Re-examining the property price premium around 'good' schools

Ellen Greaves[1] and Alberto Venturin[2]

[1]University of Exeter, Institute for Fiscal Studies, IZA
[2]European University Institute

February 2025

Abstract

The established consensus from a large worldwide literature is that access to good schools increases local property prices, which is typically interpreted as reflecting parents' demand for school quality. We show that this relationship is not universal. Replicating the boundary discontinuity design, which compares property prices on either side of a school zone boundary, we show that the effect of access to the 'good' school varies across the distribution of the difference in school quality. Price premiums are concentrated in areas where the difference in quality between schools across the boundary is large, and particularly where paying the premium avoids a 'bad' school. There is no observable premium in other areas (around 25% of the sample). These findings are essential to change the characterisation of the effect of the public school system on the housing market from widespread to concentrated. This nuance has important consequences for our understanding of the effect of the public school system on households' welfare and the general equilibrium effects of school admissions reforms.

Acknowledgement

This early work has benefited from discussions with Steve Gibbons, Olmo Silva, Hans Sievertsen, the International Workshop on Applied Economics of Education, and colleagues at the European University Institute. This work contains publicly available data licensed under the Open Government Licence v3.0. The authors alone are responsible for the interpretation of the data.

1 Introduction

Parents' residential location affects either their child's default allocation to school or the probability of admission to their chosen school in many parts of the world. This incentivises parents to choose their residential location with school access in mind. Indeed, an established literature finds evidence of higher property prices linked to 'good' schools, typically defined by academic performance. For example, in their widely cited literature review, Black and Machin (2011) state, "A one standard deviation increase in school test scores seems to generate house price increases on the order of around 3–4%". Published estimates of this relationship are remarkably consistent across contexts, in settings with and without school choice systems, and high and low levels of school accountability. The relationship between local school quality and property prices has previously typically been estimated with a linear functional form. In practice, however, there could be non-linear affects concentrated around schools with particularly high or low levels of academic performance, or areas with large or small differences in performance between schools.

Our aim is to inform how widespread and uniform residential sorting for schools is, as indicated by higher property prices. Our main research question asks under what spatial conditions differences in property prices occur. We study whether the effect is linear, as typically assumed by previous literature, or instead is non-linear with respect to the difference in school quality between the two local schools. We also study whether price premiums are driven by households avoiding 'bad' schools or competing for access to schools at the top of the distribution. Finally, we study the multidimensionality of school quality measures, asking whether price premiums suggest households are making tradeoffs between school attainment and peer composition, for example. Our primary measure of school 'quality' is pupil test scores (which combines school effectiveness and pupil intake), but we also consider pupil progress (more aligned with value-added or school effectiveness), Ofsted inspection rating, and pupil composition.

Our hypothesis is that some areas are currently disproportionately affected by residential sorting for schools, and so would therefore be disproportionately affected by school admissions reform. Our results therefore inform the likely general equilibrium effects of policy reforms to decouple residence and school access (Burgess et al. (2020)) and recent implementation in some schools and areas (Brighton and Hove, 2024; Sutton Trust, 2024).

We interrogate estimates of the property price premium around good schools using the boundary discontinuity design, in the context of English high (secondary) schools. This established method, pioneered by Black (1999), compares the prices of properties on either side of a school zone (catchment) boundary. Properties on only one side have access to a higher-performing school, while other neighbourhood attributes do not jump discontinuously. The identifying assumption is that neighbourhoods on either side of the boundary would be the same if school assignment were the same. In our setting, there is a nationwide system of school choice, but school zones are used to give pupils priority if the school becomes over-subscribed. Therefore, our interpretation of the price premium around the school zone boundary is the effect of the discontinuous increase in the probability of access to the higher-performing school.

We augment the traditional boundary discontinuity design methodology by defining a treatment indicator equal to one if the property is on the 'high' school quality side of the boundary, rather than entering school quality directly. This allows us to interact the treatment indicator with measures of the difference and levels of school quality across the boundary to explore non-linearity. We also assess households' tradeoffs between alternative school quality measures by asking how price premiums for one measure of school quality are influenced by another, rather than considering one dimension of school quality in isolation.

We have three main findings. First, we find that property prices are higher only when the difference in school quality across the boundary is sufficiently large. School quality differences below the top quartile induce weaker residential sorting, as inferred from property prices. This implies that the welfare of households choosing their residential location in these areas is not strongly affected by the geographical school admissions arrangements of the public school system. The general equilibrium effects for households in these areas would be indirect, affected only by changing residential demand from other neighbourhoods. In contrast, the price premium is around 5% where the difference in school quality is the largest.

Second, we find that the absolute quality of the schools matters: price premiums are driven by both sorting towards the best-performing schools and by sorting away from the worst-performing schools. We estimate a premium of around 13% for houses more likely to access a better school (conditional on a large enough difference in quality) where the 'better' school is at or below the median of the nationwide quality distribution. In comparison, the premium for accessing top schools is instead around 3%. This changes the existing narrative that households compete more to access 'superstar' schools (Gibbons and Machin (2006); Bibler and Billings (2020); Chan et al. (2020)).

Third, we find that there is some tradeoff between school attainment and school socioeconomic status. Where there is a large difference in school quality across the boundary, prices rise regardless of the relative school composition. Where there are smaller differences in school quality, a minimum threshold of affluent peers is necessary for prices to rise for highly attaining schools. This finding highlights the multidimensionality of school 'quality'. Improvements in school quality alone might not be enough to induce some households to compete for access.

Our paper contributes to the large literature on the effect of access to higher school quality on property prices, estimated using the boundary discontinuity design. Our study unifies the exploration of heterogeneity to determine under what conditions the price premium arises, and is strongest, along multiple dimensions. Black (1999) found that the price premium attached to the higher quality elementary school zone was around 2% for a one standard deviation increase in school quality. This seminal study has led to numerous replications worldwide, in various settings. For a one standard deviation increase in school quality, the magnitude of the estimates for price increases ranges from 0.7% in elementary and secondary schools in Vancouver, British Columbia (Ries and Somerville (2010)) to 10% for elementary schools in North Carolina, US (Kane et al. (2006)). Although there are exceptions - Leech and Campos (2003) find a premium in housing prices 16 or 20% for two popular secondary schools in Coventry, UK - most estimates are around 3%.[1]

Universally, regardless of the setting, these papers find a statistically significant positive effect on property prices on the 'good' school side of the boundary. For example, even in Finland, where school performance is not publicly available, the premium attached to the 'good' school side is around 3%. Most of these papers estimate a linear effect of school quality on property prices, while only a minority have allowed for heterogeneity. Exceptions are Gibbons and Machin (2006) who explore the interactions between school quality, over-subscription and distance in the resulting price differential, and Fack and Grenet (2010), who show that the estimated price premium for French middle schools is diluted by the presence of private school outside options. Bibler and Billings (2020) and Chan et al. (2020) find that the price premiums are highest around the highest performing schools.

Our findings have important implications for the general equilibrium consequences of reforms to school choice and admissions criteria, which will depend on the precise relationship between school quality and prices (demand). This is important to consider in the context of calls to reform school admissions due to the concern that poorer households are priced out of 'good' schools and that school zones contribute to residential segregation of cities (Davidoff and Leigh (2008), Hamnett and Butler (2013), Bernelius and Vilkama (2019), Oberti and Savina (2019)).

[1]For primary/elementary schools, estimates range from 1.8% to 10%. In order of magnitude, for a one standard deviation increase in school quality: 1.8% increase in the San Francisco Bay Area, US (Bayer et al. (2007)); 2.4% increase in Shanghai, China (Chan et al. (2020)); 2 to 4% increase in Boston, US (La (2015)); 3% increase in Helsinki, Finland (Harjunen et al. (2018)). Gibbons and Machin (2003) estimate an increase of 3.5% in England, similar to Gibbons and Machin (2006) for London and the South East, and Gibbons et al. (2013) for England. Finally, Kane et al. (2006) find a 10% premium in North Carolina, US. Only one study considers middle schools. Fack and Grenet (2010) find a one standard deviation increase in public middle school performance raises housing prices by 1.4 to 2.4% in Paris. For high/secondary schools, estimates range from 3.5 to 7% for a one standard deviation increase in school quality: 3.5% increase in the Australian Capital Territory (Davidoff and Leigh (2008)); 6% increase in Auckland, New Zealand (Cheung et al. (2022)); 7% increase in the Charlotte Mecklenburg Schools District, US (Bibler and Billings (2020)). Studying both elementary and high schools, Ries and Somerville (2010) find a 0.7% price increase in Vancouver, British Columbia. Finally, on less comparable metrics, Han et al. (2021) find around a 3% increase in prices for 'locking' access to elite senior high schools in Beijing, China, to local residents. Leech and Campos (2003) find large premiums of 16% and 20% for two popular secondary schools in Coventry, England, where the difference in school 'quality' across the boundary is not quantified.

The rest of the paper is structured as follows. Section 2 describes the education system in England and the residential incentives created for parents through commonly used over-subscription criteria. Section 3 sketches a model (in progress) of residential and school choices which will provide a framework for the subsequent empirical analysis, and quantify the likely effects of school admissions reform. Section 4 details the data used, including construction and summary statistics. Section 5 outlines the methodology to identify the causal effect of local school quality on property prices (including, in our interpretation, the effect of endogenous neighbourhood characteristics that arise due to school quality). Section 6 presents the results for the role of geographic admissions on neighbourhood sorting, before Section 7 concludes.

2 Context

Our context is admission to public high schools (alternatively called state secondary schools) in England, where school choice has been enshrined nationwide since the Education Reform Act of 1988. Everywhere in the country, parents must submit a ranking of preferred schools to their Local Authority (LA) of residence in the year before their child starts high school. These rankings can be informed by school league tables of school performance, and a wide array of indicators for pupil composition, in addition to information gathered through informal local networks. If a school becomes over-subscribed, then school admissions criteria are used to rank pupils according to their characteristics. This section describes each of these features of the English state school admissions system in detail.

Information available to parents: Parents have access to copious information about school composition and performance. For example, the Department for Education maintains two websites that make it easy for parents to find their local school options and the characteristics of these schools.[2] Characteristics include pupil performance (the share of pupils achieving the benchmark standard of performance, and specifically in English and maths, and the 'English Baccalaureate'), pupil progress (a measure of value-added, or school effectiveness), pupil absence, and pupil composition. Pupil composition includes, for example, the percentage of pupils whose first language is not English, and eligible for free school meals at any point in the past six years.

Parents' school choices: Pupils begin high school in September of the academic year in which they turn 12 years old.[3] Parents must submit their ranked list of high school choices before the deadline of 31st October in the previous academic year. These school choices are submitted to their LA of residence. Parents can choose schools from other LAs by nominating them on their LAs list. The capital, London, has a single coordinated system.

Schools' over-subscription criteria: If a school has more places than applicants, all applicants are admitted. When a school becomes over-subscribed, the over-subscription criteria give pupils priority to determine which pupils are admitted. For example, a 'sibling' criteria would give applicants with a sibling at the school priority above applicants without. These criteria must be pre-defined and published in advance. Over-subscription criteria must abide by the Government's School Admissions Code (2021), which prohibits using specific criteria likely to favour advantaged groups in society, such as interviews with parents.

The most common over-subscription criteria (aside from the sibling criteria) is geographic. Burgess et al. (2023) show that around 90% of secondary schools in England have some form of geographic school admissions criteria. Around 53% have a pre-defined catchment area, which gives pupils who live inside priority above those who live outside. Almost all schools with geographic criteria use the

[2] Get Information about Schools and School Performance Tables.

[3] The academic year runs from September until mid-July. Pupils born in September are therefore the oldest in the cohort, while pupils born in August are the youngest.

distance between home and school as a tie-breaker between pupils with otherwise equivalent priority.

Other common admissions criteria include special needs (used by 48% of secondary schools), a child of staff (used by 44%), feeder primary schools (used by 38%) and religious observance (used by 15%).

Around 11% of schools use pupil ability or aptitude as an admissions criterion. Approximately half of these schools are 'selective' or 'grammar' schools, which select all pupils according to ability. Pupils must pass an entry test to gain admission to these prestigious schools.[4] The other half of schools are partially selective - permitted to admit up to 10% of their intake according to aptitude in the school's specialist subject, for example, music, languages and sport.

School assignment: Given parents' school choices and schools' admissions criteria, each LA then runs a truth-revealing assignment mechanism to allocate pupils to schools. Overall, the assignment mechanism allows parents to make multiple school choices with few strategic distortions to their choices. Pupils are assigned to their most preferred school at which they gain admission according to the schools' capacity and (if applicable) over-subscription criteria.

More formally, the assignment mechanism is known as 'equal preferences' and is equivalent to the Gale-Shapley deferred acceptance mechanism with a short list length (between 3 and 6 choices, depending on the LA). This short list length means some parents are incentivised to misreport their true preferences to include a safe rather than ambitious school choice, coined "skipping the impossible" by Fack et al. (2019). This is because, if they are unassigned to any of their ranked schools, they would be allocated to a school with spare places (by definition, unpopular).

3 Model (work in progress)

This model will allow us to generate comparative statics exploring the role of the functional form of families' preferences for school quality, and the effect of the level and difference in school quality across the boundary on equilibrium prices. It will also allow us to (roughly) quantify the welfare effects of policy reform removing catchment areas.

Notes: we ignore the distance tie-breaking rule. We are excluding non-parents from the economy, and a private school outside option. School quality is exogenous.

3.1 Primitives

Each family has one child, who will enrol in public secondary school. Each family is endowed with exogenous income (z_i) and work location. Given this exogenous work location, each family must choose between two neighbourhoods (A and B). Each neighbourhood contains one school, with exogenous quality q. There are equal numbers of school places in school A (S_A) and school B (S_B), and equal numbers of properties in neighbourhoods A and B (M).

Each family has an exogenous partisan preference (θ_A) for neighbourhood A, which could reflect location of family/friends, for example. This partisan preference is drawn from a continuous uniform distribution between -1 and 1, implying that some families have a preference for living in A while others have a preference for living in B. Note that $\theta_B = -\theta_A$.[5]

[4] In these areas, pupils who do not pass the test, known as the '11-plus', attend 'comprehensive' schools (previously called 'secondary modern' schools). These pupils are typically around 80% of those in the area.

[5] I have written it this way to simplify equation 1. It implies that the partisan preference is stronger, but I don't think it will make a material difference to any counterfactuals - just a matter of scale.

Families derive utility from the quality of the school their child attends (q_i), the match between their partisan preferences and neighbourhood (θ_i), and consumption (c_{ij}). Consumption is a function of exogenous income minus the equilibrium rent in their choice of neighbourhood ($c_{ij} = z_i - p_j$).

$$u(q_i, z_i, p_j) = \theta_{ij} + v(q_j) + w(c_{ij}) \tag{1}$$

Where consumption is equal to exogenous income minus endogenous rent in their chosen neighbourhood $c_{ij} = z_i - p_j$.

Utility of consumption is diminishing, such that $\frac{\delta w}{\delta c} > 0$ and $\frac{\delta^2 w}{\delta c \delta c} > 0$. This means that families with higher income value consumption less, and are more willing to trade-off consumption for school quality.

In the baseline model, utility of school quality is linear, such that $\frac{\delta w}{\delta q} > 0$ and $\frac{\delta^2 w}{\delta q \delta q} = 0$. This is consistent with existing empirical literature which has estimated a linear functional form for the relationship between school quality and local property prices.

In alternative models, utility of school quality is non-linear, such that $\frac{\delta w}{\delta q} > 0$ and $\frac{\delta^2 w}{\delta q \delta q} > 0$ (or $\frac{\delta^2 w}{\delta q \delta q} < 0$).

3.2 Housing market clearing

Denote demand for housing in neighbourhood A as D_A, and demand for housing in neighbourhood B as D_B. M denotes the housing stock in neighbourhoods A and B, given the fixed and equal housing supply in both neighbourhoods.

3.2.1 Catchment area system

Where catchment areas operate, living in neighbourhood d implies the family's child attending school d, with associated school quality q_d.

Each family chooses their preferred neighbourhood to maximise their utility:

$$v(q_d) + \theta_d + w(c_d) > v(q'_d) + \theta'_d + w(c'_d) \tag{2}$$

where $c_d = Z - p_d$. In equilibrium, p_d adjusts such that demand for properties in each neighbourhood equals the supply for properties in each neighbourhood. [Note: I think we could normalise this - so rent in neighbourhood B is fixed and we find equilibrium rent in A.]

3.2.2 Lottery system

In the absence of catchment areas, a lottery system could decide which pupils attend which school. This would mean that all households have expected school quality $E(q)$, independent of their residential location.

Each family's maximisation problem therefore becomes:

$$\cancel{E(v)} + \theta_d + w(c_d) > \cancel{E(v)} + \theta'_d + w(c'_d) \tag{3}$$

As in the catchment area case, in equilibrium, p_d adjusts such that demand for properties in each neighbourhood equals the supply for properties in each neighbourhood.

4 Data

4.1 Data sources

This section describes the publicly available data used to estimate the relationship between local school quality and property prices, and variation according to market characteristics.

Property prices: We take property prices from the Price Paid Data from HM Land Registry, covering all property sales in England and Wales that are sold for full market value and lodged with HM Land Registry for registration. These are linked to data from the Energy Performance Certificate database to observe more property details, such as the floor area.[6] Properties are geolocated within postcodes, small geographic areas containing typically 15 addresses.

School zones (catchment areas): We use catchment areas collected for a related project on schools' admissions arrangements in England, from the 2021/2022 academic year. We use catchment areas from 50 Local Authorities, where Geographic Information System (GIS) files were publicly available or requested by the research team through a 'Freedom of Information' request to the Local Authority.[7]

Appendix figure B.1 shows the full sample of school zones from the data collection exercise (N = 957).[8] These are around 52% of high schools that use a school zone in their admissions criteria in England (N = 1,827), and around 29% of the total number of high schools (N = 3,249). Burgess et al. (2023) show that 53% of high schools in England have a pre-defined catchment area in their admissions criteria. Most other high schools use a distance tie-breaker rather than a school zone, although a minority of high schools - around 10% - do not have any geographical criteria. Appendix figure B.2 maps LAs in England according to the use of catchment areas and inclusion in our sample. LAs included in our sample are shaded in dark blue. LAs where a sizeable share of secondary schools (more than 17%, with an average of 68%) have a school zone are shaded in light blue. LAs where school zones are uncommon (most notably in London) are transparent.

School quality and composition: Publicly available information on secondary school performance is from the official statistics available on the 'Find and compare schools in England' website. We use the percentage of pupils that achieve at least 5 GCSEs at grades 9 to 5 (including English and Maths), the percentage of pupils that achieve the English Baccalaureate, and the schools' value-added measure (known as Progress 8). For school composition, we use the percentage of pupils eligible for free school meals. For each of these measures, we take the average across 2017/2018 and 2019/2020 academic years.[9] The most recent Ofsted rating is coded from Ofsted management information. Each continuous measure is standardised to have a mean of zero and a standard deviation of one, according to the national distribution of secondary school performance.

Appendix figure B.3 shows the variation in school quality across school zones in England, for each of our four measures of school quality. Appendix figure B.4 repeats this for one LA in England,

[6] In future work, we hope to include rental prices to explore the heterogeneity across the sales/rental market.

[7] An additional 15 LAs provided information in excel or pdf format, which could be incorporated into our analysis at a later stage.

[8] For our analysis, we drop 40 school zones with no immediate neighboring school or where there is no clean boundary between the schools. This occurs when the school zones overlap to a large extent.

[9] Our data construction accounts for changes in school identifiers over time, for example, changes occurring due to conversion to academy status.

demonstrating the local variation across school zone boundaries. There are neighboring school zones with large differences in school quality according to each measure of school quality, for example, schools with an 'Outstanding' rating bordering 'Inadequate' schools. There are also many borders with little difference in school quality, however.

Resident population: We retrieve publicly available information about the composition of residents from the 2011 and 2021 UK census waves. Information are retrieved at the output area level, which is the most granular geographic disaggregation available: on average, one output area contains around 11 postcodes. We obtain information about the housing arrangement (tenure, accommodation type, number of rooms, bedroom occupancy), demographic characteristics (gender, age, ethnicity, migration status), economic activity (employment, student share, job classification) and education/job specialization (highest qualification).

4.2 Sample description

Population: The sample we consider is composed of postcodes (i.e. location of properties) that fall within catchment areas of neighbouring schools. Since catchment areas are used more in rural areas, and almost absent in London, our sample is not representative of England as a whole. The sample is representative of the Local Authorities which use catchment areas, however. Our results are therefore generalisable to areas where the majority of schools currently use catchment areas, and not the wider country.

Appendix Table A.2 shows the average differences in shares of people in our sample and outside our sample for 2021 census characteristics. The figures shown are population-weighted averages of the shares of people living within each output area.

People in our sample is more likely to live in a house, as opposed to a flat or shared home (Panel A), more likely to own this home (both outright and with a mortgage) and less likely to live in over-crowded homes. Consistent with our less-urban sample, we also find a higher share of people identifying as white and a lower share of people identifying as black. Other variables, such as economic activity, are balanced for our sample and England as a whole.

For schools, data from the Department for Education shown in Table A.4 contains characteristics of pupils in terms of gender (and attendance, panel A), free school meals (panel B) and ethnicity and language (panel C) for academic years 2016/17, 2017/18 and 2018/19.

The sample of students we consider does not differ from the general population in terms of gender, but students in schools within our sample are less likely to be eligible for free school meals. Students identifying as white and those with English as a first language are over-represented in our sample.

House prices and school quality: Appendix tables A.6 to A.8 present some summary statistics of our final dataset. Appendix table A.6 shows that the average property price is around £250,000, which increases slightly across samples with a larger bandwidth around school zone boundaries. In comparison to appendix table A.7, which shows the full sample of prices paid by year, prices are generally higher further from the boundary (average of £350,000 in 2021, for example). This could reflect the dis-utility of being far from a school or close to a boundary, or, alternatively, other correlated attributes such as distance from the city centre.

There are a reasonable number of transactions per border, with a mean and median above 100 each year (appendix table A.7). On average, each school has 4.5 neighbouring schools that it shares a border with.

Figure 1 shows the distribution of school quality differences across the school zone boundaries, for each measure of school quality. The median difference is around 0.4 of a standard deviation (see also appendix table A.8). Only around 10% of school pairs have a difference of more than one standard deviation, which is the metric commonly reported in the literature to present the relationship between school quality and prices.

5 Methodology

We follow the standard approach in the existing literature to identify the effect of the 'good' school on property prices using the boundary discontinuity design. However, our interpretation differs from previous work, which has either assumed or estimated little evidence of household sorting across the boundary. Instead, we interpret the price effect on the 'good' school side of the boundary as due to the first-order effect of the higher school quality and the second-order effect of endogenous neighbourhood composition. For example, suppose more educated households are more likely to congregate in neighbourhoods with access to a 'good' school. In that case, this might affect the residential choices of other households, who value the amenity of this type of neighbour, rather than the school quality directly. Kane et al. (2006) summarise, "Even if houses and neighbourhoods are very similar on either side of a school border when the boundary is originally drawn, the similarity may not last long as properties are bought and sold, as neighbors change, and as houses depreciate and are improved."[10] If the interpretation of the estimated effect is the value of school quality only, then Kane et al. (2006) conclude, "To the extent this sorting occurs, it will bias boundary estimates toward finding a positive association between school quality and property value unless one fully controls for these other differences across boundaries." Instead, we interpret the estimated effect on prices as the total effect of sorting for school quality and sorting for endogenous neighbourhood attributes arising from 'good' school quality.

5.1 Boundary discontinuity design

We adapt the traditional boundary discontinuity design approach to have a binary treatment variable rather than a continuous, standardised, measure of school quality. We do so for two reasons: first, our summary statistics show that fewer than 10% of boundaries show school quality differences larger than one standard deviation, which is the traditional interpretation in the literature. In later stages of our analysis, when we explore the price premiums around boundaries with very similar neighbouring schools, this would mean inappropriately extrapolating from very small differences. Secondly, we believe that a binary treatment indicator is more informative for a policy that targets the *presence* of a discontinuous jump in school access, which would be our policy of interest. A continuous treatment variable instead implies an interest towards intervention on school quality per se rather than access to a given school.

In other aspects, we follow the methodology from previous literature. First, we exclude school zone boundaries that coincide with administrative boundaries, as all comparisons across school zones are within Local Authority. Also, boundaries that coincide with physical boundaries such as large rivers, main roads and railway tracks (see Black (1999), Gibbons and Machin (2003), Bayer et al. (2007), Davidoff and Leigh (2008)). We take school quality as an average across multiple years (see Black (1999), Kane et al. (2006), Gibbons and Machin (2006), Bayer et al. (2007), Fack and Grenet

[10] Similarly, Bayer et al. (2007) state, "Given a discontinuity in local school quality at a school boundary, one might expect that residential sorting would lead to discontinuities in the characteristics of households residing on opposite sides of the same boundary; even if a school boundary was initially drawn such that the houses immediately on either side were identical, households with higher incomes and education levels might be expected to sort onto the side with the better school."

(2010), Ries and Somerville (2010), Gibbons et al. (2013), Chung (2015), La (2015), Harjunen et al. (2018), Bibler and Billings (2020)) and use log property prices (universally done).

Let i denote a (sold) property. Let s_i be the school zone in which property i is located, and s_{-i} be the neighbouring school zone. Then, the pair $\{s_i, s_{-i}\}$ defines the boundary b. We define 'treatment' as being located on the side of the boundary with the highest school quality (SQ), where the definition of school quality varies by j:

$$T_{ib}^{j} = \begin{cases} 1 & \text{if } SQ_i^j > SQ_{-i}^j \\ 0 & \text{if } SQ_i^j < SQ_{-i}^j \\ NA & \textit{otherwise} \end{cases} \tag{4}$$

We use four measures of school quality: student attainment ($j = att$), student value added ($j = vadd$), student socioeconomic ($j = ses$) and Ofsted rating ($j = oval$).[11] For conciseness, we focus on attainment and socioeconomic status for our main analysis, which explores the non-linearity of the price premiums to boundary characteristics.[12]

The estimating equation is:

$$y_{ibt} = \tau^j T_{ib}^j + \beta H_{ibt} + d_{ib} + \theta^b + \theta^t + \epsilon_{ibt} \tag{5}$$

Where y_{ibt} is the log property price for property i near boundary b at time t. T_{ib}^j is the treatment indicator based on measure j for property i in border b. H_{ibt} are characteristics of property i near boundary b at time t (including the type of property, new build, and floor area).d_{ib} is the distance of property i from boundary b. θ^b are boundary fixed effects that account for all neighbourhood characteristics that are common to properties across the boundary, for example, the housing stock, proximity to transport links and public amenities such as parks. θ^t are year fixed effects which account for general variation in property prices over time.

Our main specification restricts the sample to properties within 320m of a school zone boundary, which is equivalent to 0.2 miles, the standard in the literature. Robustness checks vary the bandwidth around the school zone boundary (100m, 200m, 500m and optimal bandwidth computed following Calonico et al. (2014)).[13] Our identifying assumption is that within this close range, exogenous neighbourhood characteristics, such as access to public transport and the age of the housing stock, are constant.

Our main specification controls for property characteristics that are most arguably exogenous to the school admissions system and household sorting. Appendix figure B.7 shows that these property characteristics are largely balanced across the boundary. The only exception is distance from the assigned school, which is systematically further on the higher quality side of the boundary. We present alternative specifications that do not control for property characteristics and control for a fuller set of property characteristics that could arguably be endogenous to the type of households that sort into the area. For example, more affluent households could be more likely to install double-glazed windows.

[11] Please refer to the discussion in Section 4.1 for a more detailed description.

[12] Full results are in the appendix. We replicate the findings from existing literature using all four school quality measures.

[13] This is a typical approach in the existing literature. For example, Black (1999) uses bandwidths of 0.15, 0.2 and 0.35 miles. Similarly, Kane et al. (2006) use 0.4, 0.2 and 0.1 miles. Bayer et al. (2007) use 0.2 and 0.1 miles. Davidoff and Leigh (2008) use 600m, 500m and 200m, while Fack and Grenet (2010) use 250m, 300m and 350m. At the higher end, Gibbons and Machin (2003) use 1000m, although properties are matched to the nearest property on the other side of the boundary, and Cheung et al. (2022) use 500m, 800m and 1000m.

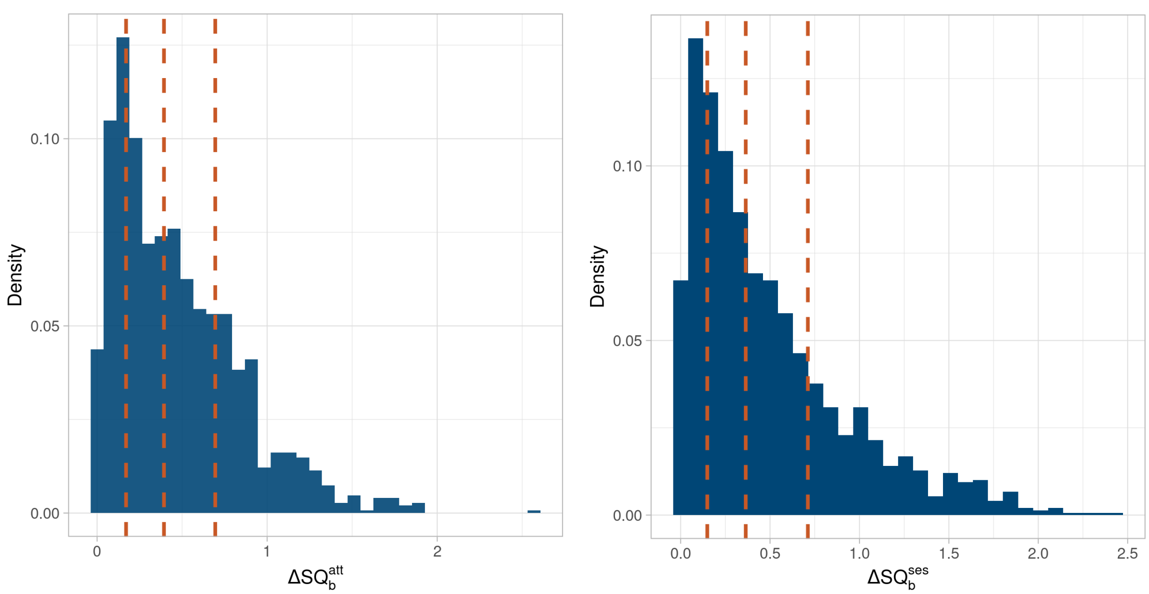
5.2 Non-linearities in the relationship school quality differences and price premiums

The standard interpretation in the existing literature is the effect of a one standard deviation increase in school quality on property prices. This might not be appropriate if there are many boundaries with only small differences in school quality. To explore this, we define ΔSQ_b^j as the difference in school quality between two neighboring schools, for each boundary b and school quality measure j.

$$\Delta SQ_b^j = |SQ^j s_i - SQ^j s_{-i}| \tag{6}$$

Figure 1 plots the distribution of ΔSQ^j for one measure of attainment (on the left) and socioeconomic status (SES, on the right), where the orange vertical lines show the quartiles. Figure 1 shows that many borders have marginal differences in school quality, and only around 10% have a difference of one standard deviation or more.

Figure 1: Distribution of school 'quality' differences across school zones in England



(a) Attainment

(b) Socioeconomic status

Source: DfE School Performance Tables and school zones collected by the authors. Notes: Orange dashed lines represent the quartiles of each distribution, where the cutoff values are {0.17,0.39,0.7} for attainment and {0.15,0.36,0.71} for socioeconomic status.

To explore whether the relationship between school quality differentials and price premiums are indeed linear (as assumed by previous literature) or non-linear, we assign each boundary a quartile of ΔSQ_b^j. Let $\Delta SQ_b^{j,Q}$ denote this variable: $\Delta SQ_b^{j,Q} = Q1$ includes boundaries where neighboring schools have similar school quality, while $\Delta SQ_b^{j,Q} = Q4$ includes boundaries where neighboring schools are very different.

We test the hypothesis that there are non-linear effects by estimating the following equation:

$$y_{ibt} = \tau^{j,Q} T_{ib}^{j} \times \Delta SQ_{b}^{j,Q} + \beta H_{ibt} + d_{ib} + \theta^{b} + \theta^{t} + \varepsilon_{ibt} \tag{7}$$

Equation 7 includes the full set of interactions between our treatment variable and the quartile indicator for school quality difference: the estimated $\widehat{\tau^{j,Q}}$ describe the price premium (if any) observed for schools in quartile Q of the distribution of ΔSQ_{b}^{j}. In other words, $\tau^{j,1}$ estimates the price premium for those boundaries defined by very similar schools and $\tau^{j,4}$ estimates it for boundaries defined by very different schools.

We interpret each $\tau^{j,Q}$ as estimates of the price premium given by different levels of similarity between neighboring schools. If there are non-linear effects, such that price premiums are concentrated where the difference in school quality is large, then we should observe a larger and significantly different $\tau^{j,4}$ from the other estimates.

5.3 Paying for 'superstars' or avoiding the worst?

We can also learn about demand for absolute school quality, even from the boundary discontinuity methodology, which by design compares the price *difference* relative to the school quality *difference* across the boundary. This is because there is variation in the level of school quality for the same difference in school quality across the boundary. For example, a boundary with a big difference in school quality could compare an average school to a 'superstar', or a 'bad' school to an average school.

Figure 2 illustrates the variation in the quality of the treated (or higher-performing) school within quartiles of $\Delta SQ_{b}^{j,Q}$, as measured by attainment (left) and socioeconomic status (SES, right). The boxes show the 25th, 50th and 75th percentiles, while the dots represent outlier schools with particularly high or low quality.

Although where the difference in school quality is larger ($Q4$) school quality of the treatment school is typically higher, this is not necessarily the case. In both panels, the higher-performing school in $Q4$ can be a below-average school. It is common for the comparison to be between a poorly and median performing school. This means that the price premium between different schools also arise from comparisons at the bottom of the distribution, in addition to 'superstar' schools at the top.

This distinction is informative about the incentives in place for people to pay a premium: if we only observe a price premium generated by top-schools versus average-schools, we can infer that people value high-quality schools *per se*. If, instead, we observe premiums arising from comparisons of average schools with low-ranking schools, we can infer that people are also willing to pay in order to 'escape' schools in the bottom of the quality distribution.

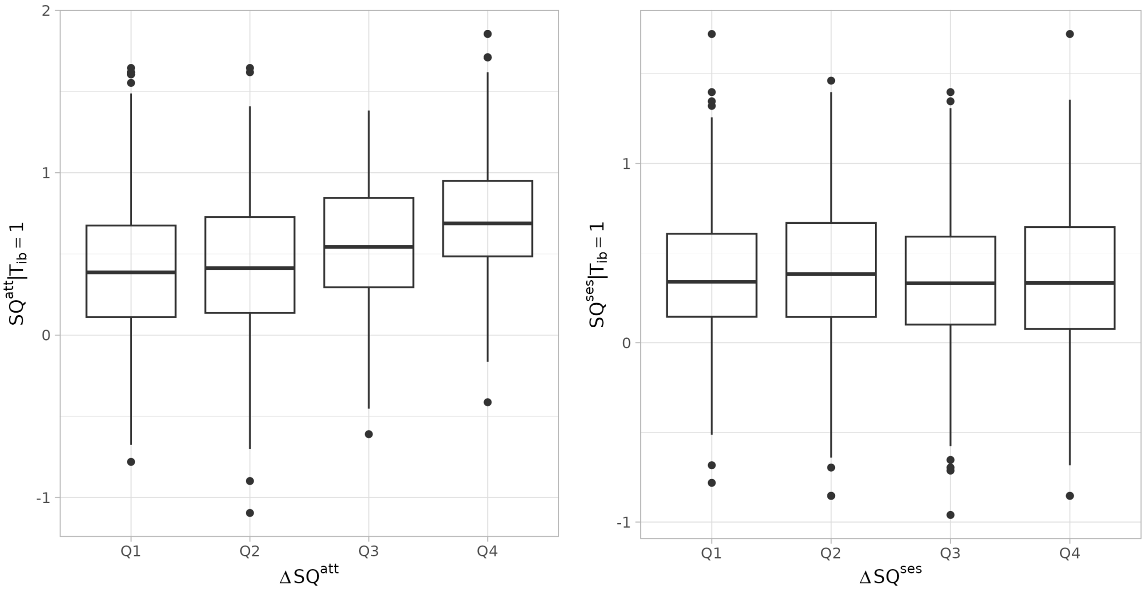
We define a variable describing the position of the neighboring schools in the quartiles of the quality distribution for each boundary, to test whether people pay a premium to avoid low quality schools. Let $SQ_{s_i}^{Q}$ define the position of school s_i in the school quality distribution and $SQ_{s_{-i}}^{Q}$ define the analogous measure for school s_{-i}. We define:

$$SQ_{b}^{comb} = SQ_{s_i}^{Q}, SQ_{s_{-i}}^{Q} \tag{8}$$

i.e. the indicator of the relative position of the two schools defining boundary b. Suppose that we have $SQ_{s_i}^{Q} = Q2$ and $SQ_{s_{-i}}^{Q} = Q1$: in this case $SQ_{b}^{comb} = Q1, Q2$

In order to test whether price premiums arise from comparisons of 'bad' schools (i.e. those for which we observe $SQ^{Q} = Q1$) and below average schools (i.e. those for which we observe $SQ^{Q} = Q2$) we estimate the following equation:

Figure 2: Distribution of school 'quality' for treated schools by ΔSQ quartile



(a) Attainment

(b) Socioeconomic status

Source: DfE School Performance Tables and school zones collected by the authors. Notes: This figure shows the school quality of the treatment school for all pairs of schools with a boundary in our final sample. The box and whisker plots represent the 25th, median and 75th percentiles for the school quality of the treatment school, with outliers represented by the dots in the whiskers. The four columns split the sample by quartiles for the difference in school quality across the boundary. For example, $Q4$ has the largest difference in school quality across the boundary.

$$y_{ibt} = \tau^{j,C} T^{j,C}_{i,b} \times SQ^{comb}_b + \beta H_{ibt} + \theta^b + \theta^t + \nu_{ibt} \tag{9}$$

where the superscript C indicates values of SQ^{comb}_b.

We estimate equation 9 within Q4 of ΔSQ^j_b to test our prediction: we expect $\tau^{j,Q1,Q2}$ to be positive and significantly different from 0 if people are willing to 'escape' bad schools at the bottom of the distribution to move to schools that are, by definition, below average.

5.4 Multidimensionality of school quality

Households consider multiple school characteristics when choosing where to send their child (Hastings et al. (2009), Borghans et al. (2015), Burgess et al. (2015), Denice and Gross (2016), Glazerman and Dotter (2017), Akyol and Krishna (2017), Beuermann et al. (2018), Fack et al. (2019), Harris and Larsen (2019), Ruijs and Oosterbeek (2019), Abdulkadiroğlu et al. (2020), Bertoni et al. (2020), Walker and Weldon (2020)). School performance and pupil composition are important factors in parents' choices, but are often highly correlated. We explore how these factors interact to shape households' residential choices, as inferred through property prices.

Although there is a significant positive correlation for high schools in England (see Figure B.8) there are cases where schools have high performance with a less advantaged socio-economic composition, and vice versa. We exploit this variation to inform whether households appear to make tradeoffs between school characteristics.[14]

Our specification compares schools across the boundaries within quartiles of difference in school quality. As in previous analyses, $Q4$ represents boundaries with the largest difference in school quality.

The second dimension is the difference in socioeconomic status. Analogously to equation 6, we define ΔSES as the absolute difference in socioeconomic status across the boundary:

$$\Delta SES = |SESs_i - SESs_{-i}| \tag{10}$$

ΔSES is then transformed into five distinct binary categories, ΔSES_C, for whether the 'treatment' school according to attainment has a 'better', the 'same' or 'worse' school composition than the 'control' school (with lower attainment), and to a 'medium' or 'large' extent.[15] These binary variables are interacted with the treatment indicator in our main specification to inform whether the difference in socioeconomic status amplifies or mediates the effect of the difference in attainment across the boundary on prices.

The estimating equation is:

$$y_{ibt} = \tau^{j,C} T_{i,b}^{j,C} \times \Delta SES_C + \beta H_{ibt} + \theta^b + \theta^t + \nu_{ibt} \tag{11}$$

where the superscript C indicates values of ΔSES_C.

6 Results

6.1 Price premiums

Table 1 shows that, in general, our results follow those from the previous literature. There is around a 2.5% increase in property prices for properties on the side of the boundary with the high-performing school when the measure of school quality is defined by pupil performance. The price premium is slightly lower when school quality is defined by socio-economic status (around 2.2%), value-added (around 1.8%) and Ofsted (around 0.1%, statistically insignificant).[16]

The results are sensitive to whether property characteristics are accounted for. Appendix tables A.13 to A.16 show the estimates whether accounting for property characteristics or not, and for different sets of property characteristics. Across all definitions of school quality, the price estimates

[14] We recognise that the socio-economic composition of pupils is not a typical or desirable measure of school quality. We classify it this way for simplicity, recognising that it is an attribute commonly considered by parents.

[15] Precisely, we define these according to the distribution of the absolute difference in socioeconomic status at the boundary. 'Better' means the 'treatment' school has a more socioeconomically advantaged population, and the absolute difference in socioeconomic status between the 'treatment' and 'control' school is greater than the 25th percentile. 'Same' means the absolute difference in socioeconomic status between the 'treatment' and 'control' school is below the 25th percentile. 'Lower' means the 'treatment' school has a less socioeconomically advantaged population, and the absolute difference in socioeconomic status between the 'treatment' and 'control' school is greater than the 25th percentile. 'Medium' means that the absolute difference is between the 25th and 75th percentile, while 'large' means that the absolute difference is greater than the 75th percentile.

[16] Note that the sample for Ofsted is necessarily smaller as school pairs are more likely to have the same value, so no 'treatment' can be defined.

Table 1: Estimation results for the relationship between the difference in local school quality and property prices

	Attainment (1)	Ofsted val. (2)	SES (3)	Value added (4)
τ^{att}	0.0255***			
	(0.0072)			
τ^{oval}		0.0112		
		(0.0089)		
τ^{ses}			0.0216***	
			(0.0071)	
τ^{vadd}				0.0182***
				(0.0064)
EPC covs	Y	Y	Y	Y
Bandwidth (m):	320	320	320	320
Observations	77,857	39,324	75,291	77,828
R^2	0.78712	0.78309	0.78766	0.78775
Within R^2	0.59611	0.58328	0.59445	0.59482
Border fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓

Notes: Signif. codes: ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered at the border level. EPC covariates include floor area, property type, construction age band and distance from assigned school. All specification control linearly for distance to the boundary.

are larger in magnitude when no property characteristics are accounted for (column 1). This suggests that properties on the 'high' school quality side of the boundary typically have characteristics that are associated with higher prices, such are more habitable rooms. Estimates are very similar between specifications that include a smaller and larger (more/less likely to be exogenous) set of property characteristics, however.

The estimates are sensitive to the bandwidth around the school zone boundaries used, typically increasing with distance from the boundary. Appendix table A.9 shows that within 100m of the boundary, there is no significant effect, with point estimates less than 1%. Within 200m of the boundary, there is a smaller effect of around 2% or under, depending on the measure of school quality (Appendix table A.10). Effects are similar to the main results for the 500m bandwidth (around 3% for attainment and socioeconomic status) and for the optimal bandwidth selection (around 2% for attainment and socioeconomic status).

6.2 Non-linearities in the relationship school quality differences and price premiums

We find that price premiums are largest in boundaries where the difference in school quality is largest ($Q4$) and are not present in boundaries where the difference in school quality is smallest ($Q1$). Table 2 shows estimates of the price premiums within quartiles of school quality difference for student attainment (left) and student SES (right). The first row presents the estimated price premium between schools that are very similar in terms of quality ($Q1$ of the difference distribution), while the fourth row estimates the premium for schools that are very different in terms of quality ($Q4$ of the difference

Table 2: Estimation results for the non-linearity of the relationship between school quality differences and property price premiums

	Attainment (1)	SES (2)
$\tau^j \times \Delta SQ^j$ = Q1	0.0079	0.0071
	(0.0144)	(0.0160)
$\tau^j \times \Delta SQ^j$ = Q2	0.0260**	0.0141
	(0.0123)	(0.0142)
$\tau^j \times \Delta SQ^j$ = Q3	0.0209*	0.0054
	(0.0110)	(0.0100)
$\tau^j \times \Delta SQ^j$ = Q4	0.0470***	0.0531***
	(0.0117)	(0.0125)
EPC covs.	Y	Y
Bandwidth (m):	320	320
Observations	77,857	75,291
R^2	0.78727	0.78794
Within R^2	0.59638	0.59498
Border fixed effects	✓	✓
Year fixed effects	✓	✓

Notes: Signif. codes: ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered at the border level. EPC covariates include floor area, property type, construction age band and distance from assigned school. All specification control linearly for distance to the boundary.

distribution).

We estimate premiums of around 1% for schools in $Q1$ and up to around 5% for schools in $Q4$. Except for $Q1$ (and $Q3$ for SES), all estimates are statistically significant at conventional levels and economically meaningful. These results suggest that, in general, the larger the difference between a school and its neighbor, the larger the average difference in house prices.

We conclude that the price premium is larger where the difference in school quality is large through a statistical test. Appendix Table A.17 shows the pairwise test for coefficient equality for all the estimates in table 2. The first column shows the hypothesis being tested, while the other two display the estimated difference and indicate the significance levels for the chosen quality measures. Table A.17 shows that, for both measures of school quality, the premiums within $Q4$ (i.e. very different schools) are significantly different from those within the other quartiles, except for $Q2$ for attainment. The pairwise differences between the lower quantiles are also smaller in magnitude and not different from each other at conventional significance levels. These findings are robust to using a wider, optimal bandwidth as shown in Appendix tables A.18 and A.19.

Overall, these results demonstrate that the effect of school quality on property prices is driven by contexts where the difference in school quality across school zone boundaries is largest. The standard interpretation of previous estimates therefore requires more nuance.

Table 3 compares the implications of the results presented by the existing literature with our estimates. In the first column, we present prices and median income for the full England sample (i.e. all areas) and compute the price premium implied by the 3% effect estimated in the literature. The second column shows the same information for the locations in our data that fall within catchment areas, using the (approximated) estimate from column (1) in table 1. The last column displays the same variable for the subsample of areas *within boundaries in the last quartile of* ΔSQ^{att}, meaning for those locations in which school quality is different enough to give rise to a premium.

These 'back-of-the-envelope' calculations show that the linearity assumption of the existing literature underestimates the magnitude of the implications: an effect of 3% for the entire sample implies premiums of around 60 to 65 percent of the median income of residents, while an effect of 5% for the sub-sample close to school zone boundaries of very different schools implies a premium of a full year of median income.

Table 3: Effect comparisons across samples, 2018 income and prices

	England	Catchment areas	Q4 of attainment difference
Population	59595327	21689487	12317144
Median income (£)	15831	15604	15742
Avg. house price (£)	348572	308713	315491
Estimated effect	3%	3%	5%
Implied price premium (£)	10457	9261	15775
Share of median income (%)	66	59	100
Months of median income	8	7	12

6.3 Paying for 'superstars' or avoiding the worst?

Price premiums could arise in competition for 'superstar' schools, as is typically implied and studied in previous literature. In contrast, price premiums could also arise to avoid the 'worst' schools, which would have different implications for parents' preferences for school quality. To explore this, we present the estimates of equation 9. We focus on boundaries within $\Delta SQ_b^j = Q4$: those boundaries where there is a large difference in school quality (typically associated with larger property premiums).

Column 1 of table 4 shows the results using student attainment as the measure of quality. Each row shows the effect for a specific combination of school-quality quartiles across the boundary. For example, the first row shows the price premiums where school quality is very different across the boundary, but where both school are located at the bottom of the quality distribution(in $Q1$).

We find a large, significant premium of around 13% for schools around the median of the distribution ($Q2$), compared to those at the bottom ($Q1$). We interpret this as suggestive that people are indeed willing to pay a sizable premium in order to avoid a school at the bottom of the distribution[17]. The price premiums are typically positive across the distribution, however only the $Q1$ to $Q2$ and $Q3$ ones are significant for attainment. Premiums in column 2 instead do not seem to reflect an 'escaping the bottom' behavior, but rather willingness to pay to access the best possible peer composition (in terms of SES).

6.4 Multidimensionality of school quality

This section explores how households respond to multiple dimensions of school quality, or characteristics. We focus on school attainment and socioeconomic status as these are the two most frequently documented attributes of schools the parents value.

[17] The result is robust to using optimally calculated bandwidths around the cutoff, although point estimates and significance levels slightly change. See Appendix table A.20

Table 4: Estimation results for price premiums across boundaries with varying levels of school quality

	Attainment (1)	SES (2)
$\tau^j \times SQ^{comb}$ = Q1-Q1	0.0052	
	(0.0177)	
$\tau^j \times SQ^{comb}$ = Q1-Q2	0.1362***	
	(0.0114)	
$\tau^j \times SQ^{comb}$ = Q1-Q3	0.0258*	0.0445
	(0.0155)	(0.0282)
$\tau^j \times SQ^{comb}$ = Q1-Q4	0.0249	0.0712*
	(0.0295)	(0.0367)
$\tau^j \times SQ^{comb}$ = Q2-Q4	0.0349	0.0428*
	(0.0258)	(0.0251)
$\tau^j \times SQ^{comb}$ = Q3-Q4	0.0258	0.0533**
	(0.0207)	(0.0233)
$\tau^j \times SQ^{comb}$ = Q4-Q4	-0.0156	0.0413*
	(0.0283)	(0.0240)
$\tau^j \times SQ^{comb}$ = Q2-Q3		-0.0502
		(0.0522)
Observations	21,145	23,013
R^2	0.79108	0.75480
Within R^2	0.61509	0.57415
Border fixed effects	✓	✓
Year fixed effects	✓	✓

Notes: Signif. codes: ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered at the border level. EPC covariates include floor area, property type, construction age band and distance from assigned school. All specification control linearly for distance to the boundary.

Where there are large differences in school attainment, the property price response is unaffected by the relative differences in socioeconomic status. The first column of table 5 focuses on boundaries where the difference in school attainment are largest ($Q4$ of the distribution of ΔSQ). The results show that households are willing to pay a premium to live in the school zone of a high-performing school with a much less advantaged pupil composition.

When there are less pronounced differences in school attainment, the property price response depends on the relative difference in socioeconomic status. The second column of table 5 shows that where the difference in school attainment is in the third quartile ($Q3$), households are only willing to pay a premium to live in the school zone of a high-performing school with a more advantaged socioeconomic status.

These findings suggest that households make tradeoffs between school characteristics. Relative socioeconomic status across the boundary becomes irrelevant only where school attainment is significantly higher on one side of the boundary. This could imply that households' preferences are lexicographic past some point of the difference in school attainment - households are willing to pay more for higher attainment, regardless of socioeconomic status.

Table 5: Estimation results for households' willingness to pay for relative socioeconomic status across the boundary, for given relative school attainment

	$\Delta SQ^{att} = Q4$ (1)	$\Delta SQ^{att} = Q3$ (2)	$\Delta SQ^{att} = Q2$ (3)	$\Delta SQ^{att} = Q1$ (4)
$\tau^{att} \times \Delta SQ^{ses}$ = Better,large	0.0593***	0.0549***	0.0807	-0.0441
	(0.0165)	(0.0207)	(0.0546)	(0.0423)
$\tau^{att} \times \Delta SQ^{ses}$ = Better,med	0.0140	0.0158	0.0132	0.0289
	(0.0169)	(0.0134)	(0.0290)	(0.0205)
$\tau^{att} \times \Delta SQ^{ses}$ = Same	0.0358	-0.0179	0.0238	0.0099
	(0.0233)	(0.0292)	(0.0163)	(0.0297)
$\tau^{att} \times \Delta SQ^{ses}$ = Worse,med	0.0594***	-0.0279	-0.0127	0.0288
	(0.0103)	(0.0474)	(0.0253)	(0.0182)
$\tau^{att} \times \Delta SQ^{ses}$ = Worse,large			0.0107	-0.0434***
			(0.0785)	(0.0096)
Observations	21,145	18,976	16,669	19,651
R^2	0.79029	0.82314	0.76487	0.81060
Within R^2	0.61363	0.68774	0.54396	0.61731
Border fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓

Notes: Signif. codes: ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered at the border level. Estimation results from equation 9. Estimation results from equation 11. Covariates include property type, total floor area, construction age band and distance from assigned school.

6.5 Residential sorting

Our interpretation of the previous results is the total effect of a discontinuous jump in the probability of accessing the 'better' school. As described in Section 5 we expect there to be residential sorting due to both the quality of the school and the endogenous residential composition. In this section we provide some evidence of residential sorting using Census data, by comparing the composition of residents on either side of the boundary.

We choose to treat the residents' composition measured in 2011 as outcomes rather than pre-treatment covariates because we acknowledge that the catchment area system has been in place for a long time, since before 2011. Unfortunately, there is no tracking of changes to school zones over time by LAs: we maintain the assumption that the current catchment areas did not dramatically change over the years.

For this reason we expect 2011 Census variables to show residential sorting due to the presence of catchment areas: hence our decision to use them as outcomes.

6.5.1 Estimation

We estimate the following equation for a set of variables describing the composition of residents on either side of a boundary, within a given bandwidth:

$$f^k_{bc} = \gamma T^j_{bc} + d_{bc} + \theta^b + \eta_{bc} \tag{12}$$

where b denotes a boundary, c denotes a census output area, f_{bc}^k is the fraction of residents in area c and boundary b who display characteristic k (e.g. f_{bc}^{emp} is the share of employed residents for area c) and T_{ib}^j is the indicator for being on the side of the higher 'quality' school. The term d_{bc} measures the average distance of postcodes (i.e. properties) from the boundary within each output area. For brevity, we focus on $j = att$ in this section.

The parameter γ measures the average difference in residents' composition (in percentage points) across boundaries.

6.5.2 Results

Table 6: Estimation results for equation 12 using economic activity, ethnicity and migration status with respect to previous year. Bandwidth of 320m.

	2011 Census		2021 Census	
	Sample avg.	Difference	Sample avg.	Difference
Panel A: Economic activity				
Employed	0.628	-0.002	0.564	-0.004
		(0.003)		(0.002)
Unemp.	0.042	-0.002***	0.026	-0.001***
		(0.001)		(0)
Active students	0.034	0.001*	0.024	0
		(0.001)		(0.001)
Inactive student	0.056	0.003**	0.054	0.002
		(0.002)		(0.001)
Student (inc. schoolchildren)	0.012	0.001***	0.198	0.003
		(0)		(0.002)
Panel B: Ethnicity				
White	0.863	-0.003	0.824	-0.003
		(0.003)		(0.003)
Mixed	0.02	0	0.026	0
		(0)		(0)
Asian	0.086	0.002	0.104	0.004
		(0.003)		(0.003)
Black	0.022	0	0.027	-0.001
		(0.001)		(0.001)
Other	0.009	0.001	0.018	0
		(0)		(0)
Panel C: Migration				
Stayers	0.871	-0.001	0.883	0.001
		(0.002)		(0.002)
Migrant UK	0.117	0	0.099	-0.002
		(0.002)		(0.001)
Migrant non-UK	0.011	0	0.01	0.001*
		(0)		(0)

Notes: Signif. codes: ***:0.01, **:0.05, *:0.1. Standard errors are clustered at the boundary level. Columns 1 and 3 display the sample means of each outcome (i.e. row) for, respectively, 2011 and 2021. Columns 2 and 4 show the estimated average difference across the border for, respectively, 2011 and 2021.

Table 6 shows the results of estimating equation 12 using demographic characteristics of residents as outcomes. Columns 2 and 3 refer to estimates using the 2011 census variables and columns 4 and 5 use the 2021 variables instead. We estimate statistically significant differences for the share of unemployed residents for both waves (a sizeable difference of 5% of baseline), showing lower rates for the higher attaining side of the boundary. In 2011 we also observe a higher average share of student for the better-performing side, of about 2% of the baseline share.

Table 7: Estimation results for equation 12 using educational qualification and Ns-SeC job categorization. 320m bandwidth.

	2011 Census		2021 Census	
	Sample avg.	Difference	Sample avg.	Difference
Panel A: Highest qualification				
None	0.222	-0.009***	0.179	-0.008***
		(0.002)		(0.002)
Level 1	0.135	-0.003***	0.098	-0.002***
		(0.001)		(0.001)
Level 2	0.153	-0.001	0.134	-0.001
		(0.001)		(0.001)
Apprenticeship	0.038	0	0.055	0
		(0)		(0)
Level 3	0.127	0.003*	0.172	0.001
		(0.002)		(0.001)
Level 4	0.272	0.011***	0.334	0.011***
		(0.003)		(0.003)
Other	0.054	-0.001**	0.027	0
		(0.001)		(0)
Panel B: Ns-SeC occupation				
High managerial	0.107	0.006***	0.132	0.007***
		(0.001)		(0.002)
Low managerial	0.206	0.005***	0.199	0.005***
		(0.002)		(0.001)
Intermediate occ.	0.128	0.002**	0.116	0.001**
		(0.001)		(0.001)
Small independent worker	0.092	0.001	0.101	0.001**
		(0.001)		(0.001)
Low technical	0.071	-0.003***	0.055	-0.002***
		(0.001)		(0)
Semi-routine	0.142	-0.005***	0.115	-0.005***
		(0.001)		(0.001)
Routine	0.114	-0.007***	0.127	-0.007***
		(0.001)		(0.001)
Unemployed	0.053	-0.003***	0.081	-0.003***
		(0.001)		(0.001)
Student	0.088	0.005**	0.076	0.003
		(0.002)		(0.002)

Notes: Signif. codes: ***:0.01, **:0.05, *:0.1. Standard errors are clustered at the boundary level. Columns 1 and 3 display the sample means of each outcome (i.e. row) for, respectively, 2011 and 2021. Columns 2 and 4 show the estimated average difference across the border for, respectively, 2011 and 2021.

Table 7 displays results for highest attainment (Panel A) and occupation (Panel B): in both panels, for both census years, we find that more qualified people (either higher educated or more professionally specialized) and higher earners sort on the side of the more desirable school in terms of attainment. The point estimates are very similar across the two census waves and they range from around 3% to around 6% of baseline shares.

We interpret this evidence as suggestive that the relevant dimension for sorting is that of earnings, consistently with the idea that people who are more likely to be able to pay a premium on house purchase are the ones who actually decide to do so. This may suggest that the public concern[18] of pricing poorer households out of higher quality education is a relevant one. Lastly, we do not find strong evidence of sorting along ethnicity, migration and economic activity: this could reflect the notion that neighbourhood amenities are equally accessible on either side of the boundary and therefore do not induce differential sorting.

[18] Sutton Trust, 2024.

7 Conclusion

This work aims to reframe our understanding of how geographic school admissions arrangements affect residential sorting. From previous literature, the conventional wisdom is that sorting for school quality is widespread. Around school zone boundaries worldwide, access to a 'good' school increases local property prices, typically around 3% for a one standard deviation increase in school quality.

This interpretation is unsuitable for the context. For English high schools, we show that across all school pair boundaries, only around 10% of school pairs have a difference of at least one standard deviation. Across most school boundaries, the difference in quality is small, and, consequently, so is the difference in prices. This nuance is important, as it changes our interpretation of how the public school system affects households' welfare.

Greaves and Turon (2023) model households' welfare from the choice of neighbourhood and school, for both parent and non-parent households. One key intuition from this model is that there can be externalities from the public school system when geographic admissions criteria are used, for both parent and non-parent households. First, for non-parent households who only directly value neighbourhood quality, their welfare can be affected by being 'displaced' from an otherwise preferred neighbourhood due to higher prices, and, if not displaced, paying higher prices in a neighbourhood with inflated prices due to the local school. Second, for parent households, an additional consideration is the probability of accessing their preferred school. The overall welfare implications for parent households, therefore, depend on their relative utility between paying higher prices and the probability of access to the 'good' school or being displaced.

Our results imply that these mechanisms are only directly relevant in a minority rather than most areas. This means that the welfare of households choosing their residential location in these areas is not directly affected by the geographical school admissions arrangements of the public school system. They remain indirectly relevant, of course, as there can be spillovers from areas with residential sorting for school quality to those without.

Our results also have implications for the general equilibrium effects of reforming school admissions criteria. Where sorting for school quality is currently minimal, the general equilibrium effects of removing geographic school admissions criteria would be second-order, affected only by changing residential demand from households previously considering other neighbourhoods. In contrast, the price premiums is around 5% where the difference in school quality is the largest.

Using very simple back-of-the-envelope calculations we show that, in areas where sorting is significant, the price premium amounts to a full year of median income. Instead, the same calculation assuming the linear 3% effect established in the literature implies a premium of 7 months of median income.

Our results show that willingness to pay a premium is not necessarily concentrated at the top of the quality distribution: we also see significant price differences in places where the alternatives being compared are not top schools. This is consistent with the idea that people are willing to pay to avoid the worst quality school and not only for a top institution, which adds to the existing understanding in the literature.

Finally, we find that households make tradeoffs between dimensions of school characteristics. Households are willing to pay for higher school attainment if the socioeconomic status of the school is also higher, unless the difference in school quality is large. This highlights the multidimensionality of school 'quality' and variations in price premiums across contexts.

We provide evidence that residential sorting is mainly detectable along the income/education dimension, supporting the idea that higher-earning households are the ones willing (and able) to pay a premium on housing, while we cannot support sorting along broader characteristics.

In many parts of the world, school access depends partly, if not completely, on residential location. Overall, our analysis implies that the welfare effects of such policies, in comparison to de-linking residential and school choices, are heterogeneous. Specific contexts imply strong demand, and therefore higher prices and displacement of households that do not value school quality. Reforms to school admissions criteria or school allocation systems are likely to lead to concentrated rather than widespread disruption of housing markets.

References

Atila Abdulkadiroğlu, Parag A. Pathak, Jonathan Schellenberg, and Christopher R. Walters. Do parents value school effectiveness? *American Economic Review*, 110(5):1502–39, May 2020. doi: 10.1257/aer.20172040. URL https://www.aeaweb.org/articles?id=10.1257/aer.20172040.

Pelin Akyol and Kala Krishna. Preferences, selection, and value added: A structural approach. *European Economic Review*, 91:89 – 117, 2017. ISSN 0014-2921. doi: https://doi.org/10.1016/j.euroecorev.2016.09.009. URL http://www.sciencedirect.com/science/article/pii/S0014292116301489.

Patrick Bayer, Fernando Ferreira, and Robert McMillan. A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, 115(4):588–638, 2007. doi: 10.1086/522381. URL https://doi.org/10.1086/522381.

Venla Bernelius and Katja Vilkama. Pupils on the move: School catchment area segregation and residential mobility of urban families. *Urban Studies*, 56(15):3095–3116, 2019. doi: 10.1177/0042098019848999. URL https://doi.org/10.1177/0042098019848999.

Marco Bertoni, Stephen Gibbons, and Olmo Silva. School Choice During a Period of Radical School Reform. Evidence from Academy Conversion in England. *Economic Policy*, 11 2020. ISSN 0266-4658. doi: 10.1093/epolic/eiaa023. URL https://doi.org/10.1093/epolic/eiaa023. eiaa023.

Diether Beuermann, C. Kirabo Jackson, Laia Navarro-Sola, and Francisco Pardo. What is a good school, and can parents tell? evidence on the multidimensionality of school output. Working Paper 25342, National Bureau of Economic Research, December 2018. URL http://www.nber.org/papers/w25342.

Andrew Bibler and Stephen B. Billings. Win or lose: Residential sorting after a school choice lottery. *The Review of Economics and Statistics*, 102(3):457–472, 2020. doi: 10.1162/rest_a_00868. URL https://doi.org/10.1162/rest_a_00868.

Sandra E. Black. Do Better Schools Matter? Parental Valuation of Elementary Education*. *The Quarterly Journal of Economics*, 114(2):577–599, 05 1999. ISSN 0033-5533. doi: 10.1162/003355399556070. URL https://dx.doi.org/10.1162/003355399556070.

Sandra E. Black and Stephen Machin. Chapter 10 - housing valuations of school performance. In Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, editors, *Handbook of the Economics of Education*, volume 3, pages 485 – 519. Elsevier, 2011. doi: https://doi.org/10.1016/B978-0-444-53429-3.00010-7. URL http://www.sciencedirect.com/science/article/pii/B9780444534293000107.

Lex Borghans, Bart H. H. Golsteyn, and Ulf Zölitz. Parental preferences for primary school characteristics. *The B.E. Journal of Economic Analysis & Policy*, 15, 01 2015. doi: 10.1515/bejeap-2014-0032.

Simon Burgess, Ellen Greaves, Anna Vignoles, and Deborah Wilson. What parents want: School preferences and school choice. *The Economic Journal*, 125(587):1262–1289, 2015. doi: 10.1111/ecoj.12153. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ecoj.12153.

Simon Burgess, Ellen Greaves, and Anna Vignoles. School places: A fair choice. *Sutton Trust Report*, 2020.

Simon Burgess, Estelle Cantillon, Mariagrazia Cavallo, Ellen Greaves, and Min Zhang. School admissions in england: The rules schools choose on which pupils to admit, 2023.

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

Jimmy Chan, Xian Fang, Zhi Wang, Xianhua Zai, and Qinghua Zhang. Valuing primary schools in urban china. *Journal of Urban Economics*, 115:103183, 2020. ISSN 0094-1190. doi: https://doi.org/10.1016/j.jue.2019.103183. URL https://www.sciencedirect.com/science/article/pii/S009411901930052X. Cities in China.

Ka Shing Cheung, Chung Yim Yiu, and Yuyu Zhang. What matters more, school choices or neighbourhoods? evidence from a socioeconomic based school zoning. *Cities*, 128:103772, 2022. ISSN 0264-2751. doi: https://doi.org/10.1016/j.cities.2022.103772. URL https://www.sciencedirect.com/science/article/pii/S0264275122002116.

Il Hwan Chung. School choice, housing prices, and residential sorting: Empirical evidence from inter- and intra-district choice. *Regional Science and Urban Economics*, 52:39–49, 2015. ISSN 0166-0462. doi: https://doi.org/10.1016/j.regsciurbeco.2015.01.004. URL https://www.sciencedirect.com/science/article/pii/S0166046215000058.

Ian Davidoff and Andrew Leigh. How much do public schools really cost? estimating the relationship between house prices and school quality*. *Economic Record*, 84(265):193–206, 2008. doi: https://doi.org/10.1111/j.1475-4932.2008.00462.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-4932.2008.00462.x.

Patrick Denice and Betheny Gross. Choice, preferences, and constraints: Evidence from public school applications in denver. *Sociology of Education*, 89(4):300–320, 2016. doi: 10.1177/0038040716664395. URL https://doi.org/10.1177/0038040716664395.

Gabrielle Fack and Julien Grenet. When do better schools raise housing prices? evidence from paris public and private schools. *Journal of Public Economics*, 94(1):59 – 77, 2010. ISSN 0047-2727. doi: https://doi.org/10.1016/j.jpubeco.2009.10.009. URL http://www.sciencedirect.com/science/article/pii/S0047272709001388.

Gabrielle Fack, Julien Grenet, and Yinghua He. Beyond truth-telling: Preference estimation with centralized school choice and college admissions. *American Economic Review*, 109(4):1486–1529, April 2019. doi: 10.1257/aer.20151422. URL https://www.aeaweb.org/articles?id=10.1257/aer.20151422.

Stephen Gibbons and Stephen Machin. Paying for primary schools: Admission constraints, school popularity or congestion?*. *The Economic Journal*, 116(510):C77–C92, 2006. doi: https://doi.org/10.1111/j.1468-0297.2006.01077.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0297.2006.01077.x.

Stephen Gibbons, Stephen Machin, and Olmo Silva. Valuing school quality using boundary discontinuities. *Journal of Urban Economics*, 75:15 – 28, 2013. ISSN 0094-1190. doi: https://doi.org/10.1016/j.jue.2012.11.001. URL http://www.sciencedirect.com/science/article/pii/S0094119012000769.

Steve Gibbons and Stephen Machin. Valuing english primary schools. *Journal of Urban Economics*, 53(2):197–219, 2003. ISSN 0094-1190. doi: https://doi.org/10.1016/S0094-1190(02)00516-8. URL https://www.sciencedirect.com/science/article/pii/S0094119002005168.

Steven Glazerman and Dallas Dotter. Market signals: Evidence on the determinants and consequences of school choice from a citywide lottery. *Educational Evaluation and Policy Analysis*, 39(4):593–619, 2017. doi: 10.3102/0162373717702964. URL https://doi.org/10.3102/0162373717702964.

Ellen Greaves and Hélène Turon. School choice and neighbourhood sorting. *mimeo*, 2023.

Chris Hamnett and Tim Butler. Distance, education and inequality. *Comparative Education*, 49(3):317–330, 2013. doi: 10.1080/03050068.2013.807165. URL https://doi.org/10.1080/03050068.2013.807165.

Jun Han, Liyuan Cui, and Huayi Yu. Pricing the value of the chance to gain admission to an elite senior high school in beijing: The effect of the ldhse policy on resale housing prices. *Cities*, 115: 103238, 2021. ISSN 0264-2751. doi: https://doi.org/10.1016/j.cities.2021.103238. URL https://www.sciencedirect.com/science/article/pii/S0264275121001384.

Oskari Harjunen, Mika Kortelainen, and Tuukka Saarimaa. Best Education Money Can Buy? Capitalization of School Quality in Finland. *CESifo Economic Studies*, 64(2):150–175, 01 2018. ISSN 1610-241X. doi: 10.1093/cesifo/ifx025. URL https://doi.org/10.1093/cesifo/ifx025.

Douglas N Harris and Matthew F Larsen. The identification of schooling preferences: Methods and evidence from post-katrina new orleans. Technical report, Technical report, Education Research Alliance for New Orleans, URL https . . . , 2019.

Justine Hastings, Thomas J Kane, and Douglas O Staiger. Heterogeneous preferences and the efficacy of public school choice. *NBER Working Paper 2145*, 2009.

Thomas J. Kane, Stephanie K. Riegg, and Douglas O. Staiger. School Quality, Neighborhoods, and Housing Prices. *American Law and Economics Review*, 8(2):183–212, 08 2006. ISSN 1465-7252. doi: 10.1093/aler/ahl007. URL https://doi.org/10.1093/aler/ahl007.

Vincent La. Capitalization of school quality into housing prices: Evidence from boston public school district walk zones. *Economics Letters*, 134:102–106, 2015. ISSN 0165-1765. doi: https://doi.org/10.1016/j.econlet.2015.07.001. URL https://www.sciencedirect.com/science/article/pii/S0165176515002748.

Dennis Leech and Erick Campos. Is comprehensive education really free?: a case-study of the effects of secondary school admissions policies on house prices in one local area. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 166(1):135–154, 2003. doi: https://doi.org/10.1111/1467-985X.00263. URL https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-985X.00263.

Marco Oberti and Yannick Savina. Urban and school segregation in paris: The complexity of contextual effects on school achievement: The case of middle schools in the paris metropolitan area. *Urban Studies*, 56(15):3117–3142, 2019. doi: 10.1177/0042098018811733. URL https://doi.org/10.1177/0042098018811733.

John Ries and Tsur Somerville. School Quality and Residential Property Values: Evidence from Vancouver Rezoning. *The Review of Economics and Statistics*, 92(4):928–944, November 2010. URL https://ideas.repec.org/a/tpr/restat/v92y2010i4p928-944.html.

Nienke Ruijs and Hessel Oosterbeek. School choice in amsterdam: Which schools are chosen when school choice is free? *Education Finance and Policy*, 14(1):1–30, 2019. doi: 10.1162/edfp_a_00237. URL https://doi.org/10.1162/edfp_a_00237.

Ian Walker and Matthew Weldon. School choice, admission, and equity of access: Comparing the relative access to good schools in england. Technical report, Lancaster University, Department of Economics, 2020.

A Appendix

A.1 Data construction

Our initial sample of school zones is 957 from 50 Local Authorities.[19]

We create pairs of neighboring school zones within Local Authorities that share a boundary (within a tolerance of 10m). Boundary indicators are assigned to each pair of school zones. Our initial sample of school pairs within Local Authorities is 2164. This decreases to 1603 when we drop boundaries where there is considerable overlap.[20] Appendix figure B.5 illustrates the construction of neighboring school pairs, with an example taken from one LA.

We join school quality to each school with a school zone, which we use to classify the treatment (or high-performing) side of the boundary. This treatment definition will vary depending on the school quality measure j used, as described in section 4. Where the measure of school quality is equal for pairs of neighboring schools, no treatment status is assigned. This is most likely to occur for the Ofsted rating, a discrete measure of school quality on a four-point scale.

To isolate the effect of a 'high quality' school zone over and above correlated neighbourhood characteristics, we focus on zip codes close to each boundary. We calculate whether zip code centroids fall within the specified bandwidths from the border. Our preferred specification uses a bandwidth of 320m, which is consistent with the modal use of 0.2 miles in the previous literature.[21] Appendix figure B.6 shows the postcode centroids that fall within a bandwidth of 200m and 800m from one school zone boundary within this LA.

Sold properties are geolocated to school zones using Geographic Information System (GIS) analysis through the software R, using the centroid of each property's zip code (postcode). As some boundaries have no sales nearby (between 2019 and 2021) we lose 99 boundaries and 9 schools from the sample. Our final sample is, therefore, 1531 boundaries from 858 schools.

A.2 Additional tables

[19] We drop 12 school zones at this stage as they have no neighboring school zone.

[20] Overlap is defined as 'considerable' when the share of the common area (contained in both schools' school zones) is greater than 0.1% of the smallest school zone. This process also leads to the sample of schools decreasing to 867.

[21] We also test the sensitivity of our results to other bandwidths: 100m, 200m, 500m and 800m and optimally chosen bandwidth based on Calonico et al. (2014), CCT hereafter.

Table A.1: Residents' composition from 2011 census, in sample and out of sample

	In sample	Not in sample	Difference	P-value
Panel A: accommodation type				
% Flat	0.123	0.168	-0.046	0
	(0.038)	(0.061)		
% House	0.872	0.826	0.046	0
	(0.04)	(0.064)		
% Mobile	0.003	0.003	0	0.397
	(0.001)	(0.001)		
% Shared	0.002	0.002	0	0
	(0)	(0)		
Panel B: bedroom occupancy				
Rating 0	0.242	0.26	-0.018	0
	(0.02)	(0.022)		
Rating -1	0.034	0.042	-0.008	0
	(0.001)	(0.002)		
Rating -2	0.006	0.008	-0.002	0
	(0)	(0)		
Rating +1	0.342	0.34	0.002	0
	(0.009)	(0.009)		
Rating +2	0.376	0.349	0.026	0
	(0.032)	(0.033)		
Panel C: ethnicity				
% Asian	0.076	0.076	0	0.898
	(0.023)	(0.018)		
% Black	0.02	0.037	-0.016	0
	(0.002)	(0.006)		
% Mixed	0.02	0.022	-0.003	0
	(0)	(0.001)		
% Other	0.007	0.011	-0.003	0
	(0)	(0.001)		
% White	0.876	0.854	0.023	0
	(0.035)	(0.041)		
Panel D: tenure				
Ownership (outright)	0.317	0.304	0.013	0
	(0.026)	(0.028)		
Ownership (tot)	0.661	0.635	0.026	0
	(0.055)	(0.061)		
Rent (private)	0.165	0.169	-0.004	0
	(0.022)	(0.021)		
Rent (social)	0.153	0.174	-0.021	0
	(0.038)	(0.044)		

The table shows the share of residents within and outside our sample for each of accommodation type, bedroom occupancy rate, ethnicity and tenures. Numbers are aggregated from the 2011 census output area figures. Each output area contains, on average, 11 postcodes.

Table A.2: Residents' composition from 2021 census, in sample and out of sample

	In sample	Not in sample	Difference	P-value
Panel A: accommodation type				
% Flat	0.124	0.162	-0.038	0
	(0.039)	(0.059)		
% House	0.832	0.784	0.048	0
	(0.056)	(0.079)		
% Mobile	0.004	0.004	0	0.032
	(0.001)	(0.001)		
% Shared	0.025	0.034	-0.009	0
	(0.005)	(0.007)		
Panel B: bedroom occupancy				
Rating 0	0.245	0.263	-0.019	0
	(0.023)	(0.025)		
Rating -1	0.032	0.039	-0.007	0
	(0.002)	(0.002)		
Rating -2	0.007	0.009	-0.002	0
	(0)	(0)		
Rating +1	0.329	0.328	0.001	0.101
	(0.009)	(0.009)		
Rating +2	0.387	0.361	0.026	0
	(0.036)	(0.038)		
Panel C: ethnicity				
% Asian	0.094	0.091	0.003	0
	(0.028)	(0.021)		
% Black	0.027	0.043	-0.016	0
	(0.002)	(0.006)		
% Mixed	0.026	0.029	-0.003	0
	(0)	(0.001)		
% Other	0.016	0.022	-0.006	0
	(0.001)	(0.001)		
% White	0.837	0.814	0.022	0
	(0.043)	(0.048)		
Panel D: tenure				
Ownership (outright)	0.341	0.323	0.018	0
	(0.029)	(0.031)		
Ownership (tot)	0.64	0.614	0.025	0
	(0.056)	(0.06)		
Rent (private)	0.199	0.202	-0.004	0
	(0.026)	(0.025)		
Rent (social)	0.151	0.172	-0.021	0
	(0.036)	(0.042)		

The table shows the share of residents within and outside our sample for each of accommodation type, bedroom occupancy rate, ethnicity and tenures. Numbers are aggregated from the 2021 census output area figures. Each output area contains, on average, 11 postcodes.

Table A.3: Residents' composition from 2021 census, in sample and out of sample

	In sample	Not in sample	Difference	P-value
Panel A: distance to work				
10 to 20 km	0.103	0.106	-0.003	0
	(0.004)	(0.003)		
20 to 30 km	0.04	0.041	0	0.181
	(0.001)	(0.001)		
2 to 5 km	0.135	0.126	0.009	0
	(0.007)	(0.006)		
30 to 40 km	0.018	0.017	0.001	0
	(0)	(0)		
40 to 60 km	0.014	0.013	0.001	0
	(0)	(0)		
5 to 10 km	0.117	0.119	-0.003	0
	(0.004)	(0.004)		
Work from home	0.297	0.304	-0.007	0
	(0.017)	(0.02)		
$<$ 2km	0.116	0.112	0.004	0
	(0.007)	(0.007)		
$>$ 60km	0.016	0.014	0.001	0
	(0)	(0)		
Other	0.143	0.147	-0.004	0
	(0.003)	(0.003)		
Panel B: economic activity				
% Employed	0.557	0.557	0	0.624
	(0.014)	(0.014)		
% Student (inc. schoolchildren)	0.209	0.205	0.005	0
	(0.013)	(0.01)		
% Active students	0.024	0.023	0.002	0
	(0.001)	(0.001)		
% Inactive student	0.06	0.056	0.003	0
	(0.009)	(0.006)		
% Unemp.	0.026	0.029	-0.003	0
	(0)	(0)		
Panel C: highest qualification				
Apprenticeship	0.054	0.053	0.002	0
	(0.001)	(0.001)		
Level 1	0.099	0.097	0.002	0
	(0.001)	(0.001)		
Level 2	0.136	0.134	0.002	0
	(0.001)	(0.002)		
Level 3	0.176	0.168	0.008	0
	(0.006)	(0.005)		
Level 4	0.329	0.338	-0.009	0
	(0.018)	(0.02)		
None	0.179	0.183	-0.004	0
	(0.008)	(0.008)		
Other	0.027	0.028	0	0.002
	(0)	(0)		
Panel D: migration				
% Migrant UK	0.098	0.096	0.002	0
	(0.006)	(0.006)		
% Migrant non-UK	0.01	0.009	0	0
	(0)	(0)		
% Stayers	0.885	0.889	-0.004	0
	(0.012)	(0.01)		
Panel E: number of rooms				
1 room	0.009	0.01	-0.001	0
Panel F: Ns-SeC				
	(0.001)	(0.001)		
2 rooms	0.082	0.092	-0.01	0
	(0.013)	(0.015)		
3 rooms	0.206	0.224	-0.017	0
	(0.024)	(0.026)		
4 rooms	0.289	0.279	0.01	0
	(0.03)	(0.029)		
5 rooms	0.246	0.24	0.006	0
	(0.031)	(0.031)		
6 rooms	0.096	0.09	0.007	0
	(0.009)	(0.008)		
7 rooms	0.042	0.039	0.004	0
	(0.003)	(0.003)		
8 rooms	0.017	0.016	0.001	0
	(0.001)	(0.001)		
9+ rooms	0.012	0.012	0.001	0
	(0.001)	(0.001)		
High managerial	0.13	0.13	0	0.305
	(0.006)	(0.007)		
Intermediate occ.	0.114	0.114	0	0.127
	(0.001)	(0.001)		
Low managerial	0.197	0.198	-0.001	0.001
	(0.005)	(0.004)		
Low technical	0.055	0.054	0.001	0
	(0.001)	(0.001)		
Routine	0.125	0.121	0.004	0
	(0.005)	(0.004)		
Semi-routine	0.114	0.114	0	0.123
	(0.002)	(0.002)		
Small independent worker	0.104	0.106	-0.002	0
	(0.002)	(0.002)		
Student	0.082	0.076	0.006	0
	(0.014)	(0.009)		
Unemployed	0.079	0.087	-0.008	0
	(0.003)	(0.003)		
Panel G: demographics				
% Male	0.491	0.489	0.002	0
	(0.001)	(0.001)		
Mean age	40.781	40.6	0.181	0
	(51.19)	(49.946)		

The table shows the share of residents within and outside our sample for a set of characteristics. Numbers are aggregated from the 2021 census output area figures. Each output area contains, on average, 11 postcodes.

Table A.4: Student composition, in sample and out of sample secondary schools

	2016-2017		2017-2018		2018-2019	
	Not in sample	In sample	Not in sample	In sample	Not in sample	In sample
Panel A: gender composition						
Part-time girls	0.001	0	0.001	0	0	0
	(0.003)	(0.001)	(0.004)	(0.001)	(0.003)	(0.001)
Part-time boys	0.001	0	0.001	0	0	0
	(0.003)	(0.001)	(0.004)	(0.001)	(0.003)	(0.001)
Full-time girls	0.497	0.494	0.499	0.494	0.499	0.494
	(0.158)	(0.048)	(0.157)	(0.048)	(0.156)	(0.048)
Full-time boys	0.5	0.504	0.5	0.505	0.5	0.506
	(0.158)	(0.048)	(0.157)	(0.048)	(0.156)	(0.048)
Girls	0.499	0.495	0.499	0.495	0.5	0.494
	(0.158)	(0.048)	(0.157)	(0.048)	(0.156)	(0.048)
Boys	0.501	0.505	0.501	0.505	0.5	0.506
	(0.158)	(0.048)	(0.157)	(0.048)	(0.156)	(0.048)
Boarding girls	0.001	0	0.001	0	0.001	0
	(0.01)	(0.002)	(0.01)	(0.002)	(0.01)	(0.002)
Boarding boys	0.001	0	0.001	0	0.001	0
	(0.015)	(0.002)	(0.015)	(0.002)	(0.015)	(0.002)
Panel B: free school meals						
FSM uptake	0.108	0.083	0.102	0.083	0.116	0.096
	(0.07)	(0.031)	(0.066)	(0.031)	(0.07)	(0.034)
FSM eligibility	0.135	0.108	0.131	0.107	0.137	0.119
	(0.084)	(0.038)	(0.08)	(0.038)	(0.082)	(0.041)
Panel C: ethnicity and language						
White	0.742	0.864	0.735	0.862	0.723	0.856
	(0.23)	(0.097)	(0.23)	(0.098)	(0.229)	(0.102)
Black	0.029	0.018	0.031	0.02	0.097	0.04
	(0.025)	(0.009)	(0.025)	(0.009)	(0.104)	(0.025)
Asian	0.036	0.025	0.038	0.026	0.04	0.027
	(0.033)	(0.012)	(0.034)	(0.014)	(0.034)	(0.014)
English language	0.81	0.895	0.809	0.894	0.802	0.89
	(0.184)	(0.079)	(0.181)	(0.081)	(0.179)	(0.083)
Panel D: total student share						
	0.746	0.254	0.733	0.267	0.717	0.283

The table shows the composition of pupils for schools in and out of our sample. Numbers refer to averages of within-school shares of the total student body, while panel D displays the share of students in sample and not in sample across all schools.

The table shows the average measures of school quality for secondary schools in sample and out of sample.

Table A.6: Summary statistics for the price paid for properties, by bandwidth around (distance to) school zone boundary

Bandwidth	N. obs.	N. sales		Price(£)	log(price)	Dist. to school (m)
100	12518	11165	mean	248276.1	12.27	2422.36
			sd	162591.6	0.54	1635.95
200	29546	25659	mean	249233.6	12.27	2371.99
			sd	162219.1	0.54	1639.00
500	85055	66280	mean	253369.8	12.28	2290.97
			sd	173145.6	0.55	1651.05
800	147055	101149	mean	254922.1	12.29	2224.60
			sd	172677.7	0.55	1673.56
1000	191796	122078	mean	255766.7	12.29	2179.13
			sd	172663.6	0.55	1670.82

Table A.7: Summary statistics for the price paid for properties, number of transactions per border, and 'neighbors' per school

Variable	year	mean	median	sd	nobs
Price paid	2019	329940.62	210500	1671201.77	69702
	2020	316933.5	224875	1126555.11	58760
	2021	350549.5	240000	1449432.25	83362
N. transactions per border	2019	158.89	139	107.13	69702
	2020	133.32	115	94.45	58760
	2021	190.63	165	131.65	83362
Neighbors per school		4.54	4	1.83	
Schools per LA		27.5	24	15.34	

Table A.8: Summary statistics for the difference in school quality across boundaries, for three alternative measures of school quality

School quality measure	Quartile	Mean	s.d.	Min	p25	Median	p75	Max
Student attainment	Q1	0.09	0.05	0.00	0.05	0.09	0.13	0.17
	Q2	0.28	0.06	0.17	0.22	0.28	0.33	0.39
	Q3	0.53	0.09	0.39	0.46	0.52	0.62	0.69
	Q4	0.99	0.28	0.69	0.79	0.90	1.13	2.57
Student value added	Q1	0.08	0.05	0.00	0.04	0.08	0.12	0.16
	Q2	0.25	0.05	0.16	0.20	0.25	0.29	0.35
	Q3	0.48	0.08	0.35	0.41	0.47	0.54	0.62
	Q4	0.88	0.23	0.62	0.70	0.83	0.99	1.85
Students eligible for FSM	Q1	0.07	0.04	0.00	0.03	0.07	0.12	0.15
	Q2	0.25	0.06	0.15	0.20	0.25	0.30	0.36
	Q3	0.51	0.10	0.36	0.43	0.50	0.60	0.71
	Q4	1.13	0.35	0.71	0.84	1.03	1.34	2.43

Table A.9: Estimation results for the relationship between the difference in local school quality and property prices

	Attainment (1)	Ofsted val. (2)	SES (3)	Value added (4)
τ^{att}	-0.0034			
	(0.0089)			
τ^{oval}		0.0165		
		(0.0124)		
τ^{ses}			0.0012	
			(0.0082)	
τ^{vadd}				0.0036
				(0.0087)
EPC covs	Y	Y	Y	Y
Bandwidth (m):	100	100	100	100
Observations	20,143	9,723	19,467	20,118
R^2	0.79436	0.78631	0.79162	0.79520
Within R^2	0.58693	0.56685	0.58347	0.58566
Border fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓

Notes: Signif. codes: ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered at the border level. EPC covariates include floor area, property type, construction age band and distance from assigned school. All specification control linearly for distance to the boundary.

Table A.10: Estimation results for the relationship between the difference in local school quality and property prices

	Attainment (1)	Ofsted val. (2)	SES (3)	Value added (4)
τ^{att}	0.0155**			
	(0.0078)			
τ^{oval}		0.0115		
		(0.0096)		
τ^{ses}			0.0103	
			(0.0073)	
τ^{vadd}				0.0153**
				(0.0069)
EPC covs	Y	Y	Y	Y
Bandwidth (m):	200	200	200	200
Observations	46,172	23,012	44,668	46,147
R^2	0.78407	0.79163	0.78390	0.78504
Within R^2	0.58731	0.59412	0.58516	0.58602
Border fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓

Notes: Signif. codes: ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered at the border level. EPC covariates include floor area, property type, construction age band and distance from assigned school. All specification control linearly for distance to the boundary.

Table A.11: Estimation results for the relationship between the difference in local school quality and property prices

	Attainment (1)	Ofsted val. (2)	SES (3)	Value added (4)
τ^{att}	0.0248***			
	(0.0068)			
τ^{oval}		0.0042		
		(0.0092)		
τ^{ses}			0.0231***	
			(0.0071)	
τ^{vadd}				0.0136**
				(0.0064)
EPC covs	Y	Y	Y	Y
Bandwidth (m):	500	500	500	500
Observations	128,732	65,906	124,679	128,881
R^2	0.79022	0.78889	0.79049	0.79100
Within R^2	0.60880	0.60231	0.60690	0.60765
Border fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓

Notes: Signif. codes: ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered at the border level. EPC covariates include floor area, property type, construction age band and distance from assigned school. All specification control linearly for distance to the boundary.

Table A.12: Estimation results for the relationship between the difference in local school quality and property prices

	Attainment (1)	Ofsted val. (2)	SES (3)	Value added (4)
τ^{att}	0.0210***			
	(0.0068)			
τ^{oval}		0.0004		
		(0.0095)		
τ^{ses}			0.0197***	
			(0.0068)	
τ^{vadd}				0.0149**
				(0.0064)
EPC covs	Y	Y	Y	Y
Bandwidth (m):	Optimal CCT	Optimal CCT	Optimal CCT	Optimal CCT
Observations	153,513	80,336	149,999	152,667
R^2	0.81593	0.82658	0.81654	0.81598
Within R^2	0.64932	0.65840	0.64866	0.64904
Border fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓

Notes: Signif. codes: ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered at the border level. EPC covariates include floor area, property type, construction age band and distance from assigned school. All specification control linearly for distance to the boundary.

Table A.13: Property price premium with 320m bandwidth, treatment defined by attainment, with and without covariates

	log(price_paid)		
	No covs (1)	Exog. covs (2)	Full covs (3)
τ^{att}	0.0440***	0.0276***	0.0266***
	(0.0096)	(0.0057)	(0.0052)
Bandwidth (m):	320	320	320
Observations	207,629	102,719	86,267
R^2	0.29253	0.78824	0.82265
Within R^2	0.00107	0.60422	0.65889
Border fixed effects	✓	✓	✓
Year fixed effects	✓	✓	✓

Sources: Energy Performance Certificates, ONS postcodes georeferenced data, Department for Education. Notes: Exogenous covariates include property type, distance from assigned school, construction age band and total floor area.
The full set of covariates further includes a more detailed property type, energy ratings, an indicator for mains gas, the proportion of windows that are glazed, the number of habitable rooms and the number of open fireplaces.

Table A.14: Property price premium with 320m bandwidth, treatment defined by Ofsted, with and without covariates

	log(price_paid)		
	No covs	Exog. covs	Full covs
	(1)	(2)	(3)
τ^{oval}	0.0261**	0.0076	0.0136*
	(0.0126)	(0.0079)	(0.0075)
Bandwidth (m):	320	320	320
Observations	106,738	51,680	42,956
R^2	0.30138	0.78591	0.82131
Within R^2	0.00036	0.59645	0.65095
Border fixed effects	✓	✓	✓
Year fixed effects	✓	✓	✓

Sources: Energy Performance Certificates, ONS postcodes georeferenced data, Department for Education. Notes: Exogenous covariates include property type, distance from assigned school, construction age band and total floor area.
The full set of covariates further includes a more detailed property type, energy ratings, an indicator for mains gas, the proportion of windows that are glazed, the number of habitable rooms and the number of open fireplaces.

Table A.15: Property price premium with 320m bandwidth, treatment defined by value-added, with and without covariates

	log(price_paid)		
	No covs	Exog. covs	Full covs
	(1)	(2)	(3)
τ^{vadd}	0.0206**	0.0175***	0.0163***
	(0.0099)	(0.0055)	(0.0051)
Bandwidth (m):	320	320	320
Observations	207,857	102,725	86,236
R^2	0.29152	0.78865	0.82272
Within R^2	0.00023	0.60285	0.65744
Border fixed effects	✓	✓	✓
Year fixed effects	✓	✓	✓

Sources: Energy Performance Certificates, ONS postcodes georeferenced data, Department for Education. Notes: Exogenous covariates include property type, distance from assigned school, construction age band and total floor area.
The full set of covariates further includes a more detailed property type, energy ratings, an indicator for mains gas, the proportion of windows that are glazed, the number of habitable rooms and the number of open fireplaces.

Table A.16: Property price premium with 320m bandwidth, treatment defined by socioeconomic, with and without covariates

	log(price_paid)		
	No covs	Exog. covs	Full covs
	(1)	(2)	(3)
τ^{ses}	0.0466***	0.0263***	0.0280***
	(0.0098)	(0.0057)	(0.0053)
Bandwidth (m):	320	320	320
Observations	202,978	99,649	83,661
R^2	0.29194	0.78760	0.82236
Within R^2	0.00119	0.60139	0.65706
Border fixed effects	✓	✓	✓
Year fixed effects	✓	✓	✓

Sources: Energy Performance Certificates, ONS postcodes georeferenced data, Department for Education. Notes: Exogenous covariates include property type, distance from assigned school, construction age band and total floor area.
The full set of covariates further includes a more detailed property type, energy ratings, an indicator for mains gas, the proportion of windows that are glazed, the number of habitable rooms and the number of open fireplaces.

Table A.17: Wald test for pairwise difference in coefficients, 320m bandwidth

	Attainment	SES
$\tau^j \times \Delta SQ^j = Q4 = \tau^j \times \Delta SQ^j = Q1$	0.039**	0.046**
	(0.017)	(0.020)
$\tau^j \times \Delta SQ^j = Q4 = \tau^j \times \Delta SQ^j = Q2$	0.021	0.039**
	(0.016)	(0.018)
$\tau^j \times \Delta SQ^j = Q4 = \tau^j \times \Delta SQ^j = Q3$	0.026*	0.048***
	(0.015)	(0.015)
$\tau^j \times \Delta SQ^j = Q3 = \tau^j \times \Delta SQ^j = Q2$	-0.005	-0.009
	(0.016)	(0.016)
$\tau^j \times \Delta SQ^j = Q3 = \tau^j \times \Delta SQ^j = Q1$	0.013	-0.002
	(0.017)	(0.019)
$\tau^j \times \Delta SQ^j = Q2 = \tau^j \times \Delta SQ^j = Q1$	0.018	0.007
	(0.018)	(0.022)

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

Table A.18: Estimation results for the non-linearity of the relationship between school quality differences and property price premiums. Optimally computed bandwidth.

	Attainment (1)	SES (2)
$\tau^j \times \Delta SQ^j$ = Q1	0.0155	0.0137
	(0.0132)	(0.0147)
$\tau^j \times \Delta SQ^j$ = Q2	0.0169	0.0034
	(0.0126)	(0.0123)
$\tau^j \times \Delta SQ^j$ = Q3	0.0059	0.0113
	(0.0121)	(0.0104)
$\tau^j \times \Delta SQ^j$ = Q4	0.0462***	0.0522***
	(0.0108)	(0.0125)
EPC covs.	Y	Y
Bandwidth :	Optimal CCT	Optimal CCT
Observations	153,513	149,999
R^2	0.81606	0.81674
Within R^2	0.64957	0.64904
Border fixed effects	✓	✓
Year fixed effects	✓	✓

Notes: Signif. codes: ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered at the border level. EPC covariates include floor area, property type, construction age band and distance from assigned school. All specification control linearly for distance to the boundary.

Table A.19: Wald test for pairwise difference in coefficients, optimally computed bandwidth

	Attainment	SES
$\tau^j \times \Delta SQ^j = Q4 = \tau^j \times \Delta SQ^j = Q1$	0.031*	0.039**
	(0.016)	(0.019)
$\tau^j \times \Delta SQ^j = Q4 = \tau^j \times \Delta SQ^j = Q2$	0.029*	0.049***
	(0.016)	(0.017)
$\tau^j \times \Delta SQ^j = Q4 = \tau^j \times \Delta SQ^j = Q3$	0.040***	0.041***
	(0.016)	(0.016)
$\tau^j \times \Delta SQ^j = Q3 = \tau^j \times \Delta SQ^j = Q2$	-0.011	0.008
	(0.017)	(0.016)
$\tau^j \times \Delta SQ^j = Q3 = \tau^j \times \Delta SQ^j = Q1$	-0.010	-0.002
	(0.017)	(0.018)
$\tau^j \times \Delta SQ^j = Q2 = \tau^j \times \Delta SQ^j = Q1$	0.001	-0.010
	(0.018)	(0.019)

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

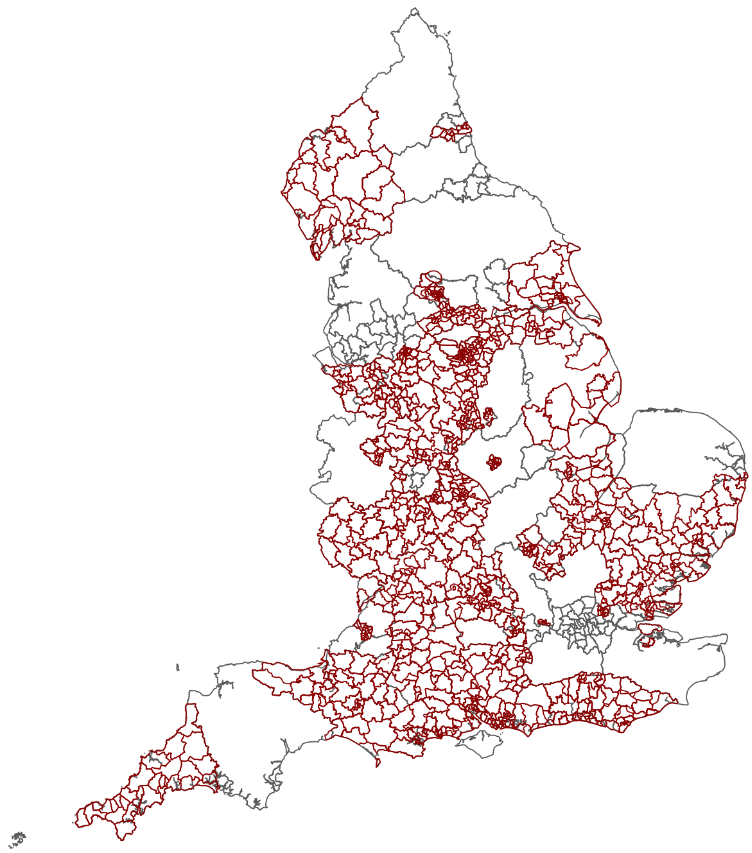
Table A.20: Estimation results for price premiums across boundaries with varying levels of school quality, optimal bandwidth

	Attainment (1)	SES (2)
$\tau^j \times SQ^{comb}$ = Q1-Q1	0.0240	
	(0.0227)	
$\tau^j \times SQ^{comb}$ = Q1-Q2	0.0894***	
	(0.0198)	
$\tau^j \times SQ^{comb}$ = Q1-Q3	0.0443**	0.0555
	(0.0173)	(0.0353)
$\tau^j \times SQ^{comb}$ = Q1-Q4	0.0372	0.1063***
	(0.0270)	(0.0340)
$\tau^j \times SQ^{comb}$ = Q2-Q4	0.0390	0.0895***
	(0.0241)	(0.0276)
$\tau^j \times SQ^{comb}$ = Q3-Q4	0.0158	0.0395*
	(0.0251)	(0.0222)
$\tau^j \times SQ^{comb}$ = Q4-Q4	0.0782***	0.0356*
	(0.0263)	(0.0207)
$\tau^j \times SQ^{comb}$ = Q2-Q3		0.1101***
		(0.0403)
Observations	36,796	35,490
R^2	0.84394	0.79310
Within R^2	0.70313	0.61829
Border fixed effects	✓	✓
Year fixed effects	✓	✓

Notes: Signif. codes: ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered at the border level. EPC covariates include floor area, property type, construction age band and distance from assigned school. All specification control linearly for distance to the boundary.

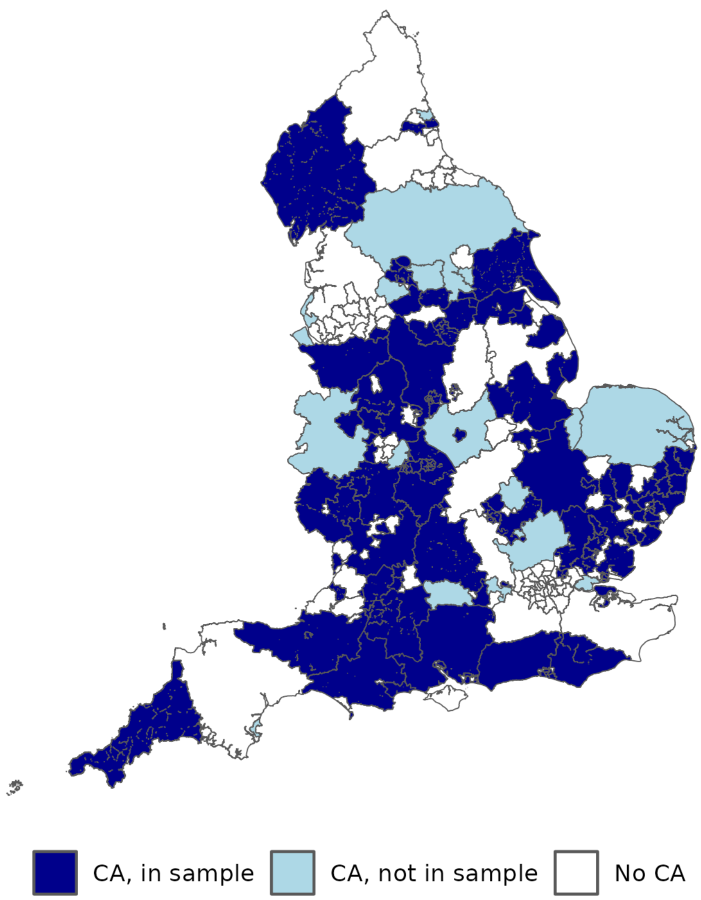
A.3 Additional figures

Figure B.1: Catchment areas used for analysis



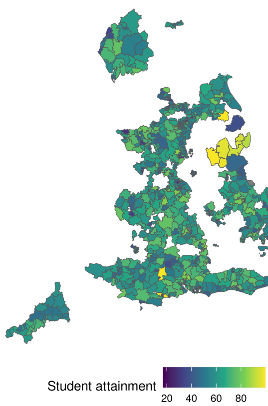
Source: School zones collected by the authors.

Figure B.2: Local Authorities (LAs) in England, by use of school zones

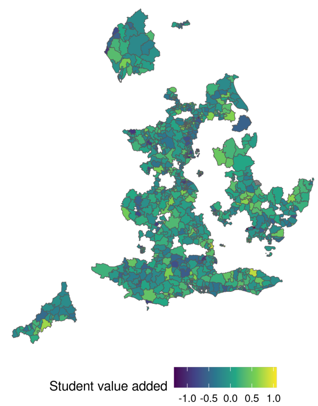


Source: School zones collected by the authors and school admissions criteria collected by Burgess et al. (2023).

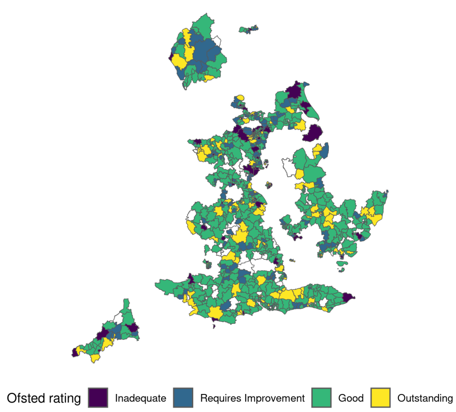
Figure B.3: School 'quality' across school zones in England



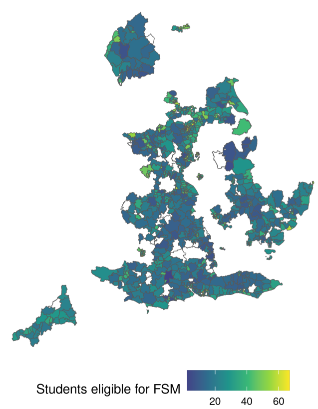
(a) % 9-5 GCSE



(b) Value-added



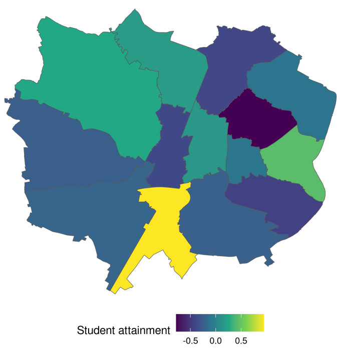
(c) Ofsted



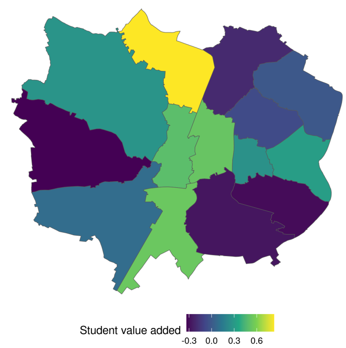
(d) % free school meals

Source: DfE School Performance Tables and school zones collected by the authors.

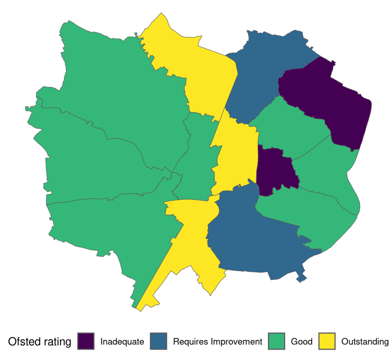
Figure B.4: School ‘quality’ across school zones in one Local Authority



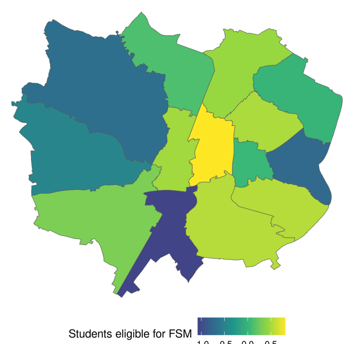
(a) % 9-5 GCSE



(b) Value-added



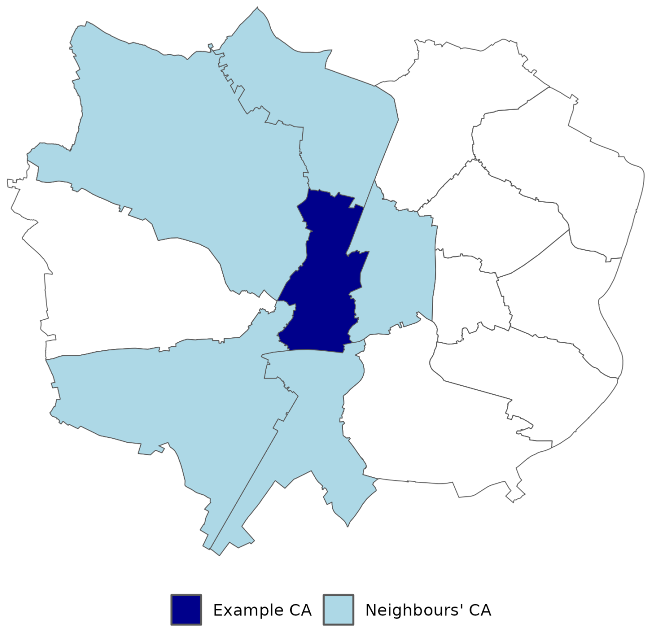
(c) Ofsted



(d) % free school meals

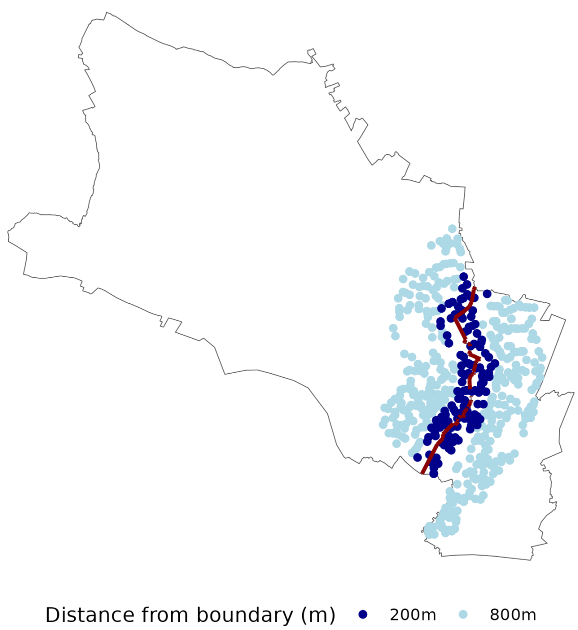
Source: DfE School Performance Tables and school zones collected by the authors.

Figure B.5: Example of construction of neighboring school zones



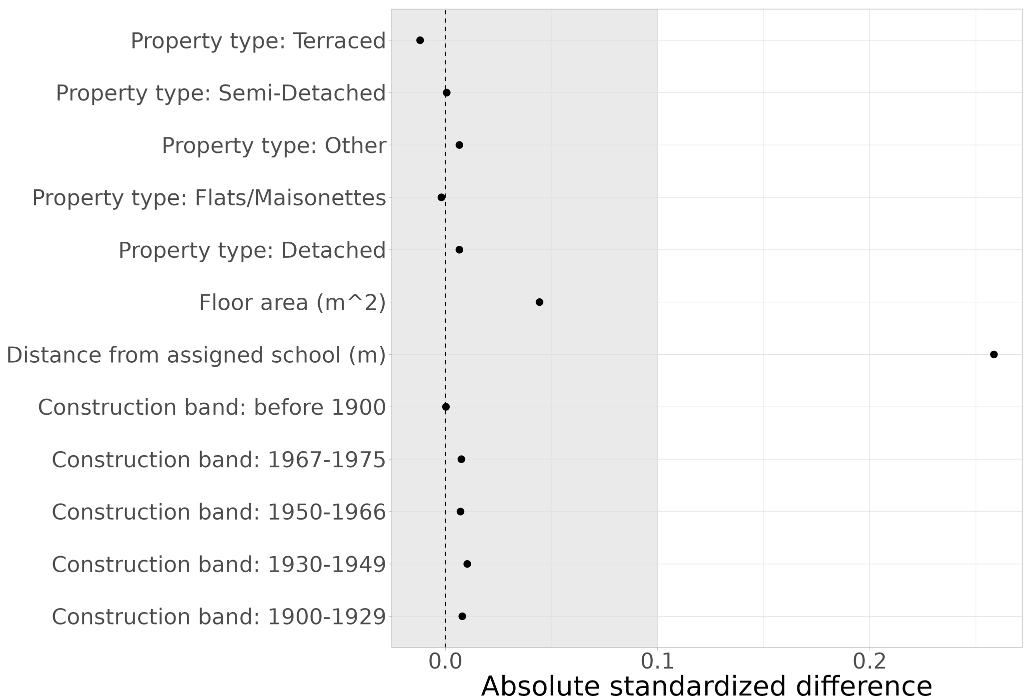
Source: School zones collected by the authors.

Figure B.6: Example of construction of postcodes close to school zone boundaries



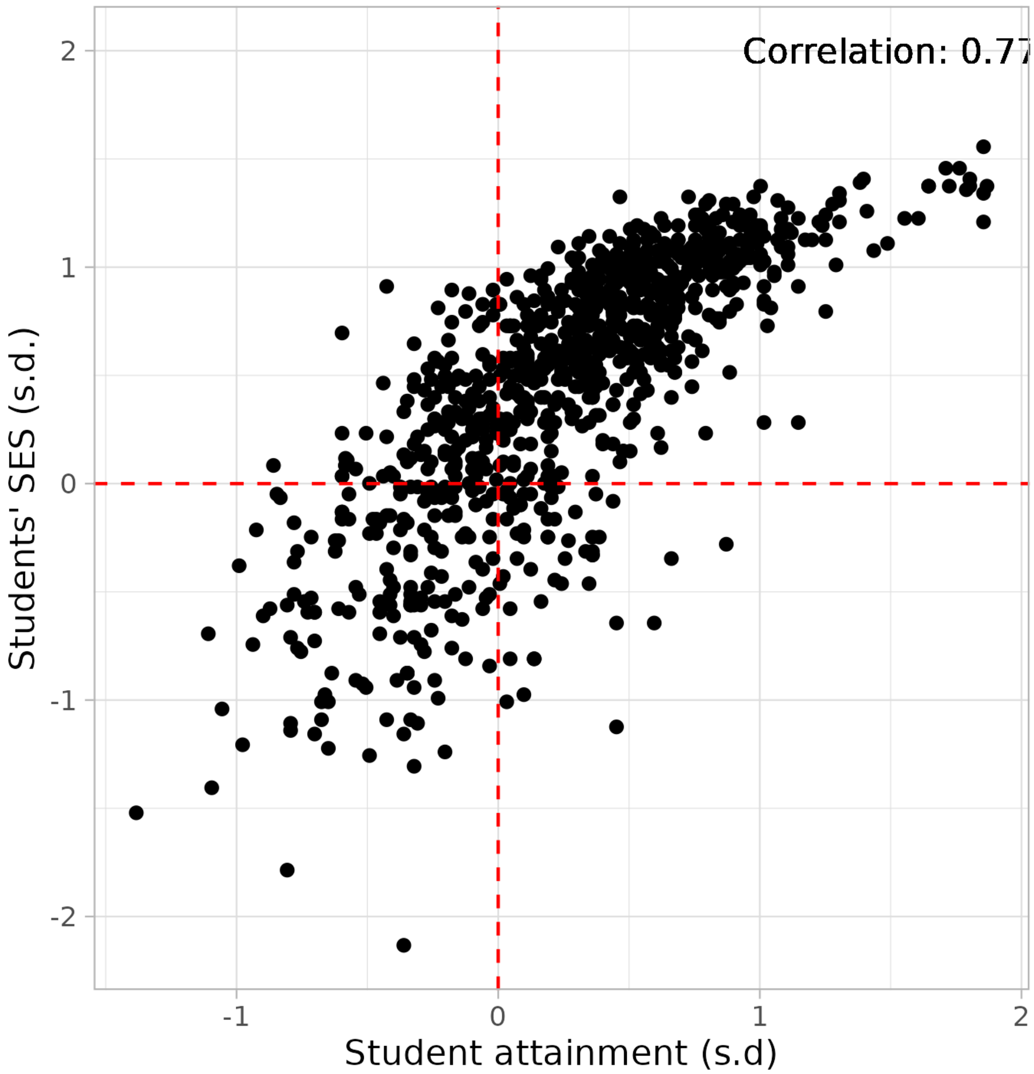
Source: School zones collected by the authors and ONS Postcode Directory.

Figure B.7: Balance of property characteristics across the boundary (absolute standardized difference)



Source: School zones collected by the authors and ONS Postcode Directory, Energy Performance Certificates, Price Paid Data, Department for Education.

Figure B.8: The correlation between school attainment and school composition



Source: DfE School Performance Tables. Notes: This figure shows the correlation between school-level attainment and pupil composition, defined by the percentage of pupils not eligible for free school meals.