Strathclyde
Discussion Papers
in Economics

School Performance, Noncognitive Skills and House Prices

Gennaro Rossi

No. 21 – 2

Department of Economics
University of Strathclyde, Glasgow

School Performance, Noncognitive Skills and House Prices.*

Gennaro Rossi†

March 30, 2021

Abstract

This paper examines parental preference for secondary schools using data from property transactions that occurred between 2015 and 2018 in Scotland, as well as a rich panel of school characteristics. By exploiting discontinuity in attendance across catchment area boundaries, I provide credibly consistent estimates of house price premiums for an array of school characteristics. In particular, whilst I show that house prices do not respond to school value added, school-average performance is well capitalised. I demonstrate that neither of these effects are driven by differences in neighbourhood amenities nor by the presence of private schools. Moreover, I show that in this specific context school "quality" is not multidimensional.

Keywords: School catchment area, Housing market, School performance, Cognitive Skills, Noncogntive Skills

JEL Codes: H75, I21, I28, J24, R21

---

*I am grateful to Markus Gehrsitz and Stuart McIntyre for their help and comments, as well as to Paul Cowell, Alex Dickson, Ian Ditzen, Erkal Ersoy, Gary Koop, Jonathan Norris and Tanya Wilson for their helpful feedback. I thank Daniel Borbely and Orion Meyer for their suggestions and technical support. I am also grateful to Andrew McHugh and Heather Sinclair from the Urban Big Data Centre at the University of Glasgow for the provision and help with the Registers of Scotland data, as well as to Nick Groome at Ordnance Survey for the provision and help with the OS data. Furthermore, I am thankful to Beverley Ferguson, Andrew King, Jill Morton, Stefanie Murr and Gary Sutton at the Scottish Government for the guidance they provided with the education data. I finally thank Sara Jiménez for her help and support. This project uses data from the Registers of Scotland provided through the Urban Big Data Centre at the University of Glasgow

2. Registers of Scotland. Economic and Social Research Council. Registers of Scotland All Sales Data, 2019 [data collection]. University of Glasgow - Urban Big Data Centre.

†Fraser of Allander Institute (FAI), Department of Economics, University of Strathclyde, Glasgow, UK. email: gennaro.rossi@strath.ac.uk

1 Introduction

Parents choose where to live, at least in part, in order to secure a place at the local school for their children. The main argument in favour of school choice relies on its impact on school productivity (Hoxby, 2003) but also on the possibility that it allows students to choose the path that best suits their needs (Hoxby, 2000). However, school choice can be a complex matter (see e.g. Burgess et al. (2015)). Examining these preferences is paramount. Not only does it provide insights on the extent to which parents are knowledgeable about their children's learning experience, it might also shed light on school performance-enhancing policies mechanisms. For instance, if choice is driven by school-level average outcomes, schools might attempt to "improve" student selection rather than effectiveness (MacLeod and Urquiola (2015); Abdulkadiroğlu et al. (2020)). Therefore, this offers clear scope for policy intervention aiming to provide parents with tailored guidance in the choice of schools that best suit their children, as well as to advise schools on how to improve their performance.

The aim of this paper is to assess which school characteristics parents seek the most, with a particular focus on secondary schools, by appraising the capitalisation of school attributes into house prices. To do so, I exploit Scotland's strict residence-based attendance system. I demonstrate that catchment area boundaries separate *de facto* the same neighbourhood, generating a credibly quasi-random source of variation in school quality. Therefore, price divergence across boundaries is causally determined by changes in school quality. I link residential property transactions to local schools and use a rich set of school information provided by the Scottish Government, alongside neighbourhood characteristics as controls. Given the large number of variables at hand, I use factor analysis and hence examine whether school quality is multidimensional.

Long-standing evidence from various contexts suggests that for a one-standard deviation increase in school average test scores, parents pay a house price premium of about 3-4% on average. Research in this area has traditionally proxied school performance using high- and low-stakes test scores see, e.g. Black (1999), Gibbons and Machin (2003), Gibbons and Machin (2006), Bayer et al. (2007), Fack and Grenet (2010), Beuermann et al. (2018); school composition, as a proxy for peer effects (Clapp et al. (2008); Fack and Grenet (2010)), often focusing on one or a few indicators. Whilst this approach is often dictated by data availability constraints, it can be problematic for two reasons. First, one might fail to capture all dimensions of school quality, ignoring school components potentially valuable to parents.

But most importantly, schools with higher scores might just have a more favourable composition. This might not be an issue when average performance is a strong predictor of effectiveness. However, it might well be that schools with an higher average performance do not provide a significant contribution to their students' learning. Another extensive, and in part more recent, strand of the literature therefore focused on measures of school effectiveness, see for example: Brasington and Haurin (2006), Hastings et al. (2009), Gibbons et al. (2013), Imberman and Lovenheim (2016), Beuermann et al. (2018) and Abdulkadiroğlu et al. (2020). One common finding is that parents still choose schools based on their average performance rather than value added (see for example Brasington and Haurin (2006) and Abdulkadiroğlu et al. (2020)). Part of the reason might be a lack of information on school effectiveness (Ainsworth et al., 2020). Imberman and Lovenheim (2016) exploit a set of school value added statistics which were newly promulgated by newspapers. Despite the fact that parents response to newly released school information has been documented to be quick and sizeable (Figlio and Lucas (2004); Fiva and Kirkebøen (2011)), Imberman and Lovenheim (2016) find no effect of value added on house prices. The question therefore remains whether parents value school effectiveness at all.

In this paper, I seek to address these issues and extend the literature in two ways. First, I use a fairly large data set of measures of in– and post–school attainment. The aim is to assess whether school performance varies along different dimensions (Heckman et al. (2006), Kautz et al. (2014), Jackson (2018)) and whether parents are aware of school performance differences (Beuermann et al. (2018)). This paper is one of only a few works examining the role of post-school attainment in school choice. Second, I isolate school-average performance from school effectiveness by means of "the virtual comparator", an index developed by the Scottish Government which compares actual school performance with performance as predicted by the demographic composition of the school. The virtual comparator is publicly available and reports its comparison alongside actual school performance. This means that parents have a good idea of how well a school is performing with respect to a counterfactual.

This is an important difference between my analysis and the existing literature in which school value added is estimated by the researcher (Brasington and Haurin (2006), Beuermann et al. (2018), Abdulkadiroğlu et al. (2020)), and despite the internal validity of the estimates, it is uncertain whether parents are in fact aware of school effectiveness. This paper adds to this strand of the literature in which value added information is widely available to parents (see for example Gibbons et al. (2013) and Imberman and Lovenheim (2016)), and it extends the literature in two significant ways. It does so by using a more contextualised measure of value added which takes into account students' characteristics which are correlates of their progress, and by examining the capitalisation of value added over a longer, post-release, time span.[1]

A further novelty of this paper lies in the type of indicators that I use. Whilst some recent work has explored school outcomes beyond academic scores,[2] parental preferences for noncognitive skills have been mostly overlooked. To the best of my knowledge, this is the first work attempting to identify and quantify parents preferences for noncognitive skills at the school level, in particular in the context of their capitalisation into house prices. Noncognitive skills have been soundly proven to be important determinants of various outcomes.[3]

A very important feature of noncognitive skills inheres to their long-lasting effects when developed early on in life (Chetty et al. (2011), Bono et al. (2016)) but also to their malleability later on (Heckman and Mosso (2014), Norris (2020)). For instance, Norris (2020) finds evidence of peer effects in learning attitudes in adolescence, and that parents' expectations can meliorate peers' influence. By focusing on the behavioural component of noncognitive skills (Lleras (2008), Bertrand and Pan (2013),Jackson (2018)) at the secondary school level, I examine whether the possible influence that peers can exert plays a role in secondary school choice.

[1]Imberman and Lovenheim (2016)'s main focus is on seven months post-release.

[2]Beuermann et al. (2018) look at parental choice in terms of, among other things, school-level incarceration rate, teen births and future employment.

[3]See Kautz et al. (2014) for a review.

Furthermore, whilst a revealed-preferences approach of this sort has the advantage of its simplicity of implementation and enables a price to be placed on certain social dynamics, it conceals the issue that observed transactions might not have occurred for school reasons. I try to address this issue, which has been mostly overlooked in the literature, with a control variables strategy.

Finally, to the best of my knowledge this is to date the only work of this sort in Scotland and one of the few works in this field[4] focusing on a nation-wide context therefore providing a contribution in terms of external validity.

Overall, my findings suggest that high average scores are often not synonymous with value added, and yet parents seem to value the former rather than the latter. Furthermore, whilst there is no evidence of multidimensionality, parents nonetheless seem to seek some outcomes but not others.

The remainder of the paper is structured as follows. In the next section I will describe the Scottish school system and its residence-based attendance system. In section 3 I will describe the data, the method used to build the main sample and the identification strategy, providing parametric evidence of its validity. In section 4 I will present and discuss the results. The final section concludes.

2 Institutional Background

The Scottish education system is predominantly composed of state funded, non-religious (or non-denominational) schools, which constitute the primary focus of this paper. Private (or independent) schools are mostly clustered in the "Central Belt" of the country, and according to the Scottish Council of Independent Schools (SCIS) census,[5] in September 2015 approximately 30,238 pupils were enrolled in private schools, which amounts to 4.3% of the total school age population. For secondary schools, the figure is 18,159, which is roughly 6% of the corresponding sub-population. These schools often follow a curriculum which is different from the national one,

[4]see for example Gibbons and Machin (2003), Gibbons et al. (2013) or Burgess et al. (2019)

[5]`http://www.scis.org.uk/facts-and-figures/`

e.g. International Baccalaureate or A-levels. Religious schools, which account for roughly 15% of the total number of students in the country, are part of the state funded school sphere and as such follow the national curriculum, but have different catchment areas.

Scottish pupils typically access their first year of primary school between their fifth and sixth year of age. Primary school includes seven grades (P1-P7) after which they move to the first of the six grades of secondary school (S1-S6). In 2010 a new national curriculum, the *"Curriculum for Excellence"*, was implemented and it consists of five levels: Early, First and Second levels take place within the seven years of primary school; Third/Fourth, during the first three years of secondary school (S1-S3) and finally the last one, Senior Phase, from S4 to S6. As students turn sixteen, usually in S4, they can either leave to pursue further education or go into employment (or unemployment), or progress to S5 (and perhaps later to S6) in order to attain further qualifications within the Scottish Qualification Authority (SQA), i.e. the national exam board.

Each of these qualifications has an equivalent level in terms of Scottish Credit and Qualification Framework (SCQF). This is a scheme which aims to facilitate credit transfer by organising degrees and qualifications in twelve levels, 12th being the most challenging one and corresponding to a PhD or a professional apprenticeship. Table A1 outlines the possible qualifications that can be attained during "Senior Phase" (S4-S6) alongside their corresponding SCQF levels. In particular, up to S4, students can attain a number of "National 5" qualifications, corresponding to SCQF level 5. For example, if a student passed tests in mathematics and English at the end of S4, she would be awarded two "National 5" qualifications, or similarly, two awards at SCQF level 5. The decision to progress further to S5 is contingent on a pupil's intention to achieve further qualifications. Within the SQA classification, these are typically "Higher" and "Advanced Higher" qualifications, which correspond to SCQF levels 6 and 7 respectively and are roughly the equivalent of the English A-levels.

As part of the admission criteria, Universities in Scotland (and rest of the UK) require a certain number of "Higher" (or above) to be attained with certain grades and within specific subjects, depending on the course and institution the student is applying to. Given this, we can regard “Higher” and “Advanced Higher” as high-stakes qualifications. One important caveat is that Scottish Government reports in its official statistics and data the number of qualifications in terms of SCQF levels, rather than in SQA nomenclature. This is a detail I will come back to later, in section 3. Table 1 shows that between school year 2013/2014 and 2015/2016 roughly 39% of pupils attained four or more SCQF level 6 awards, about 36% of leavers went to higher education whilst 66% obtained at least four awards at SCQF level 5.

Schools with higher performance in high-stakes tests might not necessarily be what parents want for their children, and they might instead opt for schools which are stronger at preparing their children for different paths such as employment after school, or further education via a college or vocational training provider. Hence, SCQF level 5 qualifications might be an interesting outcome to look at as these constitute the highest qualification prior to attaining school-leaving age.

Each of these qualifications can be achieved within a fairly large range of subjects alongside English and Mathematics. In order to frame attainment within literacy and numeracy skills, SCQF levels in literacy and numeracy are awarded to students who have achieved National 5, Higher and above in specific subjects or course/units. For example, a student would achieve SCQF level 6 in Literacy if she obtained a Higher qualification in all five unit groups for English.

With respect to school choice, Local Authorities, roughly equivalent to US counties, are usually divided into catchment areas. For example Figure 1 shows non-denominational secondary schools catchment areas for Edinburgh and Glasgow. Each dwelling is located within four nested school catchment areas: two primary and two secondary schools, one of each being non-denominational. Denominational (mostly Roman-Catholic) school catchment areas are generally larger and stretch across those of non-denominational schools. They do not require pupils to be from a particular denomination or faith but intakes criteria vary by Local Authority. For non-denominational schools, parents are advised to enrol their children in the designated local catchment area school by the Local Authority, which is obliged to enrol kids from the catchment area. However, parents can submit a “placing request” for a school different from the designated one (LearningDirectorate (2016)).

Councils prioritise children living in the catchment area, and do not have to grant a placing request by parents from outside the catchment area. They are more likely to grant such a request where this would not alter the pupils/teacher allocation, e.g. if the school needed to hire an additional teacher or set up a new classroom; or there were reasons to believe the placing request would not constitute a good fit, either for the children already in school or for the pupil whose parents have submitted the request.[6] Successful placing requests are rare and often pertain to children with additional support needs or who have sibling(s) already in the school. Moreover, Borbely et al. (2020) show catchment areas overlapping is rare.

3 Data and Research Design

The housing data used in this project are provided by the Registers of Scotland via the Urban Big Data Centre at University of Glasgow,[7] and contain every residential property transaction that occurred in Scotland from January 2008 to April 2018. The data provide detailed information about the consideration (sales price), address, postcode as well as geographic coordinates for most of the observations. Where these coordinates were missing, I integrated them using the AddressBase® Plus dataset[8] provided by Ordnance Survey. In order to match each transactions to a specific school I employed catchment areas shapefiles covering all of Scotland, provided by the Improvement Service.[9] School information is obtained from a publicly available database of the Learning Directorate of the Scottish Government (SG). In particular, for each school, information was collected on the number of students broken down by stage, percentage of pupils eligible for free school meals, pupil-teacher ratio, information about the capacity of the estates as well as the location of each school.

Attainment data are collected from the interactive dashboards provided by the Learning Direc-

[6] see, for more details, `https://www.gov.scot/publications/choosing-school-guide-parents-nov-16/`

[7] 1. © Crown copyright. Material is reproduced with the permission of the Keeper of the Registers of Scotland. 2. Registers of Scotland. Economic and Social Research Council. Registers of Scotland All Sales Data, 2019 [data collection]. University of Glasgow - Urban Big Data Centre.

[8] Ordnance Survey © Crown Copyright 2019. All rights reserved. Licence number 100034829

[9] © Crown copyright and database rights "2019" Improvement Service (licence number 100050367)

torate. They contain, among others, the following variables at the school level: *i)* The percentage of leavers attaining at least four awards at SCQF level 6 (or above). *ii)* The percentage of leavers attaining at least four awards at SCQF level 5; *iii)* the percentage of pupils on free school meals - this is meant to capture variation in school composition, which for parents might be a proxy for peer effects.[10] Table 1 shows that between school year 2013/2014 and 2015/2016 roughly 85% of pupils were not eligible for free school meals. Regarding point *i)*, the goal here is to use a measure of high-stakes attainment. As Table A1 indicates, the SQA equivalent of a SCQF level 6 award is a "Higher" qualification. A minimum of four "Higher" qualifications is a standard admission criteria applied by many universities. However, observing the number of awards at SCQF level 6, as opposed to the number of "Higher" qualifications achieved, conceals the risk of dealing with a noisy measure of academic-oriented qualifications. In fact, a SCQF level 6 award might correspond to a "Higher" qualification as well as to a Scottish Vocational Qualification (SVQ). Figure A1 shows that the percentage of leavers achieving a minimum of four awards at SCQF level 6 is highly correlated with the percentage of leavers in higher education[11], as well as with the percentage of S6 leavers.[12] This is evidence of the fact that the four-awards threshold for SCQF level 6, besides ensuring sample size gains [13] it captures variation in high-stakes attainment, being this primarily composed of "Higher" qualifications.

As I mentioned previously, a school with many students attaining "Higher" qualification is not necessary a "good" school in the sense of value added, i.e. contribution to their students' learning. This is why I also build an indicator of school 'effectiveness' (or value added) for each of the measures described in *i)* and *ii)* as well as for literacy and numeracy. These indicators consist of the difference between the actual measure of attainment at the school level and its virtual comparator. The latter has been developed by the Learning Directorate at the SG. In order to "adjust" school performance for its composition, a virtual comparator creates for each school a synthetic

[10] Please note that, unlike for some stages in primary school, free school meals in secondary school have not been made universal, therefore this variable should still capture meaningful variation in socio-economic background.

[11] After nine months from secondary school graduation.

[12] Hence those leavers attaining two extra years of schooling.

[13] The corresponding school-level percentage is normally neither too high nor too low and thus its values are rarely suppressed for disclosure check reason.

control. For each pupil in the school in question, ten other pupils are randomly picked from other local authorities in Scotland, conditional on the same demographics for the "treated" pupil such as gender, free school meal eligibility, additional support needs and Scottish Index of Multiple Deprivation (SIMD)[14]. The average attainment of these ten pupils constitutes the counterfactual score. The difference between the actual school performance and its virtual comparator should provide information about how well this school performed, given its composition. I will assess whether these measures are capitalised into house prices.

Furthermore, I will use census[15] and Scottish Index of Multiple Deprivation (SIMD)[16] variables, both at the data zone level, as controls. 6,976 data zones were defined based on the 2011 Census, each containing between 500 and 1000 inhabitants. The SIMD information pertains to the 2016 edition, which uses the data zones as redefined in the 2011 Census. Figure 2, provides an example of how Data Zones nest within secondary school catchment areas. Table 1 reports descriptive statistics for the main variable involved in the analysis, whilst Table B1 provides a brief description of the attainment variables.

3.1 Empirical Strategy

In the absence of survey data in which individuals state their preferences for schools characteristics, or data on applications ranking the desired schools, a viable way to determine parents' preferences for schools is via the capitalisation of their characteristics into house prices. The standard technique used in the literature is a revealed preferences approach in which the following model is estimated:

$$ln(p_{int}) = \beta A_n + \gamma X_{int} + \varepsilon_{int} \tag{1}$$

where the price of the i^{th} sale is a function of a vector of sale-neighbourhood-time specific character-

[14]This is an index developed by the Scottish Government and is meant to classify areas based on their degree of deprivation across seven domains, i.e. income, employment, education, health, housing, crime and access to services.

[15]Scotland Census 2011, © Crown copyright. Data supplied by National Records of Scotland.

[16]© Crown copyright 2016

istics X_{int}, an error term ε_{int} and most importantly of A_n, the amenity of interest in neighbourhood n, with β measuring the marginal willingness to pay for the amenity in question, e.g. a school with higher test scores. Obtaining a consistent and plausibly causal estimate for β is not straightforward. Amenities are not randomly assigned to neighbourhoods and estimating their valuation in a simplistic fashion like the standard hedonic pricing model in Equation 1 can lead to severe biases if confounding factors and sorting (Epple and Romano (2003); Bayer et al. (2007)) are not properly accounted for. For instance, wealthier residents tend to buy more expensive houses and this has two potential consequences: *i)* higher property tax revenue translates into more generous school funding, and potentially higher quality; *ii)* more affluent parents might devote more resources to their children's learning, thus affecting the observed performance of the local institution. Moreover, unobserved heterogeneity at the neighbourhood level in factors which affect property prices and are also correlated with school quality might induce further bias.

I attempt to disentangle these issues, following a similar approach to Black (1999), by estimating the following model:

$$ln(p_{isb,t}) = \beta Q_{s,t-1} + \gamma X_{sb,t} + \phi K_b + \varepsilon_{isb,t} \tag{2}$$

where $p_{isb,t}$ is the price of house sale i, in school catchment area s, in proximity of catchment area boundary b, in year t, whose main predictor of interest is $Q_{s,t-1}$, namely the school quality indicator (as perceived by parents) for school s, (at least) one period before the sale occurs. In particular I regress transaction prices from 2015 to 2018 on school performance from 2014 to 2016. $X_{sb,t}$ contains a set of school- and neighbourhood-specific controls at boundary b, as well as year-of-sale dummies. Finally, K_b is a vector of boundary indicators, and $\varepsilon_{isb,t}$ is the usual error term, which has to be orthogonal to $Q_{s,t-1}$, condition *sine qua non* for $\widehat{\beta}$ to consistently estimate β, the marginal willingness to pay (WTP) for a move closer to a school perceived to be "better". I focus on transactions for properties which are close enough to each other to be part of the same neighbourhood, but on different sides of a catchment area boundary. Thus, any difference in price can be argued to represent differences in school quality.

I visualise the identification strategy taking as an example the Leith Academy catchment area (Figure 3a), and its southern boundary separating it from Portobello High School (Figure 3b). The boundary on that section of Portobello Road separates two sides of the same street into different school catchment areas. Hence, any price differential between either side of the boundary can be attributed to gap in the performance between the two schools. Moreover, the inclusion of boundary fixed effects K_b should remove any unobserved variation at the boundary level. Before looking more closely at the identification strategy, I want to discuss more in depth the school "quality" component.

3.2 School "quality" indicators and multidimensionality

When discussing school quality as perceived by parents, a distinction needs to be made between high-performing schools, defined as those recording high test-scores or share of pupils attaining a certain qualification, and "effective" schools, defined as those schools best able to actively contribute to improve pupils' learning. These two scenarios do not necessarily coincide (Doris et al. (2019)). Second, even when this distinction is clear, schools might differ in terms of a vast array of other characteristics, both within the cognitive and noncognitive spheres. Whilst it might seem obvious that schools that are "good" under a certain profile (either because they have high-performing pupils, or because they are effectively enhancing student performance) should also be so under alternative profiles, there is evidence of school multidimensionality (see e.g., Beuermann et al. (2018)).

In this paper I want to assess whether: *i)* Scottish secondary schools are "multidimensional"; *ii)* parents perceive multidimensionality. Table 2 reports a factor analysis which have been run for the sixteen potential indicators of school "quality", for all 272 schools contained in the full sample. The aim is to identify subsets of variables which might be indicative of some specific school latent characteristic, as expressed by each factor. Factor loadings below .6 have been marked with blanks.

As is evident from the first column, the first factor loads pretty highly on all of the indicators - this is the first evidence of a lack of multidimensionality, in other words, "good" schools do well under most of the profiles. However, factor 1 loads higher on % of SCQF awards at level 5, 6, % of leavers in higher education and attendance rate. This suggests some commonality in variables which might reflect the “academic” orientation of a school - this in part undermines the belief that level 5s could be regarded as low-stakes qualifications, since schools with many pupils attaining Higher qualifications, which in turns are school with large shares of leavers going to university, also have many pupils attaining National 5 qualifications. Factor 2 meanwhile loads heavily on the value added measures, whilst Factor 3 loads highly, though with a different sign, on to authorised and unauthorised absence rate.[17]

Taken at face value, the results in Table 2 suggest that: *i)* school performance might be bi-dimensional, i.e. value added and average outcomes are two separate latent factors; *ii)* average academic outcomes and noncognitive components might be treated as two separate indicators. Based on this, I have built three different performance indexes by running principal component analysis on three subset of variables separately, and extracting the first component from each of these. The relevant variables for each index and their weights are reported in Table 3. The noncognitive skills index, loads negatively on attendance rate and positively for the remaining two variables. This implies that higher values of the index will correspond with higher exclusion and unauthorised absence rates. Hence, the index has been re-coded in a way that higher scores correspond to higher level of noncognitive skills. Please note that authorised absence rate is not part of this index, purely because this variable mostly captures kids missing schools due of illness.[18] Finally, each of these indexes have been standardised to have mean zero and unit-variance.

In summary, the four indicators are: *i)* an Academic Index, which is meant to capture school-level average performance; *ii)* a Value Added Index, as a measure of school effectiveness; *iii)* a Noncognitive Skills Index, as a linear combination of school-level behavioural indicators; *iv)* a Composition measure, which is purely the standardised value of the % of pupils not eligible for free school meals. These four indicators will constitute my main (endogenous) regressors.

[17] Results from a Principal Component Analysis are available on request and suggest essentially the same patterns.

[18] https://www.gov.scot/publications/summary-statistics-schools-scotland-no-10-2019-edition/pages/7/

3.3 Validity

In order to consistently estimate parents willingness to pay (WTP) for a "better" school, β, any neighbourhood characteristic which is plausibly correlated with school attributes and is a plausible predictor of house prices, must not systematically differ across the boundaries. This is the standard validity assumption of a regression discontinuity design (see for example: Lee and Lemieux (2010)). In other words, the side of the boundary which is "treated" with the better-performing school must share the same housing and neighbourhood characteristics as the opposite side of the boundary which is not "treated". Table 4 provides some evidence on this.

Each estimated coefficient originates from a model in which each of the variables presented in Table Table 4 is regressed on a binary indicator for the "good" side of the boundary, intended as the one with a larger value of the "Academic" index. Standard errors are clustered at the secondary school catchment level and reported in parenthesis. In order to explore the robustness of the results to different border distances, columns (1), (2) and (3) focus on houses which are located within 350, 300 and 250 metres of either side of the boundary, in order to create a buffer which can arguably be regarded as the same neighbourhood.[19]

We can see that the "good" side of the boundary is characterised by a performance which is approximately 1σ larger than the opposite side on average. Whilst this discontinuity occurs in a sense by construction, it is still sizeable and a common finding in the literature (see for example, Black (1999), Bayer et al. (2007) and also Gibbons et al. (2013)). However, corresponding to this jump in academic performance, there is no systematic change in a series of neighbourhood characteristics, collected at the data zone level. For example, on the "good" side, houses are not significantly bigger or smaller, [20] nor are they located in more or less densely populated areas - these differences are small and not statistically significant.

Furthermore, there is no evidence supporting the hypothesis of sorting; this is evident in the

[19]Previous literature used similar distances, see e.g. Black (1999), Bayer et al. (2007) and Fack and Grenet (2010).

[20]Unfortunately the data do not allow me to observe the characteristics of individual houses, therefore I use as a proxy the average number of rooms in the neighbourhood.

coefficients on the percentage of people with a higher education degree, without qualification, female-headed households with children. The percentage of income deprived people changes only slightly and is barely statistically significant, alongside the percentage of social housing, which is the only variable displaying a statistically significant difference across catchment area boundaries. However, it is a relatively small change within its own distribution (see table Table A3 for descriptive statistics) and it is unlikely to drive house prices by a significant and large extent. Nevertheless, this will be duly included in the main regressions as a control together with all the variables presented in Table 4. As a further robustness check, in Table A2 I repeat the above exercise when the "good" side is obtained in relation to the other performance indicators and for houses within 350 meters from the boundary - and the results are analogous.

In conclusion, whilst neighbourhoods located on the "better-performing" side of the boundary are characterised by a sharp (1σ) increase in school performance, they do not seem to be systematically different from their nearby counterparts in any dimension which is plausibly correlated with school performance and likely to affect house prices.

4 Results

Having demonstrated in Table 4 that houses in close proximity to catchment areas boundaries are comparable, in Table 5 I present the results from estimating some variations of the main model presented in Equation 2. Each estimated coefficient corresponds to $\widehat{\beta}$, obtained from regressing $ln(price)$ on one performance indicator at the time. Standard errors are clustered at the catchment area level and reported in parenthesis.

I will first focus on the models using the academic performance index as the main regressor. In column (1) I run the relevant model using the full sample, namely without restrictions based on distance from boundary. Results suggest that 1σ increase in academic performance index corresponds to a 3.2% rise in house prices. I then consider, in column (2), houses within 350 metres from the boundary controlling only for neighbourhood characteristics,[21] and $\widehat{\beta}_{Academic}$ shrinks and

[21]These include the covariates from Table 4 plus a set of year-of-transaction dummies and school size.

becomes statistically insignificant. In column (3) I control for unobserved heterogeneity through boundary fixed effects, and the WTP for higher average academic attainment is approximately 3% for 1σ increase in the index.

Columns (4)-(5) reports estimates with a further sample restriction, namely narrowing the focus on houses in urban areas, [22] and finally in column (6) I include, within the 350-metres-sample, houses in proximity of those catchment area boundaries which coincide with Local Authorities' boundaries. Their initial exclusion is justified by the fact that there might be some unobserved variation across local authority which might not be captured by boundary fixed effects.[23] However, the estimated coefficient does not change significantly and it is as precisely estimated as in the previous specifications.

A similar result is obtained with respect to $\widehat{\beta}_{Composition}$, whose magnitude hover around a 4% house price premium for a 1σ increase in the school-percentage of pupils not eligible for free school meals. For my measure of value added and the index of noncognitive skills, the point estimates are considerably smaller and not statistically significant, despite being precisely estimated. All estimates only change marginally when I restrict the sample to 300 and 250 metres from the boundary, as reported in Table A4 and Table A5, and are consistent with the relevant literature. In particular, for 1σ increase in academic performance, these results are in line with the 2.5% reported by Black (1999), 2% by Fack and Grenet (2010), and 3% by Gibbons et al. (2013).

4.1 Robustness

Results so far suggest that parents value academic performance but not effectiveness. Furthermore, I have demonstrated that my identification strategy is clean and provided results which are in line with previous work. Nevertheless, there might still be threats to the validity of the results. In this section I explain these and set out how I explore them in more detail.

First, it is necessary to isolate the effects of school resources from one purely driven by school

[22] "Other" urban areas corresponds to settlements between 10,000 and 124,999 people, whilst "large" urban areas have 125,000 inhabitants or more. These are classifications provided by the Scottish Government.

[23] I do not include Local Authorities fixed effects as these are multicollinear with boundary fixed effects.

outcomes. Second, I address an issue that has been frequently overlooked in this literature, namely that when using a revealed (school) preferences approach of this sort, I heavily rely on the assumption that all the observed housing transactions happened as a result of school choice. However, there might be areas containing only few families, hence local house prices might be unlikely to pick up school quality differentials.

Third, whilst the main focus of this project is on state funded non-denominational schools, the Scottish education system features also a private sector -Independent Schools- as well as state funded denominational schools, the vast majority of which are Roman Catholic (henceforth, RC). Whilst the latter follow the same curriculum as the non-denominational schools, I do not observe Independent schools performance.[24]

Fourth, I want to control for the possible presence of spatial trends in house prices, as well as unobserved heterogeneity across time and space. Fifth and finally, secondary school catchment boundaries can cover a large portion of territory. If houses are on opposite sides of the boundaries but far away from each other, the main identifying assumption might not hold. I therefore implement a matching framework similar to Fack and Grenet (2010):

$$\Delta ln(p_{isb,t}) = \beta\Delta Q_{s,t-1} + \gamma\Delta X_{isb,t} + \phi(K_b^* - K_b) + \Delta\varepsilon_{isb,t} \quad (3)$$

where $\Delta ln(p_{isb,t}) = ln(p_{i^*s^*b,t}) - ln(p_{isb,t})$ and i^* indicates sales on the "good" side of the boundary, whereas i is the counterfactual sale, whose price is constructed as the geometric mean of the closest three sales' prices, using inverse distance between i^* and closest three matches as weights. The identifying assumption is that i^* and i share the same unobservable characteristics, i.e. $K_b^* = K_b$.

Table 6 runs a battery of robustness checks for the academic performance indicator. The first column reports the baseline estimate, these are the same as column (3) in Table 5. Column (2) addresses the first three points outlined at the beginning of subsection 4.1 by controlling for pupil-teacher ratio, number of pupils relative to capacity, a private schools presence index similar to the

[24] Some of which follow the English system or International Baccalaureate.

one used by Fack and Grenet (2010)[25] and finally the percentage of households with dependent children.

Columns (3)-(4) addresses whether the performance of the local RC school, proxied by its score in Higher qualification exams, explains any variation in price. One thing to notice is that RC schools are not present in every Local Authority, therefore, in column (3) I first restrict the sample to all those transactions for which a RC school is in the area and check whether the results "survive" this restriction, then in column (4) I control for the local RC school performance.

In column (5) I repeat the same exercise as for the baseline model, but focusing only on those transactions which are located within the same data zone, and yet in different catchment areas. Figure 2 illustrates the fact that catchment area and data zone boundaries often overlap, but at times catchment area boundaries cut through the same data zone block. The underlying idea is to compare houses which have no differences in observable neighbourhood characteristics - which collected at the data zone level - other than the school they are served by. Moreover, this is also equivalent to comparing houses which are even closer one another. In order to exploit this variation across boundary but within neighbourhood, data zone fixed effects are used in place of boundary fixed effects. Finally, columns (6)-(7) report results from the model in Equation 3

We can see that in general there are only small changes to the main coefficient, $\widehat{\beta}_{Academic}$, when I try to verify its robustness to a series of factor. Furthermore, in Table 7 I address some concerns in relation to the presence of spatial-time trends by controlling for a function of distance from boundaries (columns (2) and (3)), and accounting for possible heterogeneity in boundary fixed effects over distance (column (4)), different SIMD domains and time (columns (5) and (6) respectively). The conclusion from these robustness checks is that these changes to the main specification do not alter the central conclusion from Table 5. Table A6 and Table A7 in appendix report the same exercise but for the school composition indicator.

[25] For each transaction I measure the distance to the closest 15 private schools and take the median of the inverse distance.

4.2 Discussion

So far in this paper I have examined preferences for some potential indicators of school quality, via their house price capitalisation. Figure 4 can shed light on the correlation between each indicator, thus providing further insights –after Table 2– on the extent to which school performance is multidimensional, and together with results from Table 5 I attempt to address the issue of whether parents are aware of multidimensionality (or absence of it). The scatter plots are built using the 159 schools present in the 350-metres-from-boundary sample and each dot is the 2014-2016 average school-level indicator. The small correlation between academic performance and value added is indicative of the fact that schools with high average performance are not, on average, those which boost pupils attainment,[26] yet from Table 5 emerges that parents prefer the previous over the latter. There are usually two possible interpretations for this: *i)* Parents are not aware of which schools are good at raising their pupils performance - however, in section 1 I have ruled out this scenario as school benchmarks are publicly available and easily accessible. In fact, virtual comparator is reported next to the academic performance indicator; *ii)* Parents simply do not consider value added as a valuable information about the quality of a school.

Furthermore, schools with higher average academic performance are also those with fewer pupils registered for free school meals (r=0.81) and approximately a similar correlation is shared between academic performance and noncognitive skills (r=0.77). Whilst the previous is strongly capitalised into house prices, with approximately the same premium of academic performance, the coefficient on noncognitive skills is small in size (even slightly negative conditional on boundary fixed effects) and hardly ever statistically significant.

This suggests two things: First, perceived potential peer effects as proxied by the share of children not eligible for free meals and average academic achievement might simply reflect the same school dimension, thus providing the same information to parents; second, despite (peer) noncognitive skills are highly correlated with academic performance, these do not appear to be

[26]Note that value added index is relative to the same qualifications the academic performance index is based upon.

capitalised into house prices, which suggests that parents do not value this dimension per se, but might be using academic performance and free meals registration as more comprehensive measures of school "quality".

5 Conclusion

Strategic residential choices based on school quality is widely understood, anecdotally even before it was demonstrated in the academic literature. This paper contributes to the sizeable literature exploring the extent to which school quality is capitalised into house prices, and extends this literature in a number of ways. First, it overcomes the issue of lack of information about value added by using an indicator which has been salient and publicly available for years; second, the indicator can be seen as a contextualised measure of value added, as opposed to the average gain in attainment by the end of school cycle; third, this paper sheds light on the role of peers' noncognitive skills in school choice.

I reach three main conclusions: First, school performance in secondary schools is not strictly multidimensional, but maybe bi-dimensional, with high performing schools not necessarily being also the most effective ones. Second, school effectiveness does not appear to be a key factor in school choice. This result confirms recent findings from MacLeod and Urquiola (2015) and Abdulkadiroğlu et al. (2020) which show that school composition and average performance might be a valid signal of quality when information on effectiveness might not be available, provided that these two dimensions are correlated with effectiveness. In this paper however, not only is performance benchmark publicly available, but it is also only weakly correlated with school-level average performance. Therefore this finding suggests that school effectiveness alone does not appear to be an important factor. One important implication of this is that parents might not be necessarily valuing schools that offer learning improvements but attach more value to schools that have high average performance, even if this is driven by selection.

This leads to the third conclusion, namely the fact that school-level average academic attainment and school composition are equally capitalised into house prices and strongly (and positively) correlated. A possible conclusion is that these are in fact two sides of the same coin, which parents might simply regard as "school quality", and that their individual effects are hard to isolate. Finally, despite the noncognitive component correlating strongly with academic performance and composition, no evidence was found of it having an independent impact on valuation. It could be that this is not part of the utility function altogether, but equally it could simply be that parents also infer noncognitive skills from composition or average outcomes.

In summary the findings of this paper suggest that parents are only responsive to a few specific school characteristics. This raises some concerns in terms of how school policy might adjust to these preferences but also, from an individual perspective, whether considering school "quality" on a single metric might undermine the success and complexity of the learning process. My estimates suggests approximately a 4% house price premium for a one-standard deviation increase in school "performance". A back of the envelope calculation results in parents paying on average a £6,800 premium at the mean.[27]

A number of areas remain for future work, including exploring the extent to which the release of school quality information in new (and different) formats affects parents' perception of school quality and preferences (if at all). The data used in this project have been publicly available from 2014, therefore exploring school preferences prior to this period would be interesting. One interesting further extension would be to consider heterogeneous preferences.[28] Finally, an under researched area remains: the long-term socio-economic and segregational consequences of residence-based attendance systems.

[27] see Table 1

[28] see, e.g. Burgess et al. (2015) Burgess et al. (2019)

References

Abdulkadiroğlu, A., P. A. Pathak, J. Schellenberg, and C. R. Walters (2020): "Do Parents Value School Effectiveness?" *American Economic Review*, 110, 1502–1539.

Ainsworth, R., R. Dehejia, C. Pop-Eleches, and M. Urquiola (2020): "Information, Preferences, and Household Demand for School Value Added," Tech. rep., National Bureau of Economic Research.

Bayer, P., F. Ferreira, and R. McMillan (2007): "A unified framework for measuring preferences for schools and neighborhoods," *Journal of political economy*, 115, 588–638.

Bertrand, M. and J. Pan (2013): "The trouble with boys: Social influences and the gender gap in disruptive behavior," *American economic journal: applied economics*, 5, 32–64.

Beuermann, D., C. K. Jackson, L. Navarro-Sola, and F. Pardo (2018): "What is a Good School, and Can Parents Tell? Evidence on the Multidimensionality of School Output," Tech. rep., National Bureau of Economic Research.

Black, S. E. (1999): "Do better schools matter? Parental valuation of elementary education," *The Quarterly Journal of Economics*, 114, 577–599.

Bono, E. D., M. Francesconi, Y. Kelly, and A. Sacker (2016): "Early maternal time investment and early child outcomes," *The Economic Journal*, 126, F96–F135.

Borbely, D., M. Gehrsitz, S. McIntyre, G. Rossi, and G. Roy (2020): "Multi-grade Classes and Pupil Attainment," *Working Paper*.

Brasington, D. and D. R. Haurin (2006): "Educational outcomes and house values: A test of the value added approach," *Journal of Regional Science*, 46, 245–268.

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

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

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

Clapp, J. M., A. Nanda, and S. L. Ross (2008): "Which school attributes matter? The influence of school district performance and demographic composition on property values," *Journal of urban Economics*, 63, 451–466.

Doris, A., D. O'Neill, and O. Sweetman (2019): "Good Schools or Good Students? The Importance of Selectivity for School Rankings," *IZA Discussion Paper*.

Epple, D. N. and R. Romano (2003): "Neighborhood schools, choice, and the distribution of educational benefits," in *The economics of school choice*, University of Chicago Press, 227–286.

Fack, G. and J. Grenet (2010): "When do better schools raise housing prices? Evidence from Paris public and private schools," *Journal of public Economics*, 94, 59–77.

Figlio, D. N. and M. E. Lucas (2004): "What's in a grade? School report cards and the housing market," *American economic review*, 94, 591–604.

Fiva, J. H. and L. J. Kirkebøen (2011): "Information shocks and the dynamics of the housing market," *The Scandinavian Journal of Economics*, 113, 525–552.

Gelbach, J. B. (2016): "When do covariates matter? And which ones, and how much?" *Journal of Labor Economics*, 34, 509–543.

Gibbons, S. and S. Machin (2003): "Valuing English primary schools," *Journal of urban*

economics, 53, 197–219.
——— (2006): "Paying for primary schools: admission constraints, school popularity or congestion?" *The Economic Journal*, 116, C77–C92.
Gibbons, S., S. Machin, and O. Silva (2013): "Valuing school quality using boundary discontinuities," *Journal of Urban Economics*, 75, 15–28.
Hastings, J., T. J. Kane, and D. O. Staiger (2009): "Heterogeneous preferences and the efficacy of public school choice," *NBER Working Paper*, 2145, 1–46.
Heckman, J. J. and S. Mosso (2014): "The economics of human development and social mobility," *Annu. Rev. Econ.*, 6, 689–733.
Heckman, J. J., J. Stixrud, and S. Urzua (2006): "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior," *Journal of Labor economics*, 24, 411–482.
Hoxby, C. M. (2000): "Does competition among public schools benefit students and taxpayers?" *American Economic Review*, 90, 1209–1238.
——— (2003): "School choice and school productivity. Could school choice be a tide that lifts all boats?" in *The economics of school choice*, University of Chicago Press, 287–342.
Imberman, S. A. and M. F. Lovenheim (2016): "Does the market value value-added? Evidence from housing prices after a public release of school and teacher value-added," *Journal of Urban Economics*, 91, 104–121.
Jackson, C. K. (2018): "What do test scores miss? The importance of teacher effects on non–test score outcomes," *Journal of Political Economy*, 126, 2072–2107.
Kautz, T., J. J. Heckman, R. Diris, B. Ter Weel, and L. Borghans (2014): "Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success," Tech. rep., National Bureau of Economic Research.
LearningDirectorate (2016): "Choosing a school: a guide for parents, revised November 2016," *Scottish Goverment*, `https://www.gov.scot/publications/choosing-school-guide-parents-nov-16/`.
Lee, D. S. and T. Lemieux (2010): "Regression discontinuity designs in economics," *Journal of economic literature*, 48, 281–355.
Lleras, C. (2008): "Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings," *Social Science Research*, 37, 888–902.
MacLeod, W. B. and M. Urquiola (2015): "Reputation and school competition," *American Economic Review*, 105, 3471–88.
Norris, J. (2020): "Peers, parents, and attitudes about school," *Journal of Human Capital*, 14, 290–342.

Tables

Table 1: Summary Statistics

	Full Sample		350 metres		300 metres		250 metres	
	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.
ln(House Price)	11.81	0.68	11.84	0.69	11.84	0.69	11.85	0.69
House Price (Thousands £)	166.74	115.98	173.87	127.61	174.70	128.58	175.77	129.83
% achieving SCQF level 6[a]	39.10	13.00	37.10	14.06	36.90	13.94	36.90	13.97
% achieving SCQF level 5[b]	66.24	11.53	63.82	12.53	63.65	12.48	63.67	12.52
% Literacy & Numeracy SCQF level 5	58.04	11.15	56.10	12.27	56.02	12.23	56.07	12.23
Dropout Rate (Prior to S6)	64.42	11.29	63.52	11.52	63.35	11.57	63.36	11.55
% of Leavers in Higher Education	36.46	11.29	35.27	12.52	35.14	12.47	35.11	12.49
% of Leavers in Further Education	24.30	6.70	24.81	6.37	24.86	6.39	24.85	6.37
% of Leavers Working	29.35	6.98	27.17	6.07	27.20	6.17	27.28	6.23
% of Leavers in Positive Destination	91.41	3.87	90.36	4.03	90.26	4.08	90.24	4.09
% of Pupils not on FSM	85.58	8.70	82.28	9.34	82.03	9.33	82.03	9.32
SCQF level 6 Value Added	−0.78	5.50	−0.99	5.35	−1.00	5.25	−0.99	5.34
SCQF level 5 Value Added	−0.40	5.02	−0.89	4.98	−0.87	5.05	−0.84	5.10
Literacy & Numeracy - Value Added	−0.64	6.38	−0.96	6.47	−0.86	6.45	−0.84	6.41
Attendance Rate	91.89	1.76	91.42	1.87	91.40	1.86	91.40	1.86
Authorised Absence Rate	5.44	1.35	5.48	1.49	5.50	1.52	5.50	1.52
Unauthorised Absence Rate	2.58	1.53	2.99	1.71	3.00	1.72	3.00	1.72
Exclusion Rate	0.08	0.08	0.10	0.08	0.10	0.08	0.10	0.08
School Size	817.65	321.36	901.79	303.37	900.92	309.14	900.73	309.16
No. of Schools	272		159		150		150	
No. of Sales	220,396		52,519		45,338		38,085	

Note: House prices refer to transactions occurred between 2015 and 2018, whereas school-level information referes to academic years 2013/14, 2014/15 and 2015/16. Attainment data are elaborated by the Learning Directorate at the Scottish Government (SG) Analytical Services and available on the SG website.

[a] At least four awards

[b] At least four awards

Table 2: Factor Analysis

Variables	Factor 1	Factor 2	Factor 3
% achieving SCQF level 6[a]	0.906		
% achieving SCQF level 5[b]	0.915		
% Attaining Literacy & Numeracy SCQF level 5	0.790		
Dropout Rate (Prior to S6)	0.784		
% of Leavers in Higher Education	0.858		
% of Leavers in Further Education			
% of Leavers Working			
% of Leavers in Positive Destination	0.727		
Attendance Rate	0.866		
Exclusion Rate	−0.651		
Authorised Absence Rate			0.737
Unauthorised Absence Rate			−0.789
% of Pupils not on FSM	0.788		
SCQF level 6 Value Added		0.698	
SCQF level 5 Value Added		0.702	
Literacy & Numeracy - Value Added		0.771	
Proportion	0.568	0.131	0.093
Cumulative	0.568	0.699	0.792

Note: This table shows the factor analysis run on the full sample of schools. The rotated factors being retained are those with eigenvalue ≥ 1 and they account for nearly 80% of the overall variance, 57% of which is in the first one.

[a] At least four awards
[b] At least four awards

Table 3: Indexes

Variables	Weights
Academic Index	
% achieving SCQF level 6[a]	0.51
% achieving SCQF level 5[b]	0.51
% Attaining Literacy & Numeracy SCQF level 5	0.48
% of Leavers in Higher Education	0.50
Value Added Index	
SCQF level 6 - Value Added	0.58
SCQF level 5 - Value Added	0.61
Literacy & Numeracy - Value Added	0.55
Noncognitive Skills Index	
Attendance Rate	-0.63
Unauthorised Absence Rate	0.56
Exclusion Rate	0.54
Composition	
% of Pupils not on FSM	1

Note: This table shows how the main indexes used in this analysis have been built. Except for *Composition*, which is fully proxied by the school-level % of pupils who are not eligible for free school meals, the other indexes are the first components of three different Principal Component Analysis run on the relevant groups of variables. The "Weights" of each variable correspond to the PCA loadings of each variable at hand.

[a] At least four awards

[b] At least four awards

Table 4: Balancing

	(1) 350 metres	(2) 300 metres	(3) 250 metres
ln(House Price)	0.0921	0.0894	0.0821
	(0.0613)	(0.0629)	(0.0638)
Academic	0.989***	0.996***	1.001***
	(0.141)	(0.144)	(0.146)
Average No. of Rooms[a]	0.130	0.137	0.124
	(0.0990)	(0.0987)	(0.0987)
Population Density	−1.408	−1.242	−0.568
	(6.397)	(6.589)	(6.673)
% On Social Renting	−3.752**	−3.890**	−3.885**
	(1.573)	(1.568)	(1.553)
% Renting	−0.685	−0.686	−0.436
	(2.100)	(2.114)	(2.130)
% Other Than White	−0.0436	−0.0937	−0.0178
	(0.613)	(0.613)	(0.632)
% No Qualification	−1.219	−1.276	−1.329
	(1.503)	(1.511)	(1.510)
% Higher Qualification	2.482	2.390	2.396
	(2.628)	(2.650)	(2.674)
Median Age	0.960	0.875	0.654
	(0.752)	(0.756)	(0.761)
% Female-Headed Households with Children	−1.132	−1.190*	−1.128
	(0.688)	(0.696)	(0.712)
Crime Rate (per 100k people)[b]	−17.33	−23.46	−1.128
	(34.47)	(36.06)	(0.712)
Income Deprivation Rate	−1.539*	−1.606*	−1.574*
	(0.817)	(0.820)	(0.821)
Overcrowding Rate	−1.609	−1.684	−1.494
	(1.079)	(1.099)	(1.094)
Observations	52,519	45,338	38,085
No. of Schools	159	150	150
Cluster	School	School	School
Sample	All	All	All

Notes: Each coefficient results from a regression of the variable at hand on a binary indicator of whether the property is located on the better-performing side of the catchment area boundary, based on the *Academic* performance indicator. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

[a]Scotland Census 2011, © Crown copyright, Data supplied by National Records of Scotland.
[b]Scottish Index of Multiple Deprivation (SIMD) © Crown copyright 2016.

Table 5: Main Results

Dependent Variable: ln(House Prices)	(1) Full Sample	(2) 350 metres	(3) 350 metres	(4) 350 metres	(5) 350 metres	(6) Cross-LAs
Academic	0.032***	0.020	0.030***	0.029***	0.034***	0.026***
	(0.008)	(0.018)	(0.007)	(0.007)	(0.008)	(0.006)
Value Added	−0.001	−0.027**	0.001	−0.000	−0.004	−0.001
	(0.006)	(0.012)	(0.006)	(0.006)	(0.007)	(0.005)
Noncognitive	0.016*	0.029*	−0.008	−0.004	−0.005	−0.006
	(0.008)	(0.017)	(0.007)	(0.007)	(0.008)	(0.007)
Composition	0.038***	0.060***	0.045***	0.042***	0.055***	0.038***
	(0.011)	(0.017)	(0.012)	(0.013)	(0.015)	(0.011)
Observations	220,396	52,519	52,519	51,426	33,946	58,379
No. of Schools	272	159	159	145	75	163
Mean	11.81	11.84	11.84	11.84	11.96	11.84
SD	0.68	0.69	0.69	0.69	0.69	0.68
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Boundary FE	No	No	Yes	Yes	Yes	Yes
Sample	All	All	All	Urban $\geq$ (10k)	Urban $\geq$ (125k)	All
Adjusted R-squared	0.467	0.500	0.561	0.565	0.577	0.559

Notes: Each Column represents a specification in which $ln(HousePrices)$ is regressed on one indicator at the time. Control variables include the covariates from Table 4 plus a set of year-of-transaction dummies and school size. "Urban $\geq$ (10k)" refers to settlements between 10,000 and 124,999 people, whilst "Urban $\geq$ (125k)" areas have 125,000 inhabitants or more. These are classifications provided by the Scottish Government. "Cross-LAs" includes houses in proximity of those catchment area boundaries which coincide with borders between Local Authorities. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Robustness Checks - Academic

Dependent Variable: ln(House Prices)	(1) Baseline	(2) Extra Controls	(3) RC Sample	(4) RC Sample	(5) "Same Block"	(6) $\Delta ln(HousePrices)$	(7) $\Delta ln(HousePrices)$
Academic	0.030***	0.028***	0.032***	0.030***	0.044***	0.047**	0.053***
	(0.007)	(0.007)	(0.008)	(0.008)	(0.016)	(0.020)	(0.018)
PT ratio		0.010					
		(0.009)					
(Roll/Capacity)×100		0.000					
		(0.001)					
Private School Index		0.073***					
		(0.020)					
Families with Children (%)		0.002**					
		(0.001)					
RC Higher				0.002*			
				(0.001)			
Observations	52,519	52,519	45,541	45,541	13,335	25,658	25,658
No. of Schools	159	159	130	130	137	128	128
Mean	11.84	11.84	11.84	11.84	11.89	.12	.12
SD	0.69	0.69	0.71	0.71	0.71	0.69	0.69
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Boundary FE	Yes	Yes	Yes	Yes	No	No	No
Datazone FE	No	No	No	No	Yes	No	No
LAs FE	No	No	No	No	No	Yes	Yes
IDW	No	No	No	No	No	No	Yes
Sample	All	All	RC Schools	RC Schools	All	All	All
Adjusted R-squared	0.561	0.562	0.580	0.580	0.593	0.336	0.277

Notes: Control variables include the covariates from Table 4 plus a set of year-of-transaction dummies and school size. Private School Index is calculated for each sale as the median of the inverse distance between the sale and the closest fifteen private schools. "Same Block" focuses on sales within the same Data Zone but on opposite side of the catchment area boundaries and still within 350 metres from it. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

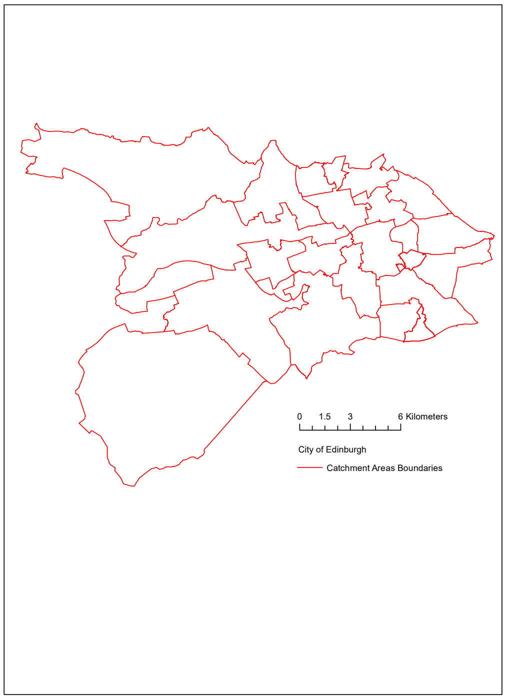
Table 7: Spatial Trends - Academic

Dependent Variable: ln(House Prices)	(1) Baseline	(2) Distance	(3) Distance Sq	(4) Distance × Boundaries	(5) SIMD × Boundaries	(6) Years × Boundaries
Academic	0.030^{***}	0.030^{***}	0.030^{***}	0.029^{***}	0.035^{***}	0.035^{***}
	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)
Distance (100s metres)		0.001	0.033^{**}	0.025		
		(0.005)	(0.016)	(0.120)		
Distance Sq.			-0.009^{**}			
			(0.004)			
Observations	52,519	52,519	52,519	52,519	52,519	52,519
No. of Schools	159	159	159	159	159	159
Mean	11.84	11.84	11.84	11.84	11.84	11.84
SD	0.69	0.69	0.69	0.69	0.69	0.69
Sample	All	All	All	All	All	All
Adjusted R-squared	0.561	0.561	0.561	0.567	0.581	0.567
F-test			2.675	2547	3253	393.6
Prob > F			0.0720	0	0	0

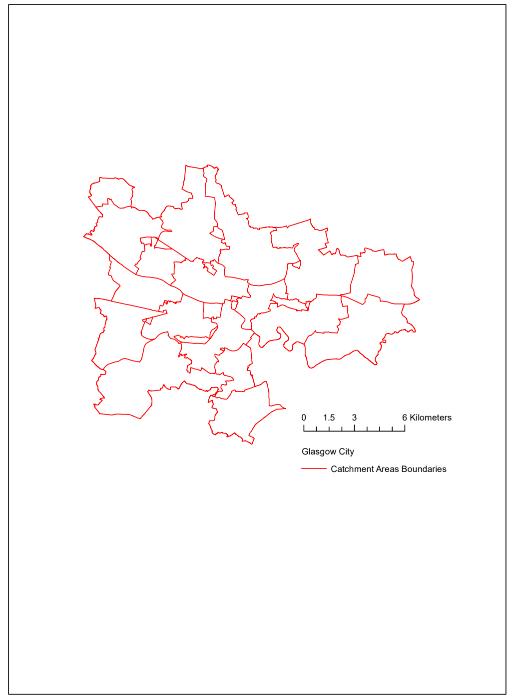
Notes: Each specification contains the covariates from Table 4 plus a set of year-of-transaction dummies and school size as well as boundary fixed effects. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

Figures

Figure 1: Secondary Schools Catchment Areas



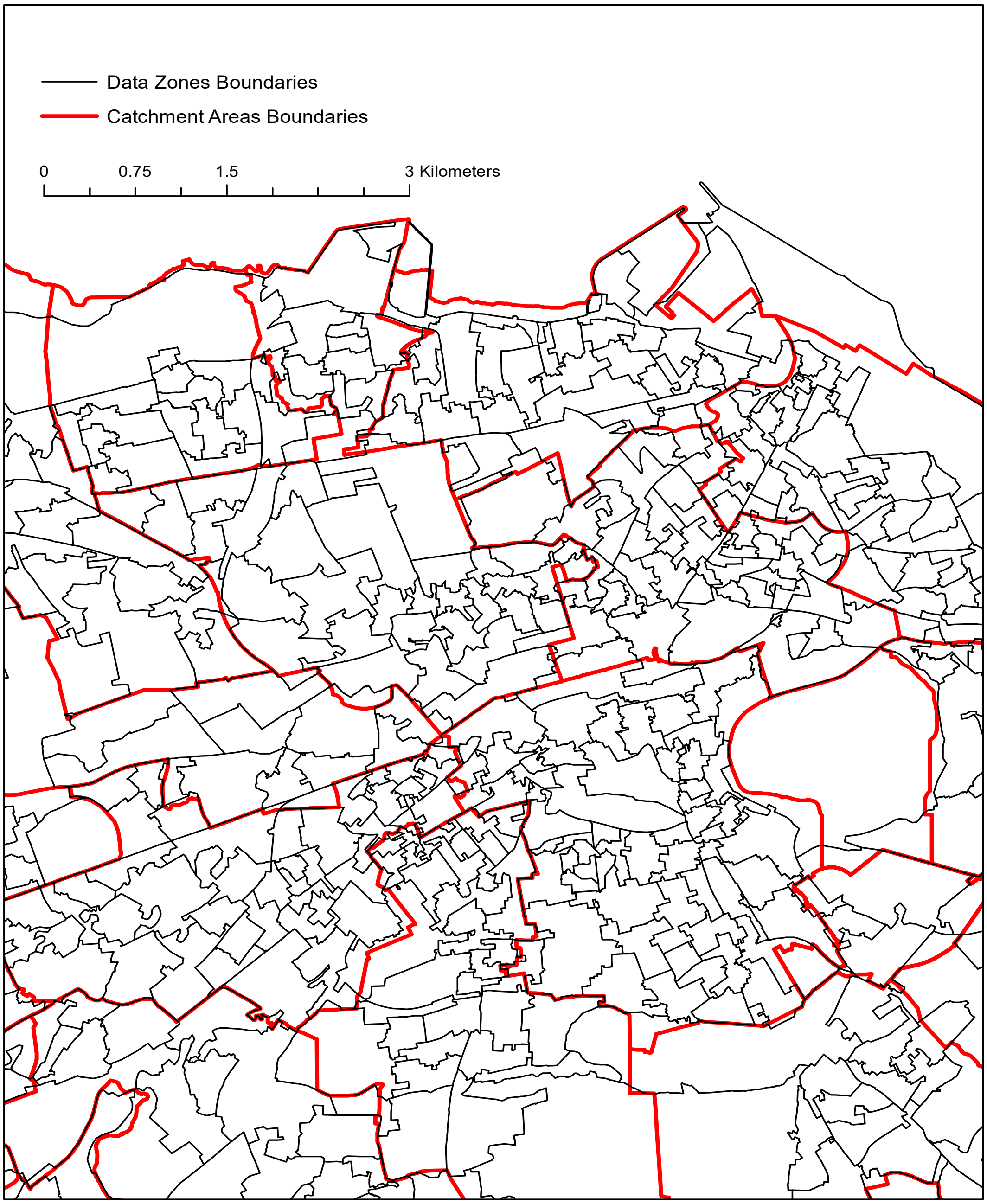
(a) City of Edinburgh



(b) Glasgow City

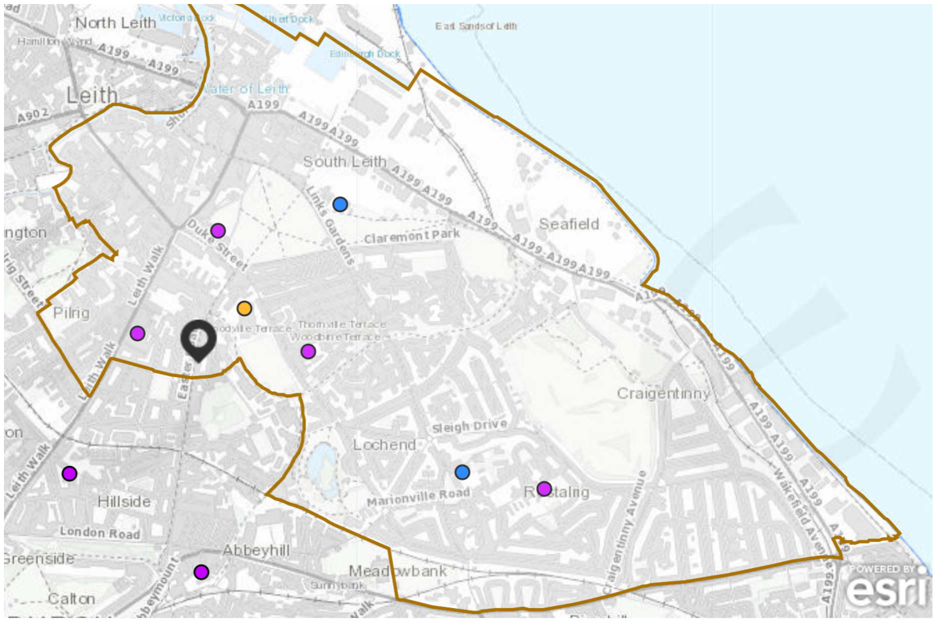
Note: Figure 1a and Figure 1b provide an example of Secondary Schools catchment areas in Edinburgh and Glasgow respectively.

Figure 2: Data Zones



Note: This figure shows how Data Zone blocks nest within Secondary Schools Catchment Areas. In particular, the figure refers to central and northern areas in Edinburgh.

Figure 3: Catchment Area Example

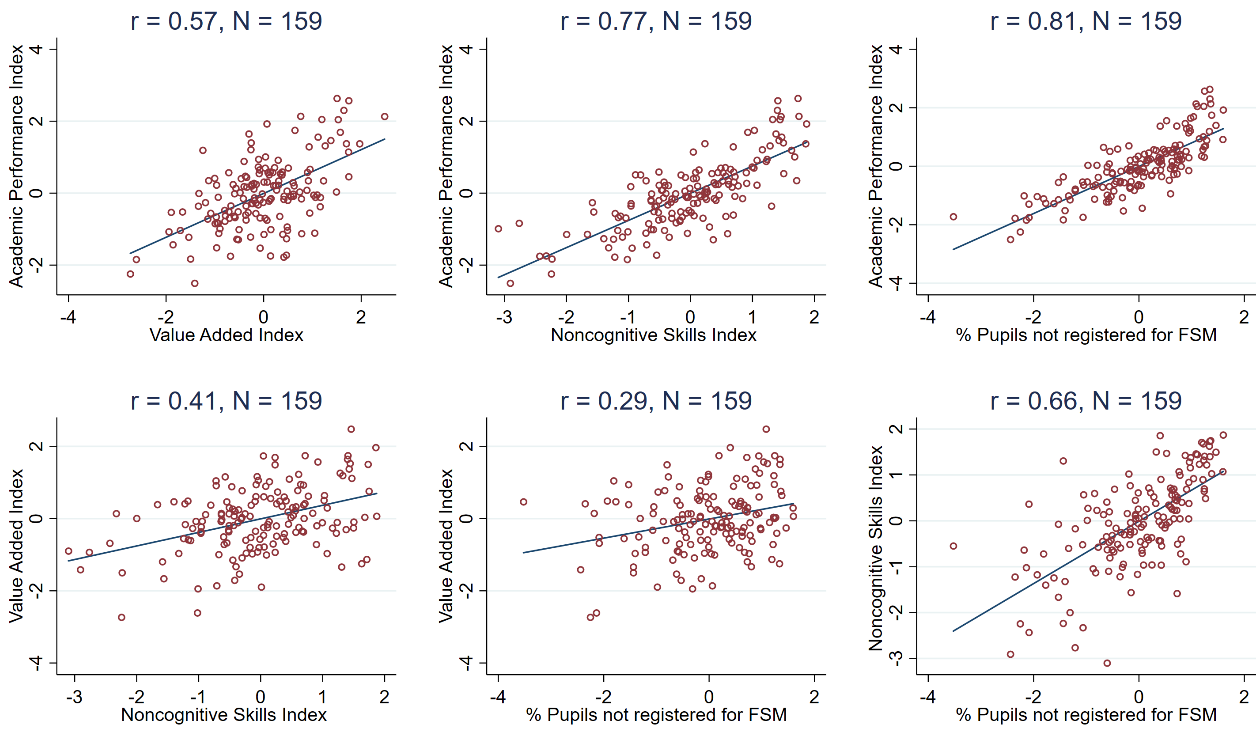


(a) Leith Academy - Edinburgh

(b) Portobello High Boundary

Note: Figure 3a shows the catchment area of Leith Academy, whereas Figure 3b shows a portion of its lower boundary, which separates Leith Academy catchment area from the one of Portobello High School.
Source: The City of Edinburgh Council website

Figure 4: Correlation Between Indicators



Note: In this set of scatter plots is reported the correlation between performance indicators at the school level. In particular, each dot corresponds to a school whose value has been averaged across all school years (2013/14, 2014/15 and 2015/16)

Appendix A

Table A1: Secondary School Qualifications

Grade	S4	S5	S6
SQA	National 5	Higher	Higher/Advanced Higher
SVQs	Modern Apprenticeship	Modern Apprenticeship	Modern Apprenticeship
	SVQ	SVQ	SVQ
SCQF	Level 5	Level 6	Level 7

Notes: Scottish Qualification Authority qualifications, Scottish Vocational Qualifications (SVQs) and their equivalent in terms of Scottish Credit and Qualifications Framework levels, by "Senior Phase" stage.

Table A2: Balancing - Other Indexes

	(1) Value Added	(2) Noncogntive	(3) Composition
ln(House Price)	0.029	0.043	0.061
	(0.055)	(0.071)	(0.067)
Index	1.013***	0.972***	0.878***
	(0.131)	(0.155)	(0.137)
Average No. of Rooms[a]	0.058	0.077	0.155
	(0.085)	(0.097)	(0.104)
Population Density	2.062	−2.700	−2.625
	(4.792)	(6.524)	(6.545)
% on Social Renting	−1.794	−1.639	−3.990**
	(1.253)	(1.836)	(1.701)
% Renting	−0.125	−1.852	−1.440
	(1.720)	(1.954)	(2.122)
% Other Than White	−0.079	−0.596	−0.282
	(0.433)	(0.606)	(0.562)
% No Qualification	−0.597	−0.292	−0.720
	(1.218)	(1.643)	(1.588)
% Higher Qualification	1.254	1.251	1.180
	(2.095)	(2.768)	(2.691)
Median Age	0.605	1.292*	1.036
	(0.650)	(0.739)	(0.761)
% Female-Headed Households with Children	−0.661	−0.934	−1.125
	(0.601)	(0.794)	(0.731)
Crime Rate [b]	−3.197	−7.276	−66.019*
	(31.265)	(54.257)	(34.809)
Income Deprivation Rate	−0.611	−1.606*	−1.480*
	(0.680)	(0.820)	(0.871)
Overcrowding Rate	−0.480	−1.537	−1.806
	(0.835)	(1.041)	(1.111)
Observations	52,519	52,519	52,519
No. of Schools	159	159	159
Cluster	School	School	School
Sample	All	All	All

Notes: Each coefficient results from a regression of the variable at hand on a binary indicator of whether the property is located on the better-performing side of the catchment area boundary, based on the *Value Added*, *Noncognitive* and *Composition* indicators in column (1), (2) and (3) respectively. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

[a] Scotland Census 2011, © Crown copyright, Data supplied by National Records of Scotland.
[b] Scottish Index of Multiple Deprivation (SIMD) © Crown copyright 2016.

Table A3: Summary Statistics - Control variables

	Full Sample		350 metres		300 metres		250 metres	
	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.
Average No. of Rooms[a]	5.02	0.88	4.78	0.89	4.78	0.90	4.78	0.89
Population Density	41.03	36.91	51.08	42.22	50.70	42.25	50.00	41.47
% on Social Renting	22.71	20.42	24.14	22.41	23.93	22.39	23.88	22.38
% Renting	11.85	10.77	13.88	12.86	13.94	12.91	13.95	12.87
% Other Than White	3.26	5.31	4.79	5.62	4.79	5.57	4.72	5.33
% No Qualification	26.77	11.75	25.92	12.80	25.85	12.86	25.85	12.86
% Higher Qualification	26.02	13.73	27.98	16.02	28.16	16.05	28.14	16.08
Median Age	41.34	6.76	39.67	7.06	39.69	7.11	39.69	7.10
% Female-Headed Households with Children	12.57	8.59	13.78	9.30	13.73	9.34	13.80	9.38
Crime Rate[b]	310.50	447.31	414.14	649.48	421.63	671.97	425.47	690.99
Income Deprivation Rate	12.34	9.64	13.45	10.57	13.40	10.59	13.41	10.61
Overcrowding Rate	10.95	7.92	13.17	8.70	13.17	8.72	13.07	8.61
Private School Index[c]	−0.10	0.88	0.31	1.15	0.31	1.15	0.32	1.15
No. of DZs	6,566		2,300		2,095		1,929	
No. of Sales	220,396		52,519		45,338		38,085	

Notes: Except for *Private School Index* which has been created at the house-level, control variables are collected at the Data Zone level, as re-defined in 2011. Census variables refers to Scotland Census 2011, whilst Scottish Index of Multiple Deprivation variables refer to 2016.

[a] Scotland Census 2011, © Crown copyright, Data supplied by National Records of Scotland.
[b] Scottish Index of Multiple Deprivation (SIMD) © Crown copyright 2016.
[c] Elaborated by the author.

Table A4: Main Results - 300 metres

Dependent Variable: ln(House Price)	(1)	(2)	(3)	(4)	(5)
	300 metres	300 metres	300 metres	300 metres	Cross-LAs
Academic	0.019	0.030***	0.029***	0.038***	0.024***
	(0.018)	(0.007)	(0.007)	(0.008)	(0.007)
Value Added	−0.026**	0.001	−0.000	−0.002	−0.003
	(0.012)	(0.006)	(0.006)	(0.007)	(0.006)
Noncognitive	0.030*	−0.008	−0.003	−0.003	−0.007
	(0.017)	(0.009)	(0.009)	(0.009)	(0.009)
Composition	0.059***	0.046***	0.043***	0.061***	0.038***
	(0.018)	(0.013)	(0.013)	(0.016)	(0.011)
Observations	45,338	45,338	44,480	29,370	50,245
No. of Schools	150	150	142	75	156
Mean	11.84	11.84	11.84	11.96	11.84
SD	0.69	0.69	0.69	0.68	0.69
Controls	Yes	Yes	Yes	Yes	Yes
Boundary FE	No	Yes	Yes	Yes	Yes
Sample	All	All	Urban ≥10k	Urban ≥125k	All
Adjusted R-squared	0.507	0.564	0.568	0.580	0.563

Notes: Each Column represents a specification in which $ln(HousePrices)$ is regressed on one indicator at the time. Control variables include the covariates from Table 4 plus a set of year-of-transaction dummies and school size. "Urban ≥ (10k)" refers to settlements between 10,000 and 124,999 people, whilst "Urban ≥ (125k)" areas have 125,000 inhabitants or more. These are classifications provided by the Scottish Government. "Cross-LAs" includes houses in proximity of those catchment area boundaries which coincide with borders between Local Authorities. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Main Results - 250 metres

Dependent Variable: ln(House Price)	(1)	(2)	(3)	(4)	(5)
	250 metres	250 metres	250 metres	250 metres	Cross-LAs
Academic	0.017	0.027***	0.026***	0.035***	0.020***
	(0.018)	(0.007)	(0.007)	(0.008)	(0.007)
Value Added	−0.028**	0.000	−0.000	−0.001	−0.005
	(0.012)	(0.006)	(0.006)	(0.008)	(0.006)
Noncognitive	0.031*	−0.010	−0.006	−0.005	−0.010
	(0.017)	(0.009)	(0.009)	(0.010)	(0.009)
Composition	0.056***	0.042***	0.040***	0.059***	0.032***
	(0.018)	(0.014)	(0.014)	(0.016)	(0.012)
Observations	38,085	38,085	37,431	24,941	42,032
No. of Schools	150	150	142	75	153
Mean	11.85	11.85	11.85	11.96	11.85
SD	0.69	0.69	0.69	0.69	0.69
Controls	Yes	Yes	Yes	Yes	Yes
Boundary FE	No	Yes	Yes	Yes	Yes
Sample	All	All	Urban $\geq$ 10k	Urban $\geq$ 125k	All
Adjusted R-squared	0.505	0.566	0.569	0.580	0.564

Notes: Each Column represents a specification in which $ln(HousePrices)$ is regressed on one indicator at the time. Control variables include the covariates from Table 4 plus a set of year-of-transaction dummies and school size. "Urban $\geq$ (10k)" refers to settlements between 10,000 and 124,999 people, whilst "Urban $\geq$ (125k)" areas have 125,000 inhabitants or more. These are classifications provided by the Scottish Government. "Cross-LAs" includes houses in proximity of those catchment area boundaries which coincide with borders between Local Authorities. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Robustness Checks - Composition

Dependent Variable: ln(House Prices)	(1) Baseline	(2) Extra Controls	(3) RC Sample	(4) RC Sample	(5) "Same Block"	(6) $\Delta ln(HousePrices)$	(7) $\Delta ln(HousePrices)$
Composition	0.045***	0.043***	0.040***	0.038***	0.042**	0.057***	0.055***
	(0.012)	(0.013)	(0.013)	(0.013)	(0.020)	(0.018)	(0.018)
PT ratio		0.009					
		(0.008)					
(Roll/Capacity)×100		0.000					
		(0.001)					
Private School Index		0.073***					
		(0.020)					
Families with Children (%)		0.002**					
		(0.001)					
RC Higher				0.002*			
				(0.001)			
Observations	52,519	52,519	45,541	45,541	13,335	26,529	26,529
No. of Schools	159	159	130	130	137	108	108
Mean	11.84	11.84	11.84	11.84	11.89	.06	.06
SD	0.69	0.69	0.71	0.71	0.71	0.69	0.69
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Boundary FE	Yes	Yes	Yes	Yes	No	No	No
Datazone FE	No	No	No	No	Yes	No	No
LAs FE	No	No	No	No	No	Yes	Yes
IDW	No	No	No	No	No	No	Yes
Sample	All	All	RC Schools	RC Schools	All	All	All
Adjusted R-squared	0.561	0.562	0.580	0.580	0.593	0.321	0.253

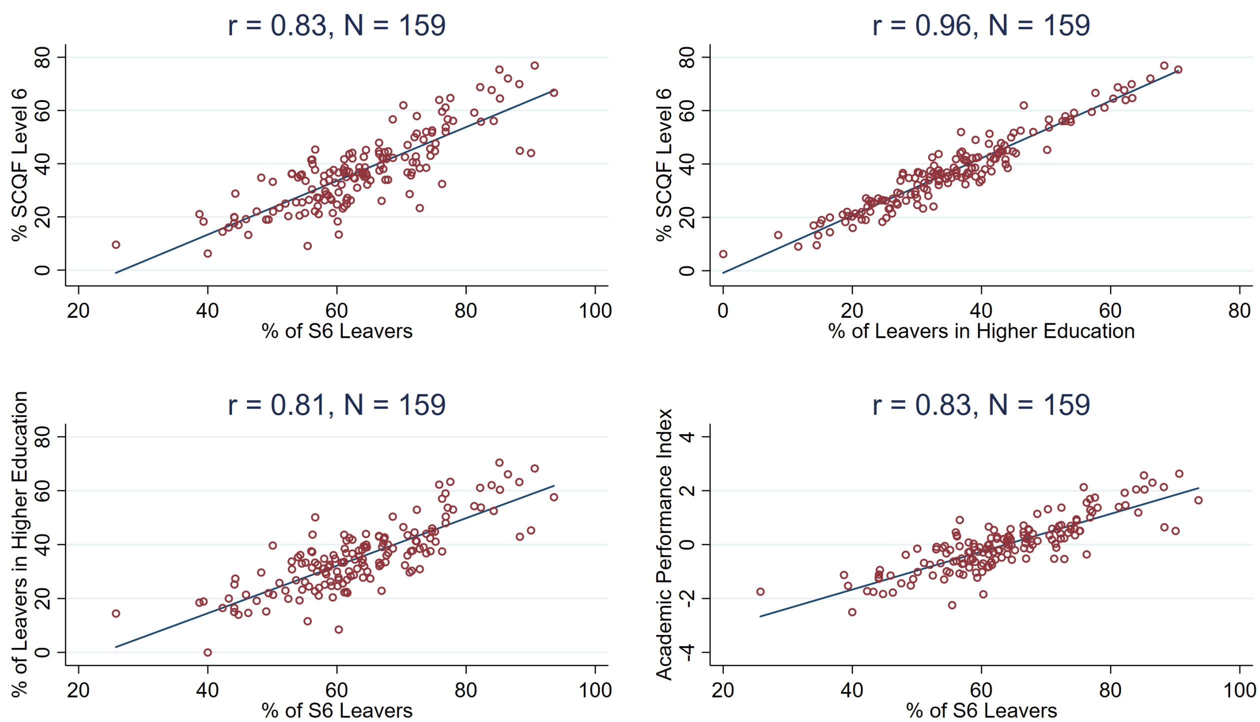
Notes: Control variables include the covariates from Table 4 plus a set of year-of-transaction dummies and school size. Private School Index is calculated for each sale as the median of the inverse distance between the sale and the closest fifteen private schools. "Same Block" focuses on sales within the same Data Zone but on opposite side of the catchment area boundaries and still within 350 metres from it. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Spatial Trends - Composition

Dependent Variable: ln(House Prices)	(1) Baseline	(2) Distance	(3) Distance Sq.	(4) Distance × Boundaries	(5) SIMD × Boundaries	(6) Years × Boundaries
Composition	0.045***	0.045***	0.045***	0.046***	0.042***	0.045***
	(0.012)	(0.012)	(0.012)	(0.013)	(0.014)	(0.012)
Distance (100s metres)		0.000	0.000**	0.000		
		(0.000)	(0.000)	(0.001)		
Distance Sq.			−0.000**			
			(0.000)			
Observations	52,519	52,519	52,519	52,519	52,519	52,519
No. of Schools	159	159	159	159	159	159
Mean	11.84	11.84	11.84	11.84	11.84	11.84
SD	0.69	0.69	0.69	0.69	0.69	0.69
Sample	All	All	All	All	All	All
Adjusted R-squared	0.561	0.561	0.561	0.567	0.581	0.567
F-test			2.676	1.370e+10	4626	3663
Prob > F			0.0720	0	0	0

Notes: Each specification contains the covariates from Table 4 plus a set of year-of-transaction dummies and school size as well as boundary fixed effects. Standard errors (in parenthesis) are clustered at the school catchment area level. *** p<0.01, ** p<0.05, * p<0.1.

Figure A1: SCQF Level 6 and other Academic Performance Indicators.



Note: In this set of scatter plots is reported the correlation between % of Pupils achieving 4 or more awards at SCQF level 6 (or better) and other, more obvious, indicators of academic performance at the school level. In particular, each dot corresponds to a school whose value has been averaged across all school years (2013/14, 2014/15 and 2015/16)

Appendix B

Table B1: Variables Description

	Time	Description
% SCQF level 6	2014, 2015, 2016	% leavers with four or more awards at SCQF level 6
% SCQF level 5	2014, 2015, 2016	% leavers with four or more awards at SCQF level 5 (or above)
% Achieving Literacy & Numeracy SCQF level 5	2014, 2015, 2016	Level 5 or above
Dropout Rate (Prior to S6)	2014, 2015, 2016	see variable name
% of Leavers in Higher Education	2014, 2015, 2016	see variable name
% of Leavers in Further Education	2014, 2015, 2016	see variable name
% of Leavers Working	2014, 2015, 2016	see variable name
% of Leavers in Positive Destination	2014, 2015, 2016	see variable name
% of Pupils not on FSM	2014, 2015, 2016	see variable name
SCQF level 6 - Value Added	2014, 2015, 2016	SCQF level 6 - SCQF level 6 Virtual Comparator
SCQF level 5 - Value Added	2014, 2015, 2016	SCQF level 5 - SCQF level 5 Virtual Comparator
Literacy & Numeracy - Value Added	2014, 2015, 2016	Lit&Num - Lit&Num Virtual Comparator
Attendance Rate	2013/2015 averaged	$(\frac{Attendance}{No.Half-dayOpenings})\times100$
Authorised Absence Rate	2013/2015 averaged	$(\frac{AuthorisedAbs}{No.Half-dayOpenings})\times100$
Unauthorised Absence Rate	2013/2015 averaged	$(\frac{UnauthAbs}{No.Half-dayOpenings})\times100$
Exclusion Rate	2013/2015 averaged	$(\frac{Half-daysmissedbecauseofexclusion}{No.Half-dayOpenings})\times100$
School Size	2014, 2015, 2016	No. of pupils in school

Notes: Data on Attendance, Absence and Exclusions are collected every two years and as such they are the average of the 2013 and 2015 values. All the other variables are used individually for school years 2013/14, 2014/15 and 2015/16.