PAYING FOR PRIMARY SCHOOLS: SUPPLY CONSTRAINTS, SCHOOL POPULARITY OR CONGESTION?

Stephen Gibbons* and Stephen Machin**

April 2004

*  Department of Geography, Centre for the Economics of Education and Centre for Economic Performance, London School of Economics

**  Department of Economics, University College London, Centre for the Economics of this deEducation and Centre for Economic Performance, London School of Economics

Abstract

School quality is capitalised in house prices if access to schools is rationed by residential location. However, the admissions system in England is such that there is no systematic, deterministic relationship between where a family lives and the school children attend. An empirical investigation of how the process and value of school choice is revealed in housing markets must therefore take account of factors that change the probabilities of school admission. We focus on two such factors: distance between school and residence, and whether a school has reached its capacity in terms of pupil numbers. We generate empirical predictions from three different theoretical approaches that link house prices to school performance, distance to school and capacity. These are respectively based upon supply constraints, school popularity and congestion effects. In line with other studies and our earlier work, we find that test-score based school performance increases property prices. However, it is only the best one in five schools that generate higher than average prices close-by. We also find evidence of ‘herd’ behaviour on the part of parents in school choice, in that prices are higher close to popular, over-capacity schools. Thus the empirical evidence is more in line with the school popularity model of the relation between house prices and primary school performance.
Executive summary

Anyone with school age children worries about choosing the best school for their child. However, the processes by which school choices are made, or indeed whether one gets an opportunity to choose, are complex. Some opt for education outside the state sector - if they can afford it. Others must adopt some other strategy to try to get their child into a decent state school. In general, this means choosing a home near the school, because admission is restricted to those who live close by.

The implications of this for house prices are well known, through anecdote and through media coverage. There is also an emerging body of harder evidence that these patterns of demand for neighbourhood schools are important. In this study we extend existing work to look more closely at how decisions over primary schooling are revealed in house price patterns.

We consider some special features of the English school admissions system that add to complexity of the decision process. For a start, catchment areas are vaguely defined, and it is usually a question of ‘nearest-in-first-in’. This leads to uncertainty about which residential locations are best. The problem is compounded by the fact that schools have limited number of pupil places, usually capped by a limit to class sizes, but parents do not know for certain whether a school will be over or under-subscribed. Parents also have imperfect and confusing information about school quality. Faced with wide disparities in state school performance, they scrutinise performance ‘league tables’, study Ofsted reports, and listen attentively to dinner party chat, to work out where the good schools are. We ask how these issues affect the process of school choice and investigate their influence on house price.

In common with other studies, we show a house price premium related to the performance of the nearest primary schools. But some of our findings run counter to common perceptions:

- A ten-percentage point improvement in the ‘league-table’ performance (at age 11, Key Stage 2) adds at least 2% to the price of properties located next to the school
• Despite this, primary schools are, in general not desirable local amenities. Only the 1-in-5 top performing schools will, on average, generate high-prices in their immediate neighbourhoods.
• The premium paid for ‘league-table’ performance is higher if local schools are full to capacity. There is also a premium for living close to a full-school, even if the league table performance is not outstanding. We interpret this as evidence of ‘herd’ behaviour in school choice. A full-school is a popular school, and so - in the eyes of eager parents - a good school.
• Although a school that was full to capacity last year will likely be hard to get into next year, this does not seem to drive up the price of houses very close to good schools relative to those more distant.
• Primary school ‘league-table’ performance has very little impact beyond 750m.

These findings show that primary school performance is a valuable local commodity. This adds weight to the argument that school admissions procedures lead to ‘selection by income’ at primary school level. A move from a weak school to a top performing school would be worth around 15% on property prices, or around £16500 in our sample from London and its surroundings. However, the results also show that some decisions seem misguided. The same improvement in less popular, under-capacity schools will cost around £4400 less.
1. INTRODUCTION

Anyone with school age children worries about choosing the best school for their child. However, the processes by which school choices are made, or indeed whether one gets an opportunity to choose, are complex. Some opt for education outside the state sector - if they can afford it. Others must adopt some other strategy to try to get their child into a decent state school. In general, this means choosing a home near the school, because admission is restricted to those who live close by.

It is therefore not surprising that the issue of which schools are better for parents to send their children to features prominently in middle class dialogue. House prices and residential location decisions are a strong point of discussion, and in this perceptions of what constitutes a better school are important. As in other group discussion environments, some talking points are likely to take place on the basis of limited (or maybe out-of-date) information and it is evident that this may shape the views of parents on schooling decisions for their children. This forms the backdrop for what we study in this paper, namely the links between primary school performance, residential location, school capacity and house prices. We set up and test various models of school choice, implementing empirical tests using detailed house price and school data in London and the South East of England.

The primary school admission system is organised in England such that there is no deterministic relationship between location of residence and the primary school attended. In fact, allocation purely on the basis of place of residence became illegal in the 1990s, after some well publicised court cases. In principle, parental preference is what counts. In practice, good primary schools in urban areas are often full or over-subscribed, and a number of geographically-based
over-subscription criteria come into play. Living within some defined ‘catchment area’ can be important\(^1\); but it is often just a case of nearest-in/first-in.

A standard approach to placing a monetary valuation on school quality is to trace out the affects of neighbourhood school quality on property values. But the ‘nearest-in/first-in’ feature of the admissions system in England generally rules out empirical strategies that exploit well defined attendance district boundaries [as used in US work like Black (1999), or Bogart and Cromwell (2000)]. For primary schools, the only case where boundaries will be fairly non-porous is at the border between Local Education Authorities (LEAs).\(^2\) These are the local government bodies (150 of them in England, around 20 percent of which are in London) that typically manage school funding and admissions. We used this particular boundary feature in Gibbons and Machin (2003) and exploit it again here. But we also need a more general approach for the majority of cases when LEA boundaries are not relevant.

The porous nature of school admissions geography implies that residence-school distance should have important impacts on the premium home-buyers pay for residences at locations close to good schools. This issue is one that has received little attention [though see Des Rosiers et al (2001) for an exception], and is a central focus of this paper. We are also interested in how admission constraints impact on the valuation of primary school performance. Parents’ willingness to pay will clearly depend on whether there are constraints, or queues, to get their children into their preferred school. We develop further the standard hedonic valuation method [see Rosen (1974) for the classic exposition, or Sheppard (1999) for a modern survey] to provide different valuations for

---

\(^1\) Obtaining systematic well-defined information on catchment areas for Local Education Authorities (LEAs) is extremely difficult. However, it is clear that a great deal of heterogeneity exists across the country. In some LEAs well-defined catchment areas do exist, in others they change over time, and in others even defining a line around a catchment area proves difficult.

\(^2\) This boundary feature is much less applicable to secondary schools, where pupils are more likely to cross LEA borders.
parents in the presence and absence of school admission constraints. In the standard approach an estimate of the implicit price of school productivity is available from a simple regression of property prices on local school performance measures where one assumes school admissions are restricted to local residents. In our analysis we generalize this to allow for the fact that just living in the local neighbourhood of a particular school may not guarantee a child’s attendance at that school, because some schools may be over-subscribed.

The remainder of the paper is structured as follows. Section 2 is devoted to model development where we build models that show how admission constraints may impact on parental valuation of primary schools. We look at three models, respectively based upon supply constraints, school popularity and congestion effects, and generate empirical predictions from each. In Section 3 we describe the data and empirical methods we use. In Section 4 we present our estimates, and relate them back to the theoretical models presented earlier. Finally, Section 5 concludes.

2. ADMISSION CONSTRAINTS: DISTANCE TO SCHOOL AND SCHOOL CAPACITY

Distance to School and School Quality
The special features of the English schooling system mean that choosing a residence in a particular location does not guarantee admission to a school. Instead, probability of admission increases as distance of residential address from the school decreases. Transport costs also increase with distance. An obvious implication of both these factors is that parents’ willingness to pay for school quality through property prices at a given residential location must be decreasing in residence-school distance. Assume further that household residential choices are constrained to a general

---

3 A parallel issue is the effects of housing supply on the rate of capitalisation of local public goods. This is explored in Cheshire and Sheppard (2003). We focus here on constraints on the supply of the local public good itself.
locality by labour market and broader housing issues, so that parents are really interested in the performance of neighbourhood schools relative to what can be expected in the general locality. The school quality \( q \) offered by a particular residential location then depends on neighbourhood school \( i \)'s test-score performance, \( s \), defined relative to the mean in the locality as \( s = s_i - \bar{s} \). But it also depends on residence-school distance, \( d \). For the moment, assume this is the distance to the nearest school.\(^4\)

We define a school quality function \( q(s,d) \) to represent the effective supply of school quality available at a residential location at distance \( d \) from a school with performance \( s \). Some plausible restrictions are that \( q(s,0) = s \), \( q(s,\infty) = 0 \). This just says that school quality at the school gate is the same as that of the school itself, but school quality at a large distance is just average performance. If \( s > 0 \) then effective quality \( q \) decreases with distance: \( q_d(s,d) < 0 \). If \( s < 0 \) the inequality is reversed.

**Admission Constraints**

In this setting school admission constraints can act as an important feature in household choice. Class size, infrastructure, resource and institutional constraints mean that schools cannot expand indefinitely in response to high demand. Let us define two types of school: admissions unconstrained, in which pupil numbers do not generally exceed institutional capacity, so over-subscription criteria are irrelevant; and admission constrained, in which pupil numbers have reached or exceeded institutional capacity, so that the school's over-subscription criteria become important.

We propose three models, each with different predictions for the ways in which school constraints have some bearing on parental choice and thus on willingness to pay:

1). Supply Constraints

\(^4\) In the empirical section we work with harmonic mean distance to the nearest three primary schools.
Firstly, in a model we term the supply constraint model, school capacity may have an impact on the school quality available at a given school-distance because it limits the supply of school quality over distance. This happens if there is an increased probability of exclusion at any distance, due to more stringent proximity-based admissions criteria. Under this scenario, expected quality erodes rapidly with distance from over-capacity good schools (and improves rapidly with distance from bad schools).\(^5\) Let us add school capacity \(\sigma \in \{0,1\}\) as an argument to the school quality function, with 1 indicating an admissions-unconstrained school. Now, the absolute value of the slope of the distance decay function is always less for under-capacity schools:

\[
|q_d(s,d,1)| < |q_d(s,d,0)| \forall d .
\]

2). School Popularity

A second possible link between parental demand and school capacity rises where schools are likely to be admissions constrained because of earlier popularity. We term this the school popularity hypothesis. In this case, full-capacity can act as a signal of expected school quality that might not be fully revealed in school performance indicators based on pupil test scores. Or it can mislead parents into believing a school is successful, even if school performance is not particularly good. In this case, high-performing schools that are at or over-capacity are likely to push up house prices more than high-performing schools that are under-capacity. School admissions constraints need not have an impact on the way the school performance premium changes with distance, but are likely to affect prices directly, or influence the sensitivity of prices to performance at any given school-distance. In this case, school over-capacity signals popularity, which interacts positively with

\(^5\) In principle, quality decreases with distance from under-capacity good schools only because of transport costs, or because there is uncertainty about availability at the time households make their choices. In any case, there could be some institutional bias in favour of applicants from residences neighbouring a school, even in under-capacity schools. But, in general, if school capacity acts as a supply constraint, then expected school quality erodes with distance more rapidly for properties close to good over-capacity schools than it does for those close to good under-capacity schools.
performance in terms if perceptions of school quality, and willingness to pay. This implies that $q_s(s,d,1) < q_s(s,d,0)$ $\forall s,d$.

The supply-constraint and school popularity models are illustrated in Figure 1. The upper Figure is the supply-constraint model, where perceived school quality declines with distance from better performing schools, and increases with distance from worse performing schools. In this model the existence of admissions constraints acts to restrict the supply of school quality such that perceived school quality erodes more rapidly if the school is over-capacity than if under capacity (solid lines compared to dotted lines).

The lower Figure shows the differing predictions of the popularity hypothesis. Here, as before, one obtains a steeper gradient between perceived quality and school test-score performance for lower distance to school residences (solid lines compared to dotted lines). But now, admissions constraints act to increase the rate at which perceived quality increases with test-score performance, so the gradients are steeper for capacity-constrained schools. These Figures make it evident that there are testable empirical predictions that emerge from these two theoretical approaches, with different signed interactions emerging in school quality functions.

3). Congestion

In addition to supply constraints and school popularity, there is a third plausible link between over-capacity schools and prices. It may actually be that over-capacity schools suffer from problems associated with overcrowding, congestion and higher pupil-teacher ratios, and this may have direct effects on parental willingness to pay for schooling. In this case we would expect under-capacity schools to be in higher demand, implying higher local house prices nearer capacity unconstrained schools.
In this congestion model, the effect of this would be to shift the curves for over-capacity schools in Figure 1 downwards, relative to the curves for under-capacity schools, thus predicting the opposite from the supply constraint and school popularity models.

**Empirical Implications**

Clearly, it is not possible, *a priori*, to use theory alone to predict quite how performance, school distance and school admissions constraints will influence parental demand, so this is a question that must be addressed empirically. We assume that the school quality function \( q(s,d,\sigma) \) defined above is the object of parental preference over schooling choices, and that this can be estimated by revealed preference methods in a hedonic property value framework. Property prices trace out the function \( q(s,d,\sigma) \), once other factors that affect housing demand are held constant.

Note that our first supply-constraint hypothesis implies that the premium on home prices paid for school performance decays more rapidly for over-capacity schools than under-capacity schools. But importantly, for properties *very close* to schools there should be no difference in the performance premium between under-capacity and capacity-constrained schools. If this were not the case, rational home buyers would always choose properties close to under-capacity schools, where the marginal cost of performance is lower. Conversely, if the second hypothesis, school popularity, is correct, the price of performance is higher in over-capacity schools than in under-capacity schools, at any school-residence distance. Finally, the congestion model predicts local house prices near capacity constrained schools to be lower than those near schools with spare places.

---

6 This assumes that parents can choose from the full set of schools. In practice, such choices may be limited. Our data shows that the over-capacity schools are concentrated in the inner metropolitan areas, whereas under-capacity schools are more dispersed.
3. METHODS AND DATA

3.1. Data sources

Our data source for house prices and housing characteristics is the Nationwide Building Society’s survey, which is based on all property sales in Great Britain for which the building society makes mortgage loans. We have the data from 1995-2002. The size of the sample varies considerably from year to year, with the size of the market and the size of the lender’s share. We restrict attention to properties in the Metropolitan area of Greater London and its surroundings.

Our schools data comes from the Department of Education and Skills, and is made up from the Annual School Census from 1996-2001 and the publicly available school performance ‘league tables’ over the same period. School performance is measured in these tables as the proportion of children reaching target levels (Level 4) in the standard age-11 tests.7

We match each property to its nearest three primary schools (and three secondary schools), using Euclidian distances derived from property and school grid references. These grid references are assigned on the basis of mailing address postcodes8 using Ordnance Survey Codepoint™ data.9 We also include some area characteristics. Firstly, the Nationwide housing database provides indicators of neighbourhood conditions based on the Acorn classification system created by CACI marketing.10 In addition, we match in dwelling density estimates from the 2001 Census, based on address postcode. All school characteristics are matched to property transactions occurring in the

7 These tests are known as the ‘Key Stage 2’ tests, as they correspond to the end of Key Stage 2 in the National Curriculum for primary school children.
8 A residential postcode is, typically, shared by 10-15 houses on one side of a residential street.
9 ‘Code Point’ provides a National Grid reference for each unit postcode in Great Britain.
10 CACI is the UK subsidiary of CACI International Inc. and provides marketing and related information systems, including the customer profiling product known as ACORN. ACORN provides geodemographic classifications for postcodes throughout the UK. We utilise those codes that indicate postcode sectors (approx 3000 households) that contain high proportions of social housing.
following year, so our database of property transactions covers 5 years from 1997-2001, plus the first quarter of 2002.

The main sample characteristics are shown in Table 1. The average distance between primary schools and residences in our sample is just 560 metres. Average ‘league table’ performance is 68% over these years, meaning that 68% of children reached the target grade between 1996 and 2000. In fact this average masks substantial increases over time, with the proportion now averaging nearly 80%.

3.2. Matching schools to properties

Since a child at a given dwelling potentially has access to a number of schools, we need some way of linking schools to dwellings according to the probability of a child living in a particular place gaining admission. We do this on the basis of how close each school is to each property. For each property in our data set, we define a local school cluster as the set of three nearest primary schools. We define the school performance available to a particular dwelling $i$ at time $t$ as the inverse-distance weighted average of the performance of these nearest three schools $j$:

$$s_i = \frac{\sum_{j \in J} \tilde{d}_{ij}^{-1} \tilde{s}_j}{\sum_{j \in J} \tilde{d}_{ij}^{-1}}$$  \hspace{1cm} (1)$$

where $J$ is the set of nearest three schools to $i$, $\tilde{s}_j$ is the school-specific performance measure and $\tilde{d}_{ij}$ is the school-property specific distance. The effective school-distance for a given dwelling is calculated as the harmonic mean of the nearest three schools, which is the natural measure given the inverse-distance weighting. So the school-distance measure is

$$d_i = \left(\sum_{j \in J} \frac{1}{\tilde{d}_{ij}}\right)^{-1}$$  \hspace{1cm} (2)$$
where $J = 3$ in our case. Thus a property that is adjacent to one school, but further away from others, will be assigned the school performance of the adjacent school, and a distance of zero. A property that is equidistant from three schools is assigned the mean performance of the nearest three, and the common distance. Any dwellings that have a school-distance greater than 1km are dropped from our estimation sample, since they are unlikely to be relevant when considering school choice issues.\footnote{This loses 17\% of the total available sample. However, the results are similar, whether or not we restrict the sample in this way.} This is certainly true for primary schools, on which we focus here, although some children do travel long distances to attend secondary schools which, on average, are much larger in terms of student numbers.

### 3.3. School admissions constraints

As our theoretical discussion has emphasised, we are also interested in the effects of school capacity, as determined by the class-size constraints. The Annual School Census data has a notional capacity measure, which is determined by the class size restrictions imposed on primary schools (a maximum of thirty). For each school we can calculate the ratio of pupils to capacity. We assume that the school quality function $q(s,d,\sigma)$ at a particular location is admissions-unconstrained ($\sigma = 1$) if the inverse-distance weighted average of this ratio in the local school cluster is less than 1. We assume schools are admissions-constrained if the ratio is greater than or equal to 1.

### 3.4. Empirical model

Putting together the theoretical reasoning of Section 2 with the information described above into an empirical model of house price determination leads us to the following empirical specification:

$$
\ln p_t = q(s_{it-1},d_t,\sigma_{it-1};\gamma) + X_t^\prime \beta + \epsilon_t
$$

(3)
where $p_{it}$ is the price of property in postcode unit $i$ at time $t$, $q$ is postcode-unit-specific expected school quality which is a function of: $s_{it-1}$, the postcode-unit-specific measure of performance in the local school cluster in the previous period; $d_{i}$, the postcode-unit-specific school-distance defined above; and $\sigma_{s-1}$, the indicator of whether the local school cluster is capacity-constrained. In (3) $x_{i}$ is a vector of other observable characteristics, $\varepsilon_{s}$ is the usual random error term and we allow for unobserved spatial fixed effects, $f_{s}$, related to the school cluster $s$.

3.5. Area effects on school performance

Any study of the effects of schools on house prices must take account of general neighbourhood factors. Anything desirable about the neighbourhood – local amenities, community attributes, or housing quality – will drive up local prices. In a world of imperfect capital markets, the rich outbid the poor for desirable neighbourhoods [e.g. Benabou (1996), Epple and Romano (2000)]. If children from richer backgrounds do better at school, then observed school quality is better in richer neighbourhoods. School performance and house prices are simultaneously determined, and regression estimates that do not take this sorting into account are biased.

Work on school price effects has tried to get round this via a number of identification strategies. These include: specifying an extensive range of neighbourhood attributes in the property value regression [Downes and Zabel (2002)]; looking at differences between neighbouring properties in different school attendance districts [Black (1999), Gibbons and Machin (2003)], exploring what happens when school district boundaries are redrawn [Bogart and Cromwell (2001)]; or using semi-parametric methods to eliminate general spatial variation [Gibbons and Machin (2003)].

Here we use a combination of approaches, but our main device is to specify school-cluster fixed effects in our regressions. Persistent effects common to all properties in the school cluster can be accounted for in our property price regressions by a standard fixed-effects strategy, using
observations in the same 3-school-cluster to calculate the fixed effects. An observation is assigned
to a school group with other observations that share the same nearest, second nearest and third
nearest primary school. Observations in the same group can be property transactions from the same
postcode in different periods, or from different postcodes in the same period.

This school-cluster fixed effect approach is illustrated in Figure 2, which maps three schools
(indicated by Δ) in Haringey, North London and the corresponding residential postcode centroids in
the related schools clusters. The postcode centroids are shaded to indicate the school cluster to
which they belong. Dwellings in Cluster 1 have the upper right-hand school as the nearest, the
lower left-hand school as the second nearest and the lower right-hand school as the third nearest.
Cluster 2 has the lower left-hand school as the nearest, the upper right hand school as the second
nearest, and the lower right-hand school as the third nearest. It is left to the reader to deduce the
rest. Each cluster is treated as a fixed effect in our regressions.

The point of this strategy is to remove as much cross-sectional variation as possible\textsuperscript{12}, whilst
retaining the scope for measuring residence-school distance effects. Indeed much of the
identification in our model comes from differences in the time trends of school-cluster performance.
But not all: Some cross-sectional variation remains between properties within 3-school-clusters.
This is because each residence has a unique residence-school distance vector, and so offers a unique
supply of school-quality within the cluster.\textsuperscript{13} Residence-school distance effects are identified here
because each 3-school cluster has a different mix of properties in each year, so the mean residence-
school distance is not constant over time.

\textsuperscript{12} An alternative approach would be to use repeated observations of sales in the same postcode unit. This
would result in a serious loss of data and information, since the sample of repeated postcodes is quite small,
and would make it impossible for us to measure distance effects.

\textsuperscript{13} In practice we can eliminate this variation without much effect on the results, by using simple means rather
than spatially weighted means.
Our second approach for eliminating area effects is to transform (3) into a spatially differenced model that eliminates the area fixed effects. For any characteristic $x_i$ associated with dwelling $i$, we calculate the characteristic $x_j$ of a geographically ‘neighbouring’ dwelling $j$, and work with the spatially differenced variable $x_i - x_j$ in our regressions. This eliminates any neighbourhood effects that are common to both dwellings. But the transformation will also eliminate any differences in school performance unless we make efforts to ensure that our neighbouring dwellings have access to different schools. If we had clearly defined neighbourhood catchment area boundaries, we could focus on dwellings that are neighbouring, but in different catchment areas [as in Black (1999)]. Since the only ‘catchment area’ boundaries that can be reliably observed are the boundaries corresponding to the districts for the admissions authorities – the Local Education Authorities in England – we use differences between matched pairs of dwellings that are ‘neighbouring’, but in different LEAs. Each dwelling is matched to its nearest neighbour in an adjacent LEA and we restrict the sample to ensure that all neighbours are within some specified distance of each other.\(^{14}\) Figure 3 illustrates the London sample of dwellings within 1km of the nearest dwelling across the LEA border. Estimation is based on cross-LEA border differences in prices and characteristics within these thin boundary zones.

4. RESULTS

4.1. Performance, distance, and prices

Our empirical results are derived from statistical regression estimates of the house price model in equation (3) under alternative specifications. The main focus is on the estimation of the

\(^{14}\) We match by distance, then by time. For a property $i$ in year $t$ we match the nearest property $j$ that is available in the same year. If a nearer property $k$ is available in the previous year, $t-1$, we replace the match
function \( q(s,d,\sigma) \), which we parameterise as a simple linear function of \( s, d, \) and \( \sigma \) with interaction terms, and we make strong efforts to deal with the potential endogeneity of school performance using the strategies laid out in Section 3.5.

The central results relating to our hypotheses about distance and admissions constraints are in Table 2. We look first at distance and performance effects only, so in Columns (a)-(d) the school quality function \( q(s,d,\sigma) \) is parameterised as a function of distance, performance and a distance-performance interaction. Column (a) is a property value model with a full set of property controls, area characteristics (as specified in the Appendix), time effects, and Local Education Authority dummy variables. Column (b) introduces school-cluster fixed effects in the regressions to account for unobserved neighbourhood differences. Columns (c) and (d) are instrumental variables estimates, respectively without and with school cluster fixed effects. In all cases, the sample is restricted to properties that are within 1km of a primary school, since we regard households living outside this range as unconcerned with primary school performance issues.\(^{15}\)

The basic effect of school performance is in Row 1. This should be interpreted as the effect of school performance on a property at the school gate (zero distance). Without school cluster effects (Column (a)) the impact on house prices is large – up to 4.8% for a 10 percentage point shift in school performance.\(^{16}\) This is close to the largest results we reported for English primary schools in Gibbons and Machin (2003) and is similar to Black’s (1999) analysis of US elementary schools. However, controlling for unobserved neighbourhood factors through school-cluster fixed effects in Column (b) more than halves the effect. Now, the main effect of performance is 2.1% on prices for

\(^{15}\) It would be a different story for secondary schools where children regularly travel further distances to school.

\(^{16}\) Calculated as \( \exp[(.047) - 1)] \times 100.\)
each 10 percentage-point improvement in the proportion of children reaching the target test grade. This is in line with the lowest results in our previous work. We consider the lower figure to be a more robust estimate of the parameter, and that the Column (a) estimate without school-cluster fixed effects is upward biased by unobserved neighbourhood effects.

Consider now the effect of school distance and its interaction with performance. Again, the estimates without school-cluster fixed effects in Column (a) are questionable. Firstly, the main effect of school distance is positive (Row 3) and the coefficient on the interaction of distance and performance is low and not statistically significant, which implies: i) that the performance in the nearest-school cluster has a constant effect on the price of neighbouring properties, regardless of their distance to the school and ii) that house prices decrease with distance, even from the best schools. This is hard to square with our theoretical discussion in Section 2, or with anecdotal evidence.

Using the school-cluster fixed effects model in Column (b) we obtain more convincing results. Again, the main effect of distance from low-performing schools is positive: prices rise by 2.1% per 100m distance from a school with an implied zero performance score. However, the significant distance-performance interaction term in Row 5 implies that this negative school impact is ameliorated by school performance: for schools with 83% and more of their pupils attaining the target grade. After this, the distance effect switches sign. It is therefore evident that people pay to move close to schools at the top of the attainment distribution. The interaction term also means that distance from a school reduces the school performance premium. As we would expect, the performance of nearest schools matters less for prices of properties that are furthest away, each 100

\[ \exp(0.020-(0.83\times0.024))-1=0 \]

---

17 The main effect of distance is 0.020 and coefficient on the distance performance premium is 0.024, so the impact of distance for a school with a performance measure of 0.83 is \( \exp(0.020-(0.83\times0.024))-1=0 \)
metres reducing the house price effect of a 10 percentage point shift in school performance by around 0.24%.

In the Column (d) specification, instrumenting school performance with salient school characteristics results in bigger coefficients on performance, but the general pattern of interactions is unchanged and re-confirms the findings. Increased coefficient estimates using Instrumental Variable methods is what we observed in our earlier work [Gibbons and Machin (2003)]. Our interpretation then, and now, is that year-on-year performance measures are less than perfect indicators of perceived long-run school performance and introduce a form of measurement error. Instrumenting with more stable school characteristics corrects for this attenuation bias due to measurement error.

4.2. …with admissions constraints

In Columns (e) and (f) we incorporate admissions constraints into the statistical within-cluster models. In (e) we include a dummy variable indicating admissions unconstrained schools and in (f) we interact this with the distance and performance variables. We structure the results in this way to first ascertain the relationship between prices and admissions constraints (predicted to be positive in the supply constraints and school popularity models, and negative in the congestion models) and then to consider the fully interactive model that enables us to distinguish between the competing models of primary school choice introduced in Section 2 above.

In the Column (e) specification the estimated coefficient on the under capacity dummy variable is negative and statistically significant, showing house prices to be higher in the proximity

---

18 The coefficient on the distance-performance interaction term is -0.024, so the magnitude of the effect on a 10 percentage point performance advantage is exp(-0.024*0.10)–1=–0.0024 or –0.24%.

19 As instruments we use indicators of the institutional age range (with nursery, junior years only), church-school status, and ‘beacon’ school status, all interacted with distance. A ‘beacon’ school is a school designated by the Department of Education and Skills as exhibiting high teaching standards and models of good practice.
of schools with admissions constraints. So, in line with the supply constraint and school popularity models, but in direct contrast to the congestion model, there is evidence that queues for primary school places do generate a premium on property prices. Interestingly, when compared to the column (b) specification the other estimated coefficients do not shift much on the inclusion of the capacity variable.

However, to properly consider which of the theoretical models is best in line with the data we need to estimate the interactive model including interactions between capacity, school performance and distance. These estimates are shown in Column (f) where there are significant interactions between admissions constraints and the effects of performance and distance. Since the coefficients are not straightforward to interpret, and to link back to the theoretical discussion, we illustrate our estimates diagrammatically in Figures 3a and 3b. These Figures are based on the significant coefficients in Column (f) and the price is expressed in terms of proportional differences from the average admissions-unconstrained school.

Figure 3a exhibits the hedonic price function of distance, by six, school-performance and admissions-constraint categories:

i) Mean performing schools (68%), admissions constrained (Mean_over)
ii) Mean performing schools (68%), admissions unconstrained (Mean_under)
iii) High performing schools (100%), admissions constrained (Top_over)
iv) High performing schools (100%), admissions unconstrained (Top_under)
v) Low-performance schools (36%), admissions constrained (Low_over)
vi) Low-performance schools (36%), admissions unconstrained (Low_under)

---

20 Viable instruments are not available for all the performance-distance and performance-capacity interactions in these models. So we simply register that, in the light of our earlier IV estimates, the performance effects and their interactions with distance and capacity may be under-estimated.
The important thing to note here is that the interactions in Table 2 Column (f), between the capacity and distance, or between capacity, distance and performance are weak, implying little or no impact from admissions-constraints on distance effects. However, as we observed before, the interactions between distance and performance are strong (Table 2, Row 2, column (f)). The clear evidence looking at Figure 3a is that schools of low and average performance are not regarded as positive neighbourhood amenities: prices actually rise as we move away from such schools. Only the best schools (the top 20%) tend to significantly pull up prices within their immediate vicinity.

Although admissions constraints have no impact on the rate at which prices change with distance, they do change the implicit price of performance at any distance. This is seen more clearly in Figure 3b which shows the hedonic price function of performance, for four school types:

i) Admissions constrained, and immediately adjacent to the dwelling (0m_over)

ii) Admissions unconstrained, and immediately adjacent to the dwelling (0m_under)

iii) Admissions constrained at mean residence-school distance of 560m (Mean_over)

iv) Admissions unconstrained at mean residence-school distance of 560m (Mean_under)

Firstly, observe that performance in admissions-unconstrained schools attracts a much lower premium than performance in admissions-constrained schools (from Table 2, Row 2, Column (f)). In fact, the premium for a 10 percentage-point improvement in school quality in admissions-constrained schools at zero distance is 50% higher than that for under-capacity schools: 2.4% as against 1.6%. So, in Figure 3b, the slope of the performance-price function for admissions-unconstrained schools in Figure 3b is less steep (0m_under compared to 0m_over). Moving from zero-distance to mean-distance schools, the performance premium declines, but the differential between admissions-constrained and unconstrained schools persists (Mean_under compared to Mean_over). Again, as we observed in Figure 3a, households pay lower prices for properties close to primary schools, unless schools succeed in getting well over 80% of their pupils up to the target grades.
4.3. **Local cross-boundary models**

The school-cluster fixed effect approach provides a strong test of the existence of school effects on prices, but does not guarantee that dwellings are properly matched to schools that are accessible only to residents in those dwellings. For this we need clearly defined mutually exclusive catchment areas. As discussed in earlier sections, we can use Local Education Authority boundaries to define catchment area boundaries. In this way we can use cross-border differences in performance and prices to identify the effects of access to different schools in a similar way to Black (1999), Bogart and Cromwell (1997, 2001), and Gibbons and Machin (2003). Table 3 presents our estimates from regressions based on this approach.

These estimates are based on differences in prices and characteristics between nearest-neighbour dwellings on either side of an LEA boundary. Other than focussing on this specific sample, the specifications are identical to comparable ones in Table 2. Columns (a) and (b) restricts the sample to properties within 1km of their matched neighbour on the other side of the LEA boundary. Columns (c) and (d) are based on properties within 750m. In each case the first column of the pair looks at performance and distance effects only. The second column introduces information on admissions constraints.

Broadly speaking the results are closely in line with this using the school-cluster fixed effects in Table 2. School performance exerts a positive impact on prices that significantly decreases with distance from the school. Prices rise with distance from the school at lower performance levels, but the price-distance gradient decreases to zero for the highest performing schools. Admissions constraints interact with performance, but not with distance in the same way as in Table 2. The results are quite robust to changes in the way the sample is restricted in terms of inter-property distances as we move across the Table.
4.4. *Interpretation and Discussion*

The reported results reveal a number of features of the house price-school performance relation. First, in line with other studies [like Black (2003) and Gibbons and Machin (2003) for primary schools, and Leech and Campos (2003) and Rosenthal (2003) for secondary schools] higher test-score based school performance is associated with increased property prices. However, it is only the best one-in-five schools that generate higher than average prices close-by. Second, there are important interactions between performance and residence-school distance and school capacity that induce systematic variations in the house price premium associated with better neighbourhood primary schools. Third, the nature of these interactions can enable us to discriminate between different theoretical reasons as to why parents are prepared to pay higher prices for living in an area that will enable to get their children admitted to a particular school.

So how are we to interpret the results in the light of our earlier theoretical discussion? The first point is that there is little evidence for the *congestion* model in which under-capacity schools are in higher demand, implying higher local house prices nearer capacity unconstrained schools. In actual fact, the under-capacity variable is negatively related to house prices in the statistical models and the only place where even an inkling of congestion effects can be seen is from the fully interacted model at the very bottom of the school performance distribution. In Figure 4 (b) one can see this for lower levels of performance where the dotted over and under capacity lines cross. The crossover-point on the graph in Figure 4b is at Key Stage 2 performance level of 0.48, which is at the 10th percentile of the KS2 distribution and even here the lines are not very far apart at all. So it seems we can rule out the possibility of congestion effects in all but the worst-performing schools.

This leaves the two hypotheses which assert that there is a higher house price premium for admission constrained schools, but which differ in the way that performance interacts with distance and capacity constraints. There are several points to note here. First, the performance premium is higher for admissions-constrained schools even at zero distance to the school. Second, note also that
admissions-constraints have no impact on the rate at which prices change with school distance. Thus, in line with the discussion in Section 2, the results are supportive of our popularity hypothesis – there are complementarities between the signal provided by a school being full and the performance of the school measured in standard pupil tests. We find no evidence that our supply-constraint hypothesis is correct: there is no premium to be paid for moving closer to an admissions-constrained school. This comparison is most easily made by looking at the Figures which are the empirical analogues to the theoretical models presented in Section 2. Comparing Figure 3a in conjunction with Figure 1a, and Figure 3b in conjunction with Figure 1b, makes it evident that the evidence is in line with the predictions of the school popularity model, rather than the model based upon supply constraints.

5. CONCLUSIONS

This paper presents new estimates on the value of primary schooling for London and the South East of England, based on property value models. We develop methodologies that are appropriate for the English setting where admission catchment areas are fuzzy and porous, and which take careful account of unobserved neighbourhood effects. Our results are largely based on changes in performance of schools over time, but also on highly localised variation in school performance within micro-geographic neighbourhoods. Our most conservative results indicate a premium of around 2% on prices for each 10 percentage point increase in the proportion of children reaching target grades in age-11 tests. This premium relates to a property notionally located outside the school gate.

Our first improvement on previous work is to present plausible estimates of the effects of school distance on this performance premium. Each 100m distance to a school erodes the performance premium by between around 0.2%, implying zero impact at about 1km. We have also
shown that all but the top 1-in-5 in the school performance league tables depress prices in their immediate vicinity. Average schools are not desirable local amenities. This may, in part, be explained by ‘flight’ from the worst schools, but environmental problems also probably contribute. The morning and evening ‘school-run’ brings traffic and congestion, and there may be additional nuisances such as playground noise that deter buyers. These environmental effects and their negative impacts on prices have been noted and estimated before [Hendon (1973)].

We expected admissions constraints to have important effects on parental decisions over schooling, and indeed there are significant interactions between measures of school over-capacity and the premium paid for high performing schools. Our first supply-constraint hypothesis was that households will pay more to live ever closer to high-performing, over-capacity schools than under-capacity schools because school places are tightly rationed. This is not borne out by our data. Nor is the notion that congestion effects make over-subscribed schools less attractive to parents. A school that is admissions-constrained is worth more than one that is not, throughout most of the performance range and this impact is not sensitive to distance.

Instead, the empirical results are consistent with our popularity hypothesis in which parents believe, rightly or wrongly, that popular schools are better than under-capacity schools. There seems to be a degree of ‘herd’ behaviour in primary school choice, with the price not perfectly reflecting the fundamentals. Parents certainly do pay to get their children into better performing primary schools, but it is evident that they prefer popular, over-subscribed schools. This seems to be the case even if their league-table results may not be up-to-scratch at the time admissions applications are made by parents.
6. REFERENCES


**Supply constraint hypothesis:**
Admissions constraints constrain the supply of school quality

Perceived school quality $q$ declines with residence-school distance ($d$) for good schools ($s_1$) and increases with distance from bad schools ($s_0$). With admissions constraints ($\sigma = 0$) the rate at which perceived quality decreases (or increases) with distance is higher than without ($\sigma = 1$). The probability of admission decreases more rapidly with distance than in under-capacity schools.

**Popularity hypothesis:** Admissions constraints signal popularity and school quality

Perceived school quality $q$ increases with test score performance ($s$). The slope is higher for low residence-school distances ($d_1$) relative to high distances ($d_0$). Admissions constraints ($\sigma = 0$) increase the rate at which perceived quality increases with test-score performance.

**FIGURE 1: CONJECTURES ABOUT THE RELATIONSHIP BETWEEN PERFORMANCE, DISTANCE AND ADMISSIONS CONSTRAINTS**
FIGURE 2: EXAMPLE OF DWELLINGS AND SCHOOL CLUSTERS IN HARINGEY, NORTH LONDON

Triangles indicate primary school locations. Squares indicate residential postcode centroids in the sample. The depth of shading indicates the school cluster to which the postcode belongs.

- Cluster 1
- Cluster 2
- Cluster 3
- Cluster 4
- Cluster 5
(a) School distance effects on property price for low, mean and top performing schools, by admissions constrained status

(b) School performance effects on property price for properties at zero and mean school distance, by admissions constrained status

FIGURE 4: ESTIMATED SCHOOL-PRICE FUNCTIONS
### TABLE 1: KEY SUMMARY STATISTICS

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-property price</td>
<td>11.570</td>
<td>0.505</td>
</tr>
<tr>
<td>Primary school performance</td>
<td>0.680</td>
<td>0.138</td>
</tr>
<tr>
<td>Primary school distance (100m)</td>
<td>5.592</td>
<td>2.136</td>
</tr>
<tr>
<td>Admissions constrained schools (proportion)</td>
<td>0.546</td>
<td>0.500</td>
</tr>
</tbody>
</table>

Notes: Properties with prices above 1million are excluded from the Nationwide sample
Performance is the proportion reaching level 4 in Key Stage 2 tests at age 10/11
Sample is for London and outer metropolitan area, restricted to properties with school-distance less than 1km
## Table 2: Performance, Distance and Capacity Interactions

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>Within cluster</th>
<th>IV</th>
<th>IV-Within cluster</th>
<th>Within cluster</th>
<th>Within cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(a)</td>
<td>(b)</td>
<td>(c)</td>
<td>(d)</td>
<td>(e)</td>
<td>(f)</td>
</tr>
<tr>
<td>School performance * under-capacity</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-6.774 (-2.83)</td>
</tr>
<tr>
<td>Distance (100s m)</td>
<td>1.215 (8.06)</td>
<td>2.000 (12.71)</td>
<td>3.727 (1.81)</td>
<td>2.068 (3.66)</td>
<td>2.000 (12.69)</td>
<td>1.961 (8.31)</td>
</tr>
<tr>
<td>Distance * under-capacity</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.084 (-0.28)</td>
</tr>
<tr>
<td>Distance * performance, *under-capacity</td>
<td>-1.819 (-0.43)</td>
<td>-2.408 (-11.03)</td>
<td>-1.882 (-1.32)</td>
<td>-3.386 (-3.23)</td>
<td>-2.403 (-11.00)</td>
<td>-2.488 (-7.66)</td>
</tr>
<tr>
<td>Under capacity</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.465</td>
<td>3.270 (1.94)</td>
</tr>
<tr>
<td>Area effects</td>
<td>52 LEA school</td>
<td>14382 school</td>
<td>52 LEA school</td>
<td>14382 school</td>
<td>14382 school</td>
<td>14382 school</td>
</tr>
<tr>
<td>Sample size</td>
<td>109136</td>
<td>109136</td>
<td>109136</td>
<td>109136</td>
<td>109136</td>
<td>109136</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.818</td>
<td>0.833</td>
<td>0.812</td>
<td>0.833</td>
<td>0.833</td>
<td>0.833</td>
</tr>
</tbody>
</table>

Notes: dependent variable is log property price
Sample restricted to properties with less than 1km school distance
Coefficients are x100
Instruments for performance are beacon school status, church school status, age-range (nursery, junior only) and their interactions with school-distance
<table>
<thead>
<tr>
<th>Maximum distance between neighbours</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1km</td>
<td>27.068</td>
<td>22.512</td>
<td>25.812</td>
<td>20.862</td>
</tr>
<tr>
<td></td>
<td>(11.38)</td>
<td>(8.08)</td>
<td>(10.81)</td>
<td>(7.45)</td>
</tr>
<tr>
<td>750m</td>
<td>-</td>
<td>-</td>
<td>-7.712</td>
<td>-4.673</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>(-2.46)</td>
<td>(-1.27)</td>
</tr>
<tr>
<td>School performance</td>
<td>1.566</td>
<td>1.422</td>
<td>1.540</td>
<td>1.378</td>
</tr>
<tr>
<td>(100s m)</td>
<td>(7.34)</td>
<td>(5.77)</td>
<td>(7.22)</td>
<td>(5.59)</td>
</tr>
<tr>
<td>Distance * under-capacity</td>
<td>-</td>
<td>-</td>
<td>-0.042</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>(-0.14)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>Distance * performance,</td>
<td>-1.582</td>
<td>-1.530</td>
<td>-1.534</td>
<td>-1.439</td>
</tr>
<tr>
<td></td>
<td>(-5.13)</td>
<td>(-4.27)</td>
<td>(-4.98)</td>
<td>(-4.01)</td>
</tr>
<tr>
<td>Distance * performance, *under-</td>
<td>-</td>
<td>-</td>
<td>0.216</td>
<td>0.268</td>
</tr>
<tr>
<td>capacity</td>
<td>-</td>
<td>-</td>
<td>(0.50)</td>
<td>(0.54)</td>
</tr>
<tr>
<td>Mean neighbour distance</td>
<td>555m</td>
<td>431m</td>
<td>555m</td>
<td>431m</td>
</tr>
<tr>
<td>Sample size</td>
<td>26790</td>
<td>19270</td>
<td>26790</td>
<td>19270</td>
</tr>
<tr>
<td>R²</td>
<td>0.783</td>
<td>0.785</td>
<td>0.783</td>
<td>0.785</td>
</tr>
</tbody>
</table>

Notes: dependent variable is log property price
Coefficients are x100
Estimates based on differences between neighbouring properties across LEA boundaries. Sample is restricted to differences between neighbouring properties, where the difference in school-distance is less than 1km.
8. APPENDIX

List of control variables used in the house price regression models

School characteristics
Inverse distance weighted proportion of non-ethnic minority children in nearest 3 primary schools
Inverse distance weighted proportion with statements of special educational needs in nearest 3 primary schools
Inverse distance weighted proportion with special educational needs (not 'stated') in nearest 3 primary schools
Inverse distance weighted average number of full-time children in nearest 3 primary schools
Inverse distance weighted average proportion with 5GCSEs Grade A-C in nearest 3 private schools
Inverse distance weighted average proportion with 5GCSEs Grade A-C in nearest 3 state secondary schools

Dwelling characteristics
Number of bedrooms
Number of bathrooms
Floor area in metres-squared
Building age in years
Property type (semi-detached) Detached, Terraced, Flat/Maisonette, Other
Central heating (full), part, none
Garage (none), single, double, parking space
New/second-hand building indicator
Freehold/leasehold

Area characteristics
Acorn group indicators for postcode sector populations that are predominantly council estate residents: better off, unemployed, greatest hardship
Distance to nearest postcode represented in the sample (decile dummies)
Census 2001 Output Area: number of dwellings per kilometre-squared (decile dummies)
Local Education Authority dummies (not in school-cluster fixed effect models)

Time periods
Year dummies
Quarter dummies