristics</i>						
FSM eligible	0.1517	0.3587	0.1342	0.3408	0.1327	0.3393
Not speaking English at home	0.4185	0.4933	0.4140	0.4926	0.4020	0.4903
White	0.4165	0.4930	0.4242	0.4942	0.4331	0.4956
Asian	0.1946	0.3959	0.2134	0.4097	0.1912	0.3933
Black	0.1618	0.3683	0.1309	0.3373	0.1196	0.3246
Special education needs	0.0079	0.0887	0.0079	0.0883	0.0089	0.0942
Female	0.4900	0.4999	0.4873	0.4998	0.4860	0.4999
Exceeding expectations at Year 0: mathematics	0.1330	0.3386	0.1376	0.3445	0.1491	0.3563
<i>Achievement outcomes</i>						
Exceeding expectations at Year 2: mathematics	0.2667	0.4422	0.2839	0.4509	0.3084	0.4619
Exceeding expectations at Year 2: reading	0.3023	0.4593	0.3272	0.4692	0.3438	0.4750
Exceeding expectations at Year 2: writing	0.2032	0.4024	0.2163	0.4117	0.2313	0.4217
<i>School choice variables</i>						
N. of schools listed	3.2069	1.8731	3.5258	1.8566	3.7768	1.8133
Offered 1st choice	0.8283	0.3771	0.6749	0.4684	0.5989	0.4902
Enroled at offered school at reception year	0.8692	0.3372	0.8543	0.3528	0.8281	0.3774
Not enroled at state schools at reception year	0.0372	0.1892	0.0529	0.2239	0.0599	0.2374
N	199,220		45,422		5,740	

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)-(2), report averages and standard deviations for all applicants. Columns (3)-(4) consider applicants who are located around the catchment boundary based on the optimal data-driven bandwidth computed following Calonico et al. (2014). Among these students, columns (5)-(6) consider applicants whose school of choice and next-best preference have the same value-added decile. Samples in columns (3)-(6) exclude applicants who are admitted to more-preferred schools, as well as applicants to undersubscribed schools or schools admitting by religion. Statistics are computed using one observation per student. All statistics are conditional on non-missing observations. See Sections 3 and 5 for details.

Table A.3: Oversubscribed schools

	Popular schools	Not popular schools	Difference (1-2)
	(1)	(2)	(3)
<i>Peer quality</i>			
Sixth grade mathematics score	0.3651	-0.5877	0.9529***
Sixth grade reading score	0.3634	-0.6027	0.9661***
<i>School effectiveness</i>			
School value added in mathematics	0.0640	-0.1018	0.1657***
School value added in reading	0.0840	-0.1337	0.2177***
<i>School type</i>			
Religious school	0.2184	0.1292	0.0893***
Academy school	0.1406	0.1930	-0.0525***
Community school	0.5394	0.6067	-0.0673***
<i>Peer composition</i>			
% FSM eligible students	0.1843	0.2882	-0.1039***
% white students	0.4767	0.3654	0.1025***
Income deprivation in student loca area (LSOA)	0.3159	0.4012	-0.0853***
N	1053	689	1742

Note. This table shows characteristics of London primary schools by oversubscription status in 2014 and 2015. Column (1) and column (2) report means for oversubscribed and undersubscribed schools respectively, while mean difference is reported in column (3). A school is coded as oversubscribed 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, while value added is estimated by regression-adjusted test scores growth at the school and averaged across subject. Both measures are computed at baseline considering previous cohorts and are standardised to have zero mean and unit variance in the working sample. 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 4 for details. *** $p < 0.01$. ** $p < 0.05$. * $p < 0.1$

Table A.4: School mobility and school attributes

	(1)	(2)	(3)
Panel A. School mobility			
Peer quality difference	0.0739*** (0.00379)	0.0667*** (0.00400)	0.0694*** (0.00396)
Distance difference	-0.0468*** (0.00671)	-0.0207*** (0.00698)	-0.0249*** (0.00665)
School value added difference	0.00673* (0.00366)	0.00807** (0.00369)	0.0103*** (0.00355)
N	63,080	61,145	58,079
Panel B. Residential mobility			
Peer quality difference	0.00420 (0.00288)	0.00406 (0.00310)	0.00333 (0.00324)
Distance difference	-0.00196 (0.00504)	-0.00128 (0.00547)	-0.00145 (0.00554)
School value added difference	-0.00145 (0.00287)	-0.000527 (0.00289)	-0.00170 (0.00291)
N	61,540	59,693	57,906
School choice controls		Y	Y
Individual characteristics			Y

Note. This table shows correlation between school mobility and school attributes. Sample is restricted to applicants to the first choice and applicants to lower ranked schools conditional on missing out on all more-preferred institutions. Reported in Panel A are estimates from linear regressions of school mobility indicator, equal to one if a student moves to another school between reception year and Year 2. Dependent variable in Panel B is an indicator variable equal to 1 if a students moves residence (observed as home postcode). Independent variable are difference between characteristics of the school of choice and of the school offered. Peer quality and school value added are standardised to have zero mean and unit variance across London primary schools. Distance is measured in kilometers. Control variables include level of school characteristics. Column (2) adds n. of preferences expressed, preference for the school, preference for the school offered, ex-post feasibility of the school. Column (3) adds individual socioeconomic characteristics: gender, free lunch eligibility, special education needs, ethnicity, language, deprivation in area of residence and baseline achievement. Standard errors are clustered at the student level and reported in parentheses. See Section 6 for details.

*** p<0.01, ** p<0.05, * p<0.1

Table A.5: Alternative empirical specifications 1

	quadratic r.v. controls			cubic r.v. controls		
	Within 500m (1)	Within 800m (2)	Within 1Km (3)	Within 500m (4)	Within 800m (5)	Within 1Km (6)
Panel A. Same school value-added						
All subjects	0.0667* (0.0382)	0.0464 (0.0305)	0.0407 (0.0275)	0.0553 (0.0475)	0.0694* (0.0401)	0.0622* (0.0365)
Mathematics	0.1359*** (0.0453)	0.0785** (0.0362)	0.0879*** (0.0325)	0.1579*** (0.0565)	0.1513*** (0.0474)	0.1171*** (0.0431)
Reading	0.0334 (0.0466)	0.0120 (0.0371)	0.0079 (0.0333)	0.0006 (0.0587)	0.0360 (0.0489)	0.0254 (0.0445)
Writing	0.0308 (0.0480)	0.0486 (0.0382)	0.0263 (0.0343)	0.0074 (0.0604)	0.0210 (0.0503)	0.0441 (0.0458)
N (All subjects)	134,361	165,984	178,797	134,361	165,984	178,797
School of application FEs	Y	Y	Y	Y	Y	Y
Parental rank FEs	Y	Y	Y	Y	Y	Y
Next-best school FEs	Y	Y	Y	Y	Y	Y
Panel B. Same school						
All subjects	0.1125** (0.0570)	0.0881** (0.0419)	0.0796** (0.0368)	0.1114 (0.0689)	0.0965* (0.0532)	0.0990** (0.0474)
Mathematics	0.2061*** (0.0689)	0.1375*** (0.0503)	0.1516*** (0.0440)	0.2578*** (0.0835)	0.2125*** (0.0640)	0.1822*** (0.0567)
Reading	0.0618 (0.0710)	0.0345 (0.0515)	0.0244 (0.0452)	0.0231 (0.0863)	0.0463 (0.0657)	0.0390 (0.0581)
Writing	0.0697 (0.0732)	0.0924* (0.0534)	0.0629 (0.0465)	0.0532 (0.0895)	0.0308 (0.0682)	0.0757 (0.0605)
N (All subjects)	137,388	169,737	182,850	137,388	169,737	182,850
School of application FEs	Y	Y	Y	Y	Y	Y
Parental rank FEs	Y	Y	Y	Y	Y	Y
School where enrolled FEs	Y	Y	Y	Y	Y	Y
Next-best school FEs	Y	Y	Y	Y	Y	Y

Note. This table shows 2SLS estimates of the effect of attending the school of choice on student learning with parametric controls for distance to the catchment boundary. Panel A reports estimates of the uninteracted school enrolment coefficient from specifications similar to Table 3 and considering difference in school value-added decile. Panel B reports estimates of the school enrolment coefficient from specifications similar to column (5) of Table 2. Applicants are considered only if residing within a given distance from the catchment boundary, indicated in column headers. Independent variables include quadratic (columns 1-3) or cubic (columns 4-6) polynomial controls of distance to the catchment boundary. See Section 6 for details. ***p<0.01. ** p<0.05. * p<0.1

Table A.6: Heterogeneous match effects on peer quality (same school value-added)

	Gender		Score at Entrance		Score at Entrance below median		
	All (1)	Female (2)	Male (3)	Above median (4)	Below median (5)	Female (6)	Male (7)
Peer quality	0.5469*** (0.0454)	0.6458*** (0.0678)	0.4596*** (0.0581)	0.5546*** (0.0694)	0.4875*** (0.0611)	0.7627*** (0.1031)	0.3311*** (0.0832)
N	48,123	23,775	24,348	22,740	25,383	11,160	14,223
Parental rank FEs	Y	Y	Y	Y	Y	Y	Y
School of application FEs	Y	Y	Y	Y	Y	Y	Y
Next-best school FEs	Y	Y	Y	Y	Y	Y	Y

Note. This table shows heterogeneous effects of attending the school of choice on peer quality of school attended holding school value-added constant. Column (1) report estimated school enrolment coefficient from a specification similar to column (2) in Panel B of Table 3, where the dependent variable is peer quality at the school where students enrol at reception year. Columns (2) to (7) report estimates of specifications similar to column (1) where the sample is restricted to students with characteristics indicated in column headers. Peer quality is measured by school-level final year test scores pooling 9 cohorts of data up to 2014 and averaged across subjects. Score at entrance is measured averaging EYFSP teacher assessments at the reception year across the 17 learning goals considered. See Section 6 for details. ***p<0.01. ** p<0.05. * p<0.1

Table A.7: Robustness to unobserved admission priority

	Mathematics	Reading	Writing	All subjects
	(1)	(2)	(3)	(4)
Panel A. Same value-added				
School Offer	0.0868** (0.0377)	-0.0037 (0.0376)	0.0262 (0.0403)	0.0332 (0.0303)
School Offer X value-added decile gain	0.0032 (0.0048)	0.0144*** (0.0048)	0.0135*** (0.0050)	0.0115*** (0.0040)
N	29,806	29,725	29,965	86,850
Panel B. Same school				
School Offer	0.2015*** (0.0755)	0.0209 (0.0753)	0.0508 (0.0816)	0.0820 (0.0603)
N	30,365	30,287	30,526	88,824
School of application FEs	Y	Y	Y	Y
Parental rank FEs	Y	Y	Y	Y
Next-best school FEs	Y	Y	Y	Y

Note. This table shows 2SLS estimates of the effect of attending the school of choice on student learning when excluding students likely enjoying admission priority. Reported are estimates obtained when considering samples that exclude students located within the catchment boundary who are likely to have a sibling at the school of choice based on the postcode of residence of all students enrolled at the time of application. Panel A reports estimates of the uninteracted school enrolment coefficient from specifications as in Table 3 and considering difference in school value-added decile. Panel B reports estimates of the school enrolment coefficient from specifications as in column (5) of Table 2. See the Appendix for details. ***p<0.01. ** p<0.05. * p<0.1

Table A.8: Robustness to differential attrition

	Dep. Var.: followed-up student		Dep. Var.: Mathematics (assign "high" achievement to non-followed students)			Dep. Var.: Mathematics (assign "low" achievement to non-followed students)		
	Non-offered (1)	Offer differential (2)	All (3)	Same value-added (4)	Same school (5)	All (6)	Same value-added (7)	Same school (8)
First choice	0.7950*** (0.0093)	0.1451*** (0.0092)						
Second choice	0.8036*** (0.0271)	0.0865*** (0.0244)						
Third choice or lower	0.8287*** (0.0349)	0.0309 (0.0304)						
All choices	0.8121*** (0.0081)	0.1224*** (0.0078)	0.0449 (0.0308)	0.0862*** (0.0325)	0.1863*** (0.0625)	0.0468 (0.0308)	0.0918*** (0.0341)	0.1982*** (0.0651)
N (all choices)	48,587		108,144	40,484	41,203	115,827	40,955	41,714
Parental rank FEs	Y	Y	Y	Y	Y	Y	Y	Y
School of application FEs	Y	Y	Y	Y	Y	Y	Y	Y
Next-best school FEs			Y	Y	Y	Y	Y	Y

Note. This table shows follow-up rates by offer status and robustness checks to selective attrition. Columns (1) and (2) report are estimates from linear regressions of follow-up indicator on school offer dummy. Dependent variable is an indicator equal to 1 if Year 2 student achievement is observed. Sample is restricted to applicants to the first choice and applicants to lower ranked schools conditional on missing out on all more-preferred institutions. Column (1) reports coefficients on the intercept, representing average follow-up rate among non-offered students. Column (2) reports coefficients on the school offer variable. Results are presented by parental preference for the school and on average across all ranked schools ("all choices"). Columns (3), (4), and (5) replicate estimations in column (2) and column (5), Panel C of Table 2, and in column (2), Panel B of Table 3, respectively, when imputing a value of one for the achievement outcome indicator to all non-offered students who are not followed (i.e., non-followed students are assumed to be working above expected standards). Column (6)-(8) report estimates analogous to columns (3)-(5) when imputing a value of zero for the achievement outcome indicator to all non-offered students who are not followed (i.e., non-followed students are assumed to be working at or below expected standards). In all regressions, controls include a local linear polynomial of distance to the catchment boundary, estimated separately on each side of the cut-off. Observations are weighted by a triangular kernel with optimal data-driven bandwidth estimated following Calonico et al. (2014). Standard errors are clustered at the student level and reported in parentheses. See the Appendix for details.

Table A.9: Covariate balance with admission propensity score

	All applicants	Applicants with risk	
	(1)	(2)	(3)
Free school meal eligible	-0.0133*** (0.0014)	-0.0042 (0.0063)	-0.0005 (0.0063)
Female	0.0037* (0.0021)	0.0115 (0.0096)	0.0117 (0.0099)
Special Education Needs	-0.0017*** (0.0004)	0.0012 (0.0016)	0.0019 (0.0016)
White	0.0193*** (0.0020)	-0.0011 (0.0096)	0.0031 (0.0077)
Black	-0.0162*** (0.0015)	-0.0023 (0.0063)	0.0025 (0.0056)
Asian	0.0116*** (0.0016)	0.0077 (0.0077)	0.0048 (0.0061)
English as additional language	0.0057*** (0.0020)	0.0098 (0.0094)	0.0102 (0.0087)
Deprivation in area of residence (LSOA)	-0.0160*** (0.0006)	-0.0055* (0.0031)	-0.0062** (0.0029)
% of population with higher education (LSOA)	0.0094*** (0.0006)	0.0160*** (0.0027)	0.0135*** (0.0025)
Achievement at Year 0	0.0287*** (0.0043)	0.0214 (0.0198)	0.0122 (0.0188)
N (Free school meal eligible)	361,880	31,209	29,587
Parental rank and n. of preferences FEs	Y	Y	Y
P-score dummies		Y	Y
Individual characteristics			Y

Note. This table shows estimates of covariate balance between students with and without school offer. Column (1) reports OLS estimates of mean difference in baseline characteristics by school offer status from column (1) of Table 1. Columns (2)-(3) restrict the sample to applicants with admission risk as measured by the admission propensity score computed starting from Abdulkadiroglu et al (2019). Column (3) adds controls for individual characteristics other than the one considered as dependent variable. Standard errors are clustered at the individual level and reported in parenthesis. See Section the Appendix for details. ***p<0.01. ** p<0.05. * p<0.1

Table A.10: Match effects of attending the school of choice with admission propensity score

	First stage (1)	Mathematics (2)	Reading (3)	Writing (4)	All subjects (5)
Panel A. Average effects					
School Offer	0.7361*** (0.0056)	0.0473*** (0.0167)	0.0289* (0.0168)	0.0240 (0.0169)	0.0334** (0.0146)
N	33,783	30,772	30,772	30,772	92,316
Panel B. Same value-added					
School Offer	0.7415*** (0.0057)	0.0446*** (0.0171)	0.0282 (0.0172)	0.0262 (0.0403)	0.0316** (0.0149)
School Offer X value-added decile gain	0.0004 (0.0016)	0.0127*** (0.0041)	0.0202*** (0.0041)	0.0135*** (0.0050)	0.0164*** (0.0036)
N	29,806	29,565	29,565	29,965	88,695
Panel C. Same school					
School Offer	0.6690*** (0.0077)	0.0148 (0.0207)	-0.0081 (0.0208)	0.0508 (0.0816)	0.0034 (0.0177)
N	31,979	30,506	30,506	30,526	92,097
School of application FEs	Y	Y	Y	Y	Y
Parental rank FEs	Y	Y	Y	Y	Y
Next-best school FEs	Y	Y	Y	Y	Y

Note. This table shows estimates of the effect of attending the school of choice on student learning when controlling for the local admission propensity score derived in Abdulkadiroglu et al. (2019). Reported are estimates from regressions controlling for propensity score dummies and considering students with nondeterministic assignment risk. Controls are included for linear polynomials of the distance to the catchment boundary estimated separately on either side of the cut-off. Column (1) report first stage estimates of school offer coefficient in a regression of school enrolment. Columns (2)-(5) report 2SLS estimates of the school enrolment coefficient instrumented with school offer in regressions of Year 2 student achievement in the different subjects assessed. Panel A reports estimates of the uninteracted school enrolment coefficient from specifications including an interaction with the difference in value-added decile between the school of choice and the next-best preference. Panel B reports estimates of the school enrolment coefficient from regressions controlling for fixed effects of the school where student enrolls at the reception year. See the Appendix for details. ***p<0.01. ** p<0.05. * p<0.1

School of Economics and Finance



This working paper has been produced by the School of Economics and Finance at Queen Mary University of London

Copyright © 2021 Marco Ovidi

All rights reserved

School of Economics and Finance Queen Mary University of London

Mile End Road

London E1 4NS

Tel: +44 (0)20 7882 7356

Fax: +44 (0)20 8983 3580

Web: www.econ.qmul.ac.uk/research/workingpapers/