Consumer Response to School Quality Information: Evidence From the Housing Market and Parents' School Choices

Iftikhar Hussain

University of Sussex

Draft

4 May 2016

ABSTRACT

This paper provides evidence on the impact of short-term innovations in school quality on both hedonic prices as well as parents' school choices. Employing a new measure of school quality, inspection ratings, I first demonstrate that changes in these ratings are related to changes in underlying school productivity. The study exploits a novel empirical strategy, which relies on the exogenous temporal variation in the release of the ratings, to overcome a number of key limitations in the prior literature on the capitalization effects of school quality. I find robust evidence of the impact of changes in ratings on house prices as well as parents' school choice decisions. Investigating heterogeneous effects reveals the importance of analyzing response in the housing market and parental decision making simultaneously.

JEL: H4, I20, R21

Key words: Housing Prices; Hedonic Regressions; School Quality; School Evaluation; Inspections

1. Introduction

Hedonic pricing models of the housing market can be a powerful tool in assessing the implicit price of local public goods and amenities. Examples from the empirical literature on the capitalization effect of such goods and bads includes analysis of school quality (Black, 1999; Figlio-Lucas, 2004; Bayer et al., 2007) environmental cleanup (Chay and Greenstone, 2005; Greenstone and Gallagher, 2008) and cancer risk (Davis, 2004).

This paper investigates the housing market capitalization effect of a novel measure of school quality, school inspection ratings. In particular, the focus is on the housing market reaction to a *change* in the inspection rating. I demonstrate that changes in inspection ratings are related to changes in the school's underlying performance – as measured by both test scores in levels as well as value added – and unrelated to changes in observable measures of the socioeconomic makeup of the student body.[1]

This study makes two key contributions to our understanding of education markets. First, as explained in greater detail below, the identification strategy as well as the measure of school quality employed offer some distinct advantages over the prior literature. Second, as well as estimating the implicit price for school quality, this study also incorporates analysis of parents' *school choice* response to ratings. In many settings, including the present one, parents have some degree of choice regarding which public school their child attends, even conditional on their location.[2] If parents switch their school choices in favor of higher rated schools, then the housing market response may be a lower bound estimate of the true willingness to pay for higher school quality. I empirically test for the significance of this mechanism.

[1] On the importance of school and teacher value added see, for example, Deming (2014) and Chetty et al (2014).
[2] This is in contrast to the Tiebout (1956) hypothesis which highlights consumers 'voting with their feet' by moving to their preferred neighborhood in order to consume their optimal level of the public good.

In particular, using parents' rankings of local schools from administrative applications data, I estimate a simple discrete choice model in order to assess the strength of preference for school quality, conditional on the family's location.[3] Arguably this approach, which takes the home location as given, is appropriate in a setting where either the change in school quality is marginal (so that transaction costs imply that families are unlikely to move home, but they may switch their school choice) or where the focus is on the immediate or very short term effect of the information treatment. Related to this, I also investigate the impact of ratings on total school enrolment.

The setting for this study is the English public (state) schooling sector, where schools are subject to inspections every few years by external evaluators or inspectors. Inspectors visit each school at very short notice, observe lessons, interview school leaders as well as parents and assess students' written work.[4] At the end of the inspection they provide an overall grade – on a four point scale ranging from 'outstanding' to 'fail' – and write a report which is made available on the Internet.[5] As documented in Hussain (2015), the inspection body is a highly independent institution and the system does not suffer from the phenomenon of 'ratings compression' typical of many subjective evaluation systems in both the private as well as public sectors (refXX). One reason why we might expect to see a market response to a change in a school's rating is that they usefully summarize changes in test score performance that are not readily discernable to consumers. Evidence from the literature suggests that even when test score and value added performance data are publicly available, as is the case in the current

[3] The details of the school assignment mechanism are discussed below.

[4] Guidelines published by the inspecting body, Ofsted, suggest that inspectors care about both hard data – test scores in levels as well as value added – as well as the softer measures of school quality gathered during the inspection visit. The precise weighting scheme is not made explicit.

[5] In the analysis below I undertake a number of exercises in order to shed light on the information content of the ratings produced by the inspectors.

setting, parents are unable to distinguish signal from noise in short-term changes in these measures.[6]

The empirical strategy employed in this paper relies on the exogenous timing of inspections: inspections take place through the year and I demonstrate that timing is unrelated to a school's performance. I use this timing policy rule to identify the immediate impact of the inspection ratings. In addition, schools are not inspected every year, whereas students are tested, and the results disclosed, on an annual basis. This temporal variation in the availability of different types of information allows me to carry out falsification exercises in order to rule out alternative explanations for my main findings.

This empirical approach offers a number of advantages over the prior literature. In particular, any estimated effect is uncontaminated by 'sorting bias' associated with alternative empirical research designs such as the boundary fixed effect approach (see for example, Black, 1999, and Bayer et al., 2007; see also Meghir and Rivkin, 2011, for a discussion of this literature).[7] This study focuses on the impact immediately after the school quality information is revealed, so that there is little opportunity for communities to re-sort and hence this channel is shut down. In addition, because the changes in school quality are marginal ones, it can be credibly argued that the hedonic price function does not shift in response to the treatment.[8]

Second, by focusing on changes in ratings, which are correlated with changes in the school's test score and value added but not with its changing demographics, the analysis presented here does not conflate peer quality and school productivity. This is in contrast to the

[6] See for example, Kane, Staiger and Samms (2003). A second reason why ratings may have an impact on the housing market is that they may contain information about school quality that is not readily available in the public realm. For example, inspectors asses the quality of lessons being provided. These aspects of school quality may have an impact on aspects of human capital acquisition not readily captured by test scores, such as non-cognitive skills (Heckman, ..).

[7] For studies outside the US, see for example, Fack and Grenet (2009), Fiva and Kirkboen (2011), Gibbons and Machin (2003) and Gibbons et al. (2013).

[8] Card et al (2008) provide evidence that even small changes can lead to tipping effects. However these are likely to evolve over time and so be less of a concern in the current research design.

vast majority of the school quality literature which largely focuses on schools' long-term test scores in levels, making it difficult to assess to what extent parents value peers rather than underlying school productivity.[9]

Finally, the research design allows for a simple test of heterogeneity in marginal willingness to pay for school quality. This is an important issue in theoretical models of neighborhood stratification (e.g. Ellickson, 1971) which employ the 'single crossing' assumption which implies that that richer households have a higher marginal willingness to pay for local public goods.

Summarizing the main findings of the study, the evidence shows that a unit increase in the rating leads to a rise of half of one percent in local property prices.[10] This result is precisely estimated and is robust to a variety of specification tests. Although seemingly small, the fact that there is any market reaction at all is interesting given that first, there is also an enrolment response (see below) and second, changes in inspection ratings are signals of short term innovations in quality, which may be reversed in the next inspection round (approximately three years later). The evidence also suggests that these effects persist until the next inspection, three or more years later.

Investigating heterogeneity in marginal willingness to pay for school quality reveals that the effect for properties located near schools serving low proportions of free lunch students is 1.5 percent for each unit change in the rating; for properties located near schools serving very high proportions of free lunch students the effect is close to zero. This stark finding *could* be interpreted to imply either that poorer families are insensitive to marginal changes in school

[9] Researchers have also investigated the capitalization effect of value added test scores, which are arguably less correlated with student quality, although the evidence of any significant and lasting effects is mixed, see for example, Imberman and Lovenheim (2014). Studies which have assessed the housing market impact of school characteristics other than test scores include Cellini et al. (2010) who estimate the impact of school facility investments. For evidence on the importance of teacher characteristics valued by parents, see Jacob and Lefgren (2007).

[10] Ratings range from 1 to 3 (fail category schools are excluded), so the changes in ratings range from -2 to +2. Properties lie inside a 500m radius around the school.

quality or that they are unaware of these changes. However, the school choice and enrolment analysis below suggests an alternative explanation.

The first notable finding from the analysis of parents' ranked preferences data, which exploits the variation in ratings over time in a discrete choice framework, is that a rise in a school's rating has a statistically and economically significant effect on the probability of choosing that school. This revealed preference can also be assessed via the tradeoff families are willing to make with respect to extra travel distance. This shows that parents are willing to travel approximately 7 percent farther for a unit rise in the rating. Analyzing heterogeneity in treatment effect by poverty status shows that the probability of choosing a school improving its rating is *at least as high* for poor families as it is for non-poor families.[11]

The results for the school's total enrolment outcome mirror these finding. This analysis finds robust evidence demonstrating that schools expand and contract in response to better and worse ratings, respectively. In addition, enrolment rises more for schools serving less advantaged students.

The analysis of parents' preferences data and the school enrolment outcome has clear implications for the hedonic analysis: the estimated housing market capitalization likely understate the true implicit prices for improvements in school quality. Second, this evidence also suggests an alternative explanation for the earlier finding that there is a negligible price impact for housing units located near schools serving the most disadvantaged students: if there is greater excess capacity at these types of schools, then any change in demand may be felt through school switching (at application) rather than through the hedonic price mechanism.

This study builds on and extends the related literature in a number of directions. In particular, this is the first study to combine an analysis of the housing market capitalization

[11] Although I cannot shed light on *why* disadvantaged families would respond more, one hypothesis may be that these families are initially less well informed and hence the signal from the ratings shifts the priors for these families to a greater extent than for more advantaged families.

effect of school quality with an analysis of parents' school choice decisions. Examples from the literature include Figlio and Lucas (2004), who provide evidence on the impact of disclosing test-based school report cards in the housing market, and Hastings and Weinstein (2008), who shed light on the impact of experimentally varying schools' test score information on parents' school choices.[12] Another contribution of this study is its focus on a measure of quality which is better aligned with school quality than test scores in levels, the typical focus of most studies in both the hedonics literature as well as the school choice literature.[13]

The research design employed in this study also offers a number of improvements on the prior literature. By focusing on the immediate impact of marginal changes in school quality, this study is less prone to the critique that exploiting large variations in local amenities over long periods of time leads to shifts in the hedonic price function, making any housing market capitalization effect difficult to interpret (Kuminoff and Pope, 2014).[14]

In addition, although extensive evidence suggests that the housing market is sensitive to long run differences in quality among schools (see the survey by Black and Machin, 2011), there is scant evidence on the dynamic effects of school quality on house prices. For example, Kane, Staiger and Samms (2003) find no impact of variation in test scores over time on house prices. This finding may appear to be puzzling given that school performance is unlikely to be fixed over time. This paper seeks to address this issue and fill this gap in the literature.

---

[12] There is also a small literature demonstrating that the housing market response to school quality measures is dampened when there is greater school choice (e.g. Brunner at al., 2012, Fack and Grenet, 2010, and Reback, 2005).

[13] Researchers have also investigated the capitalization effect of value added test scores, which are arguably less correlated with student quality, although the evidence of any significant and lasting effects is mixed, see for example, Fiva and Kirkboen (2011) and Imberman and Lovenheim (2014). Studies which have assessed the housing market impact of school characteristics other than test scores include Cellini et al. (2010) who estimate the impact of school facility investments. For evidence on the importance of teacher characteristics valued by parents, see Jacob and Lefgren (2007).

[14] Furthermore, with respect to the literature on school choice and parents' preferences for school attributes, studies typically estimate correlates of choices and, for example, schools' test scores. Such associations may be a function of correlated unobservables such as a school's reputation. In this study, by exploiting variation in ratings over time, I can more credibly claim to have a handle on the causal impact of school quality on parents' choices.

Finally, the results highlight the potential for relaxing information constraints using a top-down approach to monitoring and disclosure.[15] This is especially striking given that a number of influential papers have highlighted the problem of test score volatility and mean reversion for test-based accountability regimes. Figlio and Lucas (2004) show that the housing market in Florida initially responds strongly to state-administered school grades, but these effects fade as the market learns and adapts to the volatile nature of these grades.[16] Other studies highlighting the notion that schools may be rewarded or sanctioned on the basis of noise in test-based accountability regimes include Chay, McKewan and Urquiola (2005) and Kane and Staiger (2002). Given this body of evidence, inspection ratings would appear to address an important shortcoming in measures of quality of public goods which rely purely on hard performance data.

The remainder of is this paper is organized as follows. Section 2 lays out a simple hedonic model for school quality. Section 3 describes the context for this study and the data. Section 4 lays out the empirical strategy adopted to evaluate the impact of inspection ratings on house prices and reports the results. Section 5 reports the findings from the school choice and enrolment analysis and section 6 concludes.

[15] On the importance of information constraints in education markets, see for example Hastings and Weinstein (2008) who show that providing simplified information can generate large effects on parents' school choices. Other studies have found no effects (Bettinger et al., 2012) whilst still others have demonstrated that welfare may sometimes decline as a result of greater information dissemination leading to strategic response on the part of service providers (Dranove et al., 2003).

[16] Figlio and Lucas (2004) note that in Florida schools with large idiosyncratic gains from one cohort to the next struggle to match this performance in subsequent years, leading to large fluctuations in the assigned grades. Using Norwegian data, Fiva and Kirkboen (2011) also find evidence of very short lived housing market effects (fading within three months) to information disclosure of adjusted test scores.

2. School Quality and the Hedonic Pricing Model

Following Rosen's (1974) analysis of hedonic prices for differentiated products, assume that a consumer residing in a house with characteristics $z_1, z_2, \dots, z_n$ attains utility $U(x, z_1, z_2, \dots, z_n)$, where x is the numeraire good. The equilibrium price of a house with these characteristics is determined by the hedonic price function, $P(z_1, z_2, \dots, z_n)$. The implicit marginal price for each characteristic z_i is then define by $\frac{\partial P}{\partial z_i}$. The consumer's budget constraint is given by $y = x + P(z)$, where y is total income and z is the vector of the characteristics of the house. Maximizing utility subject to this constraint yields the optimization condition that the consumer selects z_i such that marginal rate of substitution with respect to the numeraire good is set equal to the marginal implicit price of z_i:

$$\frac{\frac{\partial U}{\partial z_i}}{\frac{\partial U}{\partial x}} = \frac{\partial P}{\partial z_i}.$$

Using the budget constraint to substitute for x in the utility function implicitly defines an individual's bid or value function θ as follows:

$$U(y - \theta, z_1, z_2, \dots, z_n) = u,$$

where u is some constant level of utility. Thus $\theta(z; u, y)$ is the amount the consumer is willing to pay for housing attributes z at a given level of income and holding utility constant. This yields the optimality condition, $\frac{\partial \theta}{\partial z_i} = \frac{\frac{\partial U}{\partial z_i}}{\frac{\partial U}{\partial x}}$. Thus at the optimum, the bid function will be tangential to the hedonic price function:

$$\frac{\partial \theta}{\partial z_i} = \frac{\partial P}{\partial z_i}.$$

Figure 1 illustrates tis optimum for the case where one of the housing attributes, z_1 say, is represented by school quality, Q. The figure shows bid functions for two households, located in two separate neighborhoods, initially consuming school quality q_a and q_c and paying implicit prices p_a and p_c, respectively.

Consider now an exogenous increase in quality of the nearby school for each of these two locations, so that the two households are now forced to consume q_b and q_d. There is a higher house (or rental) price associated with these higher levels of amenities, p_b and p_d, respectively. In the absence of transaction costs, and given enough time to move, households would relocate to a house providing the original level of schooling services at the original prices, i.e. points q_a, p_a and q_c, p_c, thus attaining maximum utility. However, assuming that the price changes associated with the marginal quality changes are small, transaction costs imply that these two households consume q_b and q_d school quality units. The change in house prices yields the welfare gain for homeowners from the rise in school quality and is an upper bound on the total welfare gain (since the loss in welfare for consumers forced to consume higher Q than their optimum would dictate, must be subtracted from this quantity).

On the other hand, households moving into these two neighborhoods following the improvement in school quality, who would have paid moving costs in any case, are willing to pay the marginally higher price for the gain in school quality and are located at their optimum at q_b, p_b and q_d, p_d, satisfying the 'kissing' property of the hedonic equilibrium.

An important point illustrated by Figure 1, and relevant to the empirical analysis below, is that the (marginal) change in school quality takes place across initially low and high quality schools (indeed, for the application below, these changes occur across the full range of different types of schools). This implies that one can estimate marginal implicit prices across the full spectrum of public schools, serving a varied set of communities. In addition, the analysis will identify the impact of school quality changes which are 'relevant' to families located near these schools. This may be contrasted with analysis of differences in willingness to pay for test score levels across school attendance zones. In this case, poorer families may rationally be unwilling to pay substantial amounts for lare test score gains if transferring to such schools leads to poor matching of student to school and small or no gain for their children.

3. Institutional Context and Data Description

The English public (state) school system combines centralized testing with a school inspection regime. For primary schools – the focus of this study – a key performance measure is the age-11 'Key Stage 2' test taken in May of each year, before students transition to secondary schools. Test scores are publicly disseminated, via government and official websites as well as via rankings of 'league tables' in newspapers. Since the early 2000's, information on schools' value added has also been publicly available.

In addition to test scores, the market also has access to inspection ratings provided by the Office for Standards in Education, or Ofsted. Over the period relevant to this study schools are usually inspected once during an inspection cycle, lasting from three to five years. Inspections entail a visit by two or more inspectors over a number of days who assess the quality of education being provided by the school. Inspectors spend a large proportion of their time observing classroom teaching but they also interview school leaders, examine students' written work and speak to parents. At the end of their visit inspectors write a report which includes a headline grade for the school.[17] These reports are made available in the Internet. These grades are on a four-point scale, ranging from 'Outstanding' to 'Fail.' As indicated in their official documents, the overall grade for a school reflects both the hard test performance data as well as the qualitative evidence gathered by inspectors during their on-site visit. The

[17] The following summarizes the role of inspections: "The inspection of a school provides an independent external evaluation of its effectiveness and a diagnosis of what it should do to improve, based upon a range of evidence including that from first-hand observation. Ofsted's school inspection reports present a written commentary on the outcomes achieved and the quality of a school's provision (especially the quality of teaching and its impact on learning), the effectiveness of leadership and management and the school's capacity to improve." (Ofsted, 2011, p.4, quoted in Hussain, 2015).

exact weights attached to the objective versus subjective measures are not clearly set out. For further details on the inspection process see Hussain (2015).

Anecdotal evidence suggests that the market does pay attention to inspection ratings. For example, Figure 1 shows screenshots from two of the UK's leading real estate search websites (Rightmove and Zoopla). Following a property search, consumers can view the property's location on a map which also displays local schools. As the figures show, the Ofsted rating and link to the report are readily available on these commercially provided maps.

2.1 Descriptive Statistics and Evidence on the Exogenous Timing of Inspections

I initially focus on inspections outcomes data as well as the timing of inspections. The focus of this study is the 2006 to 2008 school inspection round.[18] At the beginning of this period, September 2005, the inspectorate introduced a simplified, 'plain English' reporting style to its inspection reports, where the headline rating was reported upfront on the first page of the main body of the report, reports were made much more succinct and easier to decipher for the average parent. Furthermore, this was also the period when short notice inspections were introduced. In earlier years, schools had many months of notice of the exact date the inspectors ere due to visit the school. Over this period, as well as in earlier years, the timing of inspections was exogenously determined (see below). In years after 2008, the inspectorate moved to a regime where schools receiving worse inspection outcomes or those whose test score performance showed rapid deterioration were visited earlier and more frequently.

Table 1 shows the key characteristics for the approximately 8,000 primary schools with valid inspection and test score data.[19] Table 2 shows the transition matrix for inspection ratings

[18] Note that '2006' refers to the academic year 2005/06; '2007' refers to the academic year 2006/07, and so on.
[19] In general Primary schools in England cater to five to eleven year olds.

from one round to the next.[20] This clearly shows that there is a great deal of flux in ratings. For example, just over half of schools rated 'good' (grade 2) in the previous inspection round were also rated good in the current round; one tenth were uprated (to 'outstanding', or grade 3); and around one third were downgraded to 'satisfactory' (grade 1).

On the timing of inspections, Table 3 provides indicative evidence on the exogenous timing of inspections. This table shows that schools inspected early in the previous round are also the ones inspected early in the 2006 – 2008 round. Further analysis shows that any remaining differences in timing of inspections are largely unrelated to school performance. For example, regression analysis shows that the prior inspection year is a strong predictor of the current inspection year whilst prior test score performance has an economically very weak effect.[21] Furthermore, investigating correlates of the *month* of inspection conditional on year of inspection shows that prior year of inspection has an impact but lagged test score performance has no influence.[22]

2.2 Correlates of Inspection Ratings

---

[20] Note that schools failed during the 2006 to 2008 inspection round have been excluded from the analysis in this paper as they are subject to increased scrutiny, repeat inspections and higher turnover of the school leadership. See Hussain (2015) for further details of the workings of this this punitive aspect of the inspection regime.

[21] The estimated regression equation is as follows:

$$\underset{}{Year_s} = \hat{a} + \underset{(0.005)}{0.39} * PriorYear_s - \underset{(0.0003)}{0.001} * Test_s - \underset{(0.0003)}{0.0005} * FreeLunch_s,$$

where for school s, $Year_s \in \{2006,2007,2008\}$ is the current year of inspection; $PriorYear_s$ is the year of the previous inspection; $Test_s$ is the school's mean 2004 and 2005 test score national percentile rank; $FreeLunch_s$ is the school's national percentile rank on the percentage of students' eligible for free lunch; standard errors clustered at the Local Education Authority level reported in parentheses; N=8,287. These results show that a 3-year difference in the prior inspection year raises the expected current inspection year by more than 1 year; an increase in test percentile rank of 50 percentile points lowers the predicted year of inspection by 0.05 of a one year (i.e. around 3 weeks).

[22] For example, the results for schools inspected in 2006 are as follows:

$$Month_s = \hat{b} + \underset{(0.062)}{0.19} * PriorYear_s - \underset{(0.0035)}{0.0025} * Test_s + \underset{(0.0037)}{0.0021} * FreeLunch_s,$$

where for school s, $Month_s$ is the month of inspection with September coded as 1, October coded as 2, up to June coded as 10; see previous footnote for other definitions; N=2,185. These results show that prior inspection year has an impact in the expected direction whilst the lagged test score performance has a small and insignificant effect on the timing of inspections. Similar results are obtained for 2007 and 2008 inspected schools.

In the hedonics analysis below, the key parameter of interest is the impact of the change in inspection ratings on house prices. This section assess whether changes in inspection ratings from one inspection round to the next are correlated with observable changes in school characteristics, such as test score performance, value added, composition and size of the student body.

The main finding is that test scores – in levels as well as value added – are strongly associated with inspection outcomes, whilst the effects of student demographics are small and insignificant. If observable measures of SES such as the percent of students eligible for free lunch do not have any impact on the estimated effect of schools' test rank on inspection ratings then it would seem plausible to argue that unobservable changes in the makeup of the student body are also unlikely to influence changes in inspection ratings (see Altonji et al., 2005, for a formal discussion). This result is important in the context of interpreting the house price results reported below.

In order to undertake this analysis consider, for example, schools which were inspected in 2008, with a prior inspection in 2004. Using data corresponding to the 2008 and 2004 inspections, the following model can be estimated:

$$rating_{st} = \alpha_0 + \alpha_1 TestPercentile_{st} + \alpha_2 VApercentile_{st} + \gamma' W_{st}$$
$$+\alpha_3 1(t = 2008) + \lambda_s + u_{st}\ , \qquad (1)$$

where $t \in \{2004, 2008\}$ and λ_s is the fixed effect for school s; hence, this model is equivalent to a first differences model. $rating_{st}$ is the inspection outcome for school s in inspection year t and $TestPercentile_{st}$ is the school's mean test score performance in the two years prior to the inspection, measured in national percentiles. $VApercentile_{st}$ is the school's mean value added performance in the two years prior to the inspection, measured in national percentiles.[23]

[23] As explained below, the VA variable is not available for earlier inspection years.

W_{st} is a vector of three variables: the proportion of students receiving free lunch; the proportion of minority students; and a measure of school size, the number of full-time equivalent students. Finally, α_3 captures changes over time, due perhaps to changes in overall standards in schools.

Model (1) represents the setup for schools inspected in 2008. In practice, all inspections from the three years 2006 to 2008 are combined, along with data from the prior inspections, in order to estimate the school fixed effects model.

Table 4 reports the results of this analysis. Note that the value added variable is excluded from the model in columns 1 and 2 since this variable is not available for the prior inspection for the majority of schools inspected in 2006 and 2007.[24] The results in column 1 show a strong and statistically significant relationship between changes in inspection ratings and changes in schools' test performance: a rise in a school's test rank of 20 percentile points from one inspection to the next (which corresponds to approximately one standard deviation of the change in test percentile rank between the two inspections) is associated with a 0.24 rise in the inspection rating.

Importantly, the results in column 2 show that the impact of test rank is unchanged when controls for the composition and size of the student body are included. In fact the coefficient on the measure of SES, the school's percentile rank on the proportion of students eligible for free lunch, is small, *positive* and insignificant. The effect of the proportion of minority students at the school also appears to be small, although it is marginally significant. The finding that changes in school demographic covariates do not affect changes in the inspection ratings is also supported by the results in column 3, which now includes the school's performance as measured by mean student value added for those schools inspected in 2008.

[24] Student value added data were first published in 2003. Therefore, this variable cannot be included for schools inspected in between 2006 and 2008 where the prior inspection is from 2003 or earlier.

Both test scores in levels as well as value added have strong effects on inspection outcomes, whilst the effect of students' SES composition is small and insignificant.[25]

2.3. The Information Content of Inspection Ratings

In this section I ask whether the information embodied in inspection ratings adds any value over and above information already in the public realm. I undertake a forecasting exercise in order assess whether gains in school performance for highly and poorly rated schools revert to mean. In answering this question I rely on an institutional aspect of the English testing regime whereby inspections can sometimes fall in a window between the testing of students and the revelation of those test results. The details of this setting are as follows. All age-11 primary mschool students take the 'Key Stage 2' test in the second week of May each year. These tests are centrally administered, are hand written by students and sent off for marking by external examiners. Results are revealed to all stakeholders at the same time, in July. Inspections take place throughout the academic year, including in the weeks straight after the test. I focus on those schools inspected in June, after the test takes place, but before the results are revealed in July. Figure 1 illustrates the timing of events in a given academic year.

Using this setup, I ask whether the May test performance for schools up- or down-graded in the June inspection is better or worse than would be expected given historical performance.

The question I address is whether the June inspection ratings can forecast the yet-to-be-disclosed May test performance. The key idea here is that the May test is unaffected by the inspection outcome but is not known to the inspectors (nor teachers or parents).

[25] The results in column 3 suggest that a rise in a school's value added rank of 35 percentile points (corresponding to the standard deviation for the change in value added rank) is associated with a 0.19 rise in the inspection rating.

Table 5 provides some preliminary evidence on this issue. To begin, Panel A depicts all schools inspected between 2006 and 2008 and does not exploit this natural experiment. The first row of Table 5 shows that there are large losses (gains) in performance between the two inspections for donwngraded (upgraded) schools. For example, for the 317 schools experiencing a two-unit drop in their rating (from 'outstanding' to 'satisfactory') the mean performance on the mathematics and English test for the two years immediately prior to the current inspection is 12.7 percentile points below the mean for the two years prior to the previous inspection. This is in line with the notion that inspectors are attentive to past performance. Rows 2, 3 and 4 show how performance evolves in the year of inspection, one year after and two years after inspection, respectively.

By focusing on June inspected schools, the evidence in Panel B addresses the concern that inspection outcomes may be driving the results in Panel A (e.g. teacher effort rises following a poor inspection rating). The main results from this analysis are in the final row of Panel B and these suggest that there is little evidence of mean reversion in test scores following good and poor inspection outcomes.[26]

Table 6 reports results from a regression analysis performed for schools inspected in June. The dependent variable is the change in test percentile points between the May test in the year of inspection and the prior year. Column 1 shows the result of regressing this outcome on the change in percentile points between the test from the year prior to the current inspection and the test from the year prior to the previous inspection. The estimate of 0.34 implies that each percentile gain between the two inspections is reversed by 0.34 points in the year of the current inspection. This result corroborates the general finding in the literature that test scores exhibit strong mean reversion (refs XX). Column 2 suggests that a 3.1 percentile gain is

[26] For example, the change in test percentiles between the test in May of the inspection year and the prior inspection year for schools experiencing a unit decline in their rating is negative (-1.6), compared to a decline of 6.2 in the periods between the current and prior inspection.

sustained for schools experiencing a unit rise in their May inspection rating. A 3 percentile point gain is sustained even after controlling for changes in other observable controls (column 3) as well as a detailed set of school controls (column 4).

The key finding from this analysis is that schools uprated (downgraded) by the inspectors are partly selected on the basis of past improvements (declines) in test performance; as would be expected under a mean reversion scenario, some of this improvement (decline) is reversed in the May test in the year of the inspections. But, importantly, upgraded (downgraded) schools perform better (worse) than would be predicted conditional on their recent test score trajectories (as well as a long list of covariates).

4. The Impact of Inspection Ratings on House Prices

4.1 Empirical Strategy

Estimating the impact of school quality ratings on house prices is not straightforward. The concern with regressing house prices on inspection ratings would be that correlated unobservables drive both ratings as well as house prices. The empirical approach adopted in this paper is to compare house prices just after an inspection for houses located near schools experiencing an improvement or decline in their rating versus houses located near schools experiencing no change. As explained in detail below, the credibility of this strategy critically hinges on the fact that the timing of inspections is unrelated to school performance.

Using data from the months straight after the inspection (the post-treatment period) and from the academic year before the inspection (the pre-treatment period), the following difference-in-differences model can be estimated:

$$y_{ist} = \beta_0 + \beta_1 1(t = post\ year) + \delta ChangeRating_s * 1(t = post\ year)$$
$$+\gamma_1' X_{st} + \gamma_2' Z_{ist} + \lambda_s + u_{ist}, \qquad (2)$$

where y_{ist} is the log of the sale price for house i near school s in year t. $ChangeRating_s$ is the rating in the 'post' year minus the one from the previous inspection; $ChangeRating_s \in \{-2, -1, 0, 1, 2\}$.[27] The dummy $1(t = post\ year)$ is switched on for house sales in the months straight after the inspection and is switched off for sales in the academic year immediately before the inspection.[28] λ_s is a school fixed effect. X_{st} is a vector of time-varying school characteristics, including mean performance on the age-11 'Key Stage 2' test, test score value added,[29] the proportion of students receiving free lunch and the proportion of minority students. Z_{ist} are characteristics of the property, including proxies for its size.[30] The main analysis below is conducted using property transactions within a 500m radius of the school. Errors, u_{ist}, are clustered at the school level.

The parameter of interest is δ, the impact of a unit improvement in the inspection rating on house prices. The identification assumption is that the counterfactual change in prices for houses near schools experiencing an increase or decrease in their inspection rating is captured by the change in prices experienced by homes located near schools experiencing no change in their rating. I probe the common trends assumption by testing whether there is a treatment effect in the years immediately prior to the inspections.[31]

---

[27] Ratings take on the values 1,2 and 3 ('satisfactory', 'good' and 'outstanding', respectively). As discussed above, schools receiving a fail rating are dropped from the analysis.

[28] In practice data from the inspection years 2006, 2007 and 2008 are combined and year dummies included. All regressions also includes dummies for month of house sale.

[29] Test scores, in levels and value added, are included with a lag of one year. House sale in late summer (July and August) may be influenced by contemporaneous test scores (which are released in July of each year). In a set of robustness tests (not reproduced here to conserve space), I also include contemporaneous test scores. Inclusion of these has virtually no effect on the main results.

[30] The property control variables are dummies for whether the property is detached, semi-detached, terrace or flat (apartment); freehold or leasehold; and newbuild or not.

[31] An additional concern is that the inspection outcome changes the composition of houses sold straight after an inspection. I assess the importance of this type of selection bias by comparing estimates with and without controls for house characteristics.

A potentially more serious threat to identification is that even if the common trends assumption holds in previous years, the possibility remains that correlated unobservables are driving the results. For example, a change of leadership at the school may lead to changes in perceptions of school quality, which then lead to changes in the inspection rating as well as house prices. In order to address such concerns I exploit the fact that over the period of analysis the timing of inspections is exogenously determined. Under this assumption, a simple test for the importance of omitted variables is to assess whether there is an impact of a change in inspection ratings on house prices in the months *immediately prior* to the inspection. Any significant effect in this placebo regression would suggest that the main results may be subject to bias.

Figure 4 illustrates the main idea. Consider the set of schools receiving an inspection in February 2009, leading to an uprating, say. For the treatment analysis, sale prices for nearby houses from April through to August are used for the post-treatment sample (March sales are ignored because it may take up to a month to release the report). For this difference-in-differences model, the pre-treatment sample consists of houses located near this set of schools in the prior academic year (September 2007 to August 2008). For the placebo analysis, house sales from September 2008 through to January 2009 are used for the 'post' sample; the pre-treatment sample is as before, i.e. sales in the academic year 2007/08.

If omitted variables are a real threat, then we would also expect to see a significant treatment effect for the placebo regression. The reason for this is that 'good news' (unobserved by the econometrician) for the uprated schools arrives throughout the year (and arguably is more likely to arrive at the beginning of the academic year when new personnel are in place, new curricula and programs may be up and running, etc.). If inspectors and house prices both respond to this good news (the omitted variable), we would expect some market reaction to inspection ratings even before the inspection takes place.[32] I interpret a finding of no effect in the placebo regression as evidence that correlated unobservables do not lead to bias in the main results.

4.2 Results

Table 7 presents estimates of equation (2). The first specification in Column 1 simply includes school fixed effects. Column 1 reveals a raw gain of 0.4 percent for each unit improvement in the inspection rating, which is statistically significant at the 10 percent level. The regression in Column 2 introduces controls for type of property. House characteristics include dummies for whether the property is detached, semi-detached, terrace or a flat (apartment); freehold or leasehold; and whether it is newly built or not. Addition of these controls increases the point estimate to around 0.5 percent and improves the precision, so that estimates are now significant at the 5 percent level. The fact that the fit (r-squared) of the model rises shows these housing controls are important. The fact that the estimates do not decline and in fact rise somewhat, show that the house price results are not biased upward by the selection of types of properties which are placed on the market following an inspection.[33]

Column 3 includes time-varying school characteristics (test scores in levels; value added; percent students eligible for free lunch; percent minority students; all measured in national percentiles). As the evidence above demonstrates, changes in ratings are correlated

[32] If timing is non-random, i.e. the month in which inspectors arrive at the school is related to changes in school quality, then the placebo regression will yield a zero treatment effect, even if the change in house prices is a result of the unobserved change in school quality. Thus, the fact that timing of inspections is indeed exogenous is of critical importance for this identification strategy to work.

[33] If better quality or bigger homes are put up for sale following a positive inspection, but the true impact of the change in rating is zero, column 1 would show a positive impact, shrinking towards zero in column 2. An alternative hypothesis might be that *worse* properties are on the market following a positive inspection (if for example, current families in the local school put off selling their properties). Under this scenario, the estimate in column 1 would be a *downward* biased estimate of the true impact of ratings. When property controls are introduced in column 2, the estimate should rise. The evidence presented in Table 7 is arguably in line with this hypothesis.

with changes in test scores and value added. However, inclusion of these key variables changes the estimate very little.[34] This evidence bolsters the hypothesis that the impact of changes in ratings on house prices is real.

The evidence in columns 1 to 3 of Table 7 is highly suggestive that the causal impact of a unit rise in inspection ratings is around half of one percent. But in order to provide a fully convincing argument that these estimates are not subject to bias, column 4 reports results from rthe important placebo regression described above. This regression investigates whether there is a treatment effect for property transactions in the months immediately prior to the inspection. As explained in detail above, this test is used to rule out the possibility that the results in column 3 are contaminated by any remaining omitted variable bias. The small and insignificant estimate for the placebo treatment effect suggests that this is not a major threat to the identification strategy. Thus, we can rule out that the possibility that parents are reacting to some unobserved measure of school quality rather than the ratings per se.

Column 5 reports estimates from an alternative placebo regression, this time employing data from the year before inspection (classified as the post or 'treatment' year for the purposes of this placebo analysis) and two years before inspection (classified as the prior year). The results for column 5 suggest that schools up-/downgraded and those experiencing no change in their rating, all experienced the same changes in house prices immediately prior to inspection. These findings support the common trends assumption underlying the difference-in-differences model.

Heterogeneous Effects

[34] Note that using mean test performance from the two years prior (instead of from one year alone) leaves the estimates virtually unchanged.

In order to investigate whether the market response to ratings varies by family background, Table 8 reports results for the full model which now also includes a triple interaction term, that between the change in rating, the post dummy and the percent of students eligible of free lunch at the school.[35] Column 1 reproduces the basic result. [36] Column 2 indicates that there is striking heterogeneity in this mean response. The main effect and the interaction with the free lunch percentile rank variable are significant at the 1 percent and 5 percent levels, respectively. For the least deprived schools, a unit rise in the rating leads to an approximately 1.5 percent rise in house prices, three times the mean effect reported in column 1. For school serving the poorest households, ratings have no impact.[37]

Evidence on Medium-Term Effects

In order to analyze the medium-term effects, I estimate model (1) using transactions data from one year before as well as one and two years after the inspection. Unlike the previous empirical strategy, this simple fixed effects strategy does not allow for the test employed to rule out the threat posed by omitted variables (described in Figure 4). Nevertheless, as demonstrated next, there would appear to be a strong case in favor of interpreting these medium-term estimates as causal effects.

Table 9 reports the results. The basic regressions in columns 1 and 2, without and with property characteristics, respectively, suggest a statistically significant gain of between 0.4 and

[35] [-The assumption underlying the use of this measure of SES is that …local houses..local families in school… Given that more than half of studnets attend their nearest school [reference], this assumption seems realistic.]
[36] For the triple interaction term, I use the school's percentile rank on the 2004 and 2005 mean of percent students eligible for free lunch. Schools with missing prior free lunch or test score data are dropped and thus there is a small fall in the sample size compared to that employed in column 3, Table 7.
[37] Column 3 repeats the analysis with the school's prior test percentile in the interaction term, in place of free lunch percentile. Although a similar pattern of heterogeneous response emerges (an effect of around 0.8 percent for schools scoring in the top end of the prior test score distribution; close to zero for those in the bottom of this distribution), the results are not significant.

0.5 percent for each unit rise in the rating. Importantly for the fixed effect strategy, when key school quality controls – which might be expected to influence both the rating and property prices – are added in column 3, there is almost no change in the estimated effects. This lends credibility to the argument that these estimates are not contaminated by omitted variable bias.[38] Before turning to the results in column 4, the placebo results in table 5 – which employ data from one (the 'post-treatment' year) and two years (the 'pre-treatment' year) before inspection – demonstrate that there is no evidence of differential trends in house prices for schools experiencing an up, down or no change in the rating.

As explained above, the estimates in columns 1 to 3 use transactions data from one and two years after the inspection, thus these results are the mean effect over these two years.[39] It is also useful to estimate the dynamic effects of the rating. The impact of the rating may diminish over time, perhaps because the information becomes less salient or 'newsworthy' in subsequent years. On the other hand, the initial response may be sustained or even increase over time if, for example, consumers learn via social networks. In column 4, row 2, the model now includes estimates of a triple interaction term, Change-in-rating * Post-year * Second-year, where the Second-year dummy is switched on for transactions from the second year after inspection. The results in the first row indicate that one year after inspection, a unit rise in the inspection rating leads to 0.4 percent rise in property prices (in line with results presented in Table 7). By the second year, this premium *increases* by more than a third (estimate reported in the second row, column 4) to nearly 0.6 percent. Note however, that this difference in the one and two year premium is not statistically significant.

---

[38] See Altonji et al. (2005) for a formal statement of this intuition. As in Table 7, the within-school changes in test and value added percentile are not statistically significant, although there is some evidence that the market does respond to changes in the percent of students receiving free lunch (significant at the 10 percent level). The latter may be a consequence of the fact that the panel dimension is longer (relative to that used in Table 1) and so the signals are easier to pick up over this longer horizon. Note that this variable is likely correlated with changes in the socioeconomic composition of local residents and therefore it is unclear whether the estimated impact reflects a response to changing neighborhood conditions or school quality.

[39] Note that *three* years after an inspection, most schools can expect a new inspection visit.

Further Robustness Tests

[to be added]

5. School Choice Analysis

4.1 Data and descriptive statistics

The data consist of parents' ranked preferences for primary schools from applications made in the fall of each year, 2001 to 2008. The school choice data from the borough also include the full home postcode of the applicant and whether the child is offered a spot in one of the listed school.[40] This latter piece of information is critical, as it is used to construct the 'cutoff' distance for each school, as explained below.

Parents list up to four schools, ranking them in order of preference. The assignment rule prioritises children with special needs and children with a sibling already in the school.[41] The next priority is based on distance from home to school: for secular schools, children living closest to the school are given priority. A given secular school facing excess demand is assumed to be in a child's choice set if her home falls inside the cutoff radius for that school (assuming the child does not qualify for special needs and does not have a sibling at the school). For secular schools the cutoff distance is determined using information on the child without special needs or a sibling, living furthest away from the school who was allocated a place at the school in the previous year. For religious schools, spots in the school are allocated on the basis of

[40] There are 1.8 million individual postcodes in the UK, with an average of 16 households per postcode. Thus, using the postcode to construct the home-school distance variable should result in minimal measurement error.
[41] Special needs children are excluded from the analysis. Most of the analysis below is on the sample where there is no sibling in a local public school. I also report some results for the full sample, i.e. including applicants with a sibling.

religious affiliation. In the absence of data on religious affiliation (not available at the borough level), whether a religious school is in the child's choice set cannot be determined. Consequently, religious schools and students who apply to a religious school are dropped from the analysis. See the data appendix for further details. School-level information on geographical location, test score performance, percent of students eligible for free lunch and inspection ratings are sourced from administrative data as described in section [3xx].

Table 11 shows some simple statistics from these data. Coulmn 1 shows means for the full sample and column 2 shows means for the sub-sample of applicants who do not have an older sibling already attending one of the preferred schools. The majority of the results reported below are for the latter sample, although I also report some results for the full sample so that comparisons can be made.

Table 11 shows that the mean number of schools listed by parents are around 2.3 to 2.7 whilst just over 26 schools – of the full set of 46 secular public schools in the borough – are available to parents (i.e. are in their choice set).[42] 39 percent of applicants already have an older sibling in one of the schools requested. Around 22 to 24 percent of students are eligible for free lunch.[43] Turning to the characteristics of the first choice school, this is on average just over 1km from the student's home, and is approximately the third nearest school, indicating that families are making an active choice and not automatically enrolling their child into the nearest school. All schools were inspected once during the 2006 to 2008 inspection cycle. Of the 46 schools in the analysis sample, 13 were inspected in 2006, 19 in 2007 and 14 in 2008. The mean change in ratings of -0.28 indicates that ratings were lower on average in the 2006 – 2008 round than in the earlier inspection round, but nevertheless there is substantial variation around this mean (standard deviation of the change is 0.96). The rankings on test score and

[42] Recall that a school is unavailable if the child does not have a sibling at the school and she resides outside the cutoff distance for the school, determined using the previous year's marginal entrant. In practice I add 200m to the cutoff distance to allow for the fact that the cutoff distance can vary from year to year.

[43] Free lunch status is obtained from post-enrolment data linked back to the applications data.

percent of students eligible for free lunch indicate that the first choice schools come from close to the middle of the test score and SES borough-level distributions. Further analysis, not reported here, shows that on average the first choice school performs better on inspection ratings and test scores and also has a lower proportion of students eligible for free lunch compared to the mean for the nearest three schools in applicants' choice sets

4.2 Conditional logit model and identification

The approach to the empirical analysis is the standard conditional logit model (McFadden, 1974). Parents of student *i*, applying for a school in year *t* are assumed to choose from the available set of alternatives, i.e. school $j \in \{1, 2, .., J\}$, in order to maximize utility,

$$U_{ijt} = \delta ChangeRating_j * Post_t + x'_{ij}\beta + \gamma_j + e_{ijt}.$$

The deterministic part of utility is represented by the change in school *j*'s inspection rating, interacted with a dummy which is turned on if the application is made in the year after the rating is revelaed ($Post_t$), and $x'_{ij}\beta$, where x'_{ij} represents a vector of school and match-specific attributes (test scores and the percentage of students eligible for free lunch at school *j*, as well as distance from *i*'s home to school *j*). Importantly, alternative-specific (i.e. school) fixed effects are also included in the form of γ_j. The parameter of interest is δ, the impact of a unit change in the rating. The error term e_{ijt} is the random component of utility, assumed to be i.i.d. and from a type I extreme value distribution. This framework yields the conditional logit model, where the probability that student *i* chooses school *j* is given by

$$\Pr(Y_i = j \mid ChangeRating_j, t, x'_{ij}) = \frac{\exp(\delta ChangeRating_j * Post_t + x'_{ij}\beta)}{\sum_m \exp(\delta ChangeRating_m * Post_t + x'_{im}\beta)} \quad (1),$$

where *m* is the set of schools available to applicant *i*.

Identification of the impact of the change in the rating, δ, relies on the within-school variation in ratings over time. In fall 2004 the $Post_t$ dummy is switched *off* for all schools; in fall 2008 it is switched *on* for all schools. Inspections take place in one of the three years 2005/06, 2006/07 or 2007/08, and so for applications made in the years between 2004 and 2008, the $Post_t$ dummy will be zero for some schools unity for others. I also run a placebo regression using applications data from fall 2004 and earlier to test whether there is any impact from the future change in inspection rating.

4.3 School choice results

Table 11 reports coefficient estimates from the conditional logit model of first choice school outcome. Columns 1 through 7 use the sample of applicants with no older sibling attending any school listed on the application form. Estimates in the last two columns are based on the full sample.

All versions of the model include school fixed effects except for columns 1 and 2. The model in column 1 includes only distance (linear and quadratic terms) and the change in inspection, interacted with the *post* dummy. The model estimated in column 2 includes the school's test decile as well as percent students eligible for free lunch, but not alternative-specific fixed effects. This in effect reproduces the traditional specification for the school choice model estimated in the literature. In line with many previous studies (e.g. Hastings et. al, 2009, and Burgess et. al, 2009), these results show that families value the school' proximity; performance as measured by test scores; and place a negative weight on the proxy for student SES composition, the proportion of students eligible for free lunch. Of course, these measures may be correlated with omitted school characteristics, such as a school's reputation, and are not necessarily informative about the causal impact of, for example, test scores. Finally, columns 1 and 2 provide some initial evidence that families value schools which improve their inspection ratings.

Turning to the main specification which includes school fixed effects, column 3 demonstrates that parents are responsive to changes in inspection ratings: the estimate for δ is significant at the 1% level and has an economically meaningful impact. The value that families place on a unit rise in the rating can be measured in terms of the tradeoffs they are willing to make with respect to extra travel distance. A coefficient of 0.1 for δ implies a willingness to travel approximately an extra 80m or 7 percent (assuming the mean travel distance of 1.1km) in order to attend a school which gains a unit rise in its rating. Although this may seem small, it comparable to the preference for a decile rise in a school's test score rank: the results reported in column 2 suggest that one decile rise in test scores is associated with a willingness to travel an extra 110m.[44]

Introducing school fixed effects in the model renders school's test decile as well as percent students eligible for free lunch insignificant (column 4). Both of these variables are two-year moving averages from the year before application. These results suggest once again that parents are unable to distinguish underlying trends from noise in noisy test score metrics. On the othet hand, introducing these variables leaves the parameter estimates for δ unchanged.

Columns 5 to 7 report on heterogeneity in treatment effect by poverty status, measured by eligibility for free lunch status. Columns 5 and 6 stratify the ample by the student's free lunch eligibility status. Column 7 introduces a triple interaction treatment in the second row, testing for a differential effect for disadvantaged families. The striking finding here is that if anything, poorer families respond *at least as much* as non-poor families. In column 7 the

[44] Given that residential location and school choice may be jointly determined, disutility of distance is likely overstated in this model and hence these willingness to travel estimates likely understate true preferences for school characteristics.

interaction is statistically significant at the 10% level and economically relevant: the response from poorer families is 60 percent larger than for non-poor families.

All of the estimation up to column 7 is based on the sample of applicants without an older sibling in a local primary school. Column 8 now estimates the impact of the change in rating on the full sample, including families which already have a child in primary school. The costs of responding to new signals of quality by switching schools are likely higher for families with an older child already enrolled in primary school. Hence the expected response to the change in rating is expected to be smaller for this group. The results in column 8 corroborate this intuition. The estimate for the full sample is now around a third smaller than for the earlier sample, although it is still statistically significant at the 5% level.

Finally, column 9 reports results for the placebo analysis. Application data from fall 2002 to fall 2004 are used for this analysis, with the *post* dummy switched on for the 2004 applications. The change in inspection rating continues to be for the inspections carried out in one of the years 2006 to 2008, relative to the previous inspection round. In effect, the issue is whether future changes in rating have an 'impact' in the years before they were actually revealed. This may the case, for example, if parents independently respond to the same unobservables that inspectors respond to in forming their judgements. The results in column 9 suggest that there is no evidence of this: the coefficient estimates on the placebo treatment dummy is small and statistically insignificant. This finding bolsters confidence in the interpretation that the results reported in Table 11 provide evidence on the causal impact of the change in rating on consumer preferences.

5.4 Enrolment Analysis

The analysis of parents' application data yields useful insights into how parents respond to school quality ratings. I now assess to what extent such responses lead to changes in aggregate school enrolment at the national level. A rise in enrolment in response to a higher inspection rating may imply that, to some extent at least, higher demand may be met by expansion of school places (perhaps at those schools where there is spare capacity), implying that the housing market response is a lower bound estimate of the true willingness to pay for higher school quality.

Table 12 reports the results from this analysis. Once again, I exploit the variation in timing of inspections. Using enrolment data for English primary schools, the question addressed now is whether log-enrolment changes in response to a change in the school's rating. School-level data are employed from 2005, the year just before the first inspection in the 2006 to 2008 inspection cycle, through to 2009 (see the appendix for further details of the data). The 'post year' dummy (row 1 in Table 12) is switched on in the years after the inspection. Column 1 shows that without any school controls, a unit rise in the inspection rating leads to a 1.6 percent rise in enrolment. Adding controls (column 2) shows that this estimate is unchanged. There appears to be a modest, statistically significant impact on enrolment from short-term changes in the school's rank on the percent students eligible for free lunch; there is a very small and marginally significant effect of the school's test score decile.

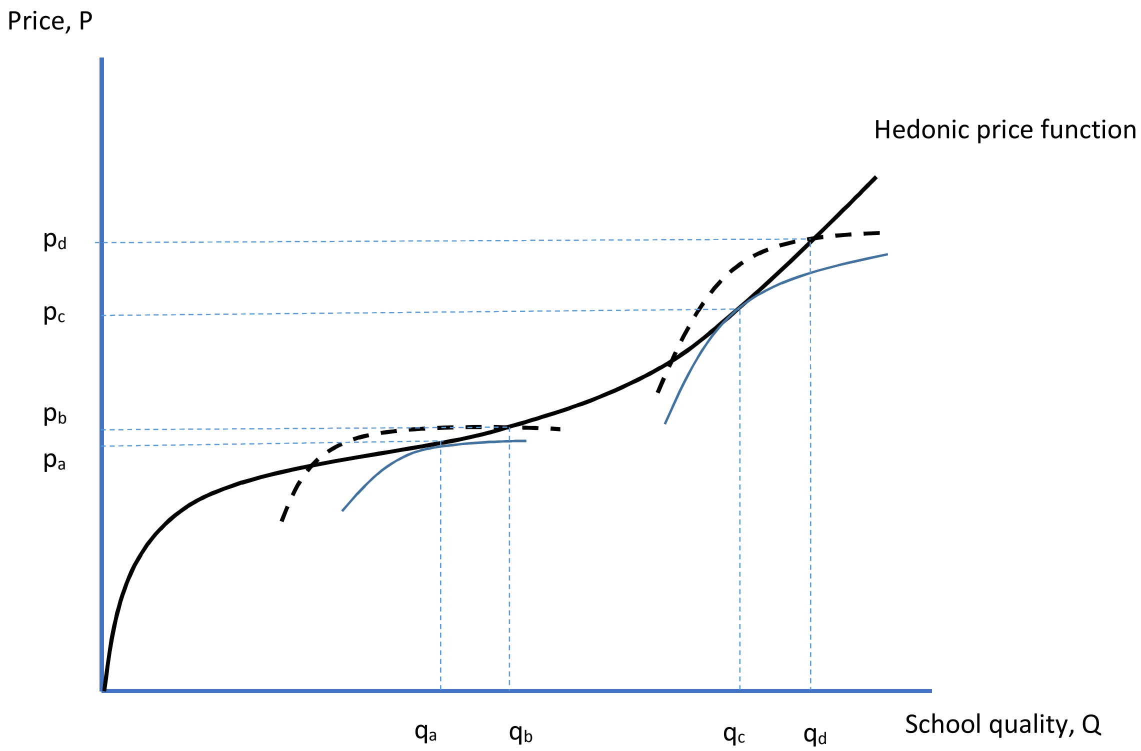
Column 3 includes a triple interaction term between the change in rating, post dummy and the school's decile rank on the percent of students eligible for free lunch. This shows that the impact on enrolment from a unit change in the rating is 70 percent bigger for schools with the highest proportion of students eligible for free lunch than for ones with the lowest. Finally, the results for the placebo analysis in column 4 employ data from the years 2003 to 2005, with the post dummy switched on in 2005. This shows that there are no differential trends for schools improving their rating versus those experiencing a decline in their rating in the years prior to the inspection.

There are two key messages from the analysis of the enrolment outcome. First, there is robust evidence that schools expand and contract in response to better and worse ratings, respectively. This has clear implications for the hedonic analysis reported above: the estimated housing market capitalization likely understate the true implicit prices for improvements in school quality.

Second, the finding that enrolment rises by more for schools serving lower SES students suggests an alternative explanation for the earlier finding that there is a negligible price impact for housing units located near these schools. If there is greater excess capacity at these sorts of schools, then any change in demand may be felt through school switching (at application, when switching costs are lower) rather than the price mechanism.

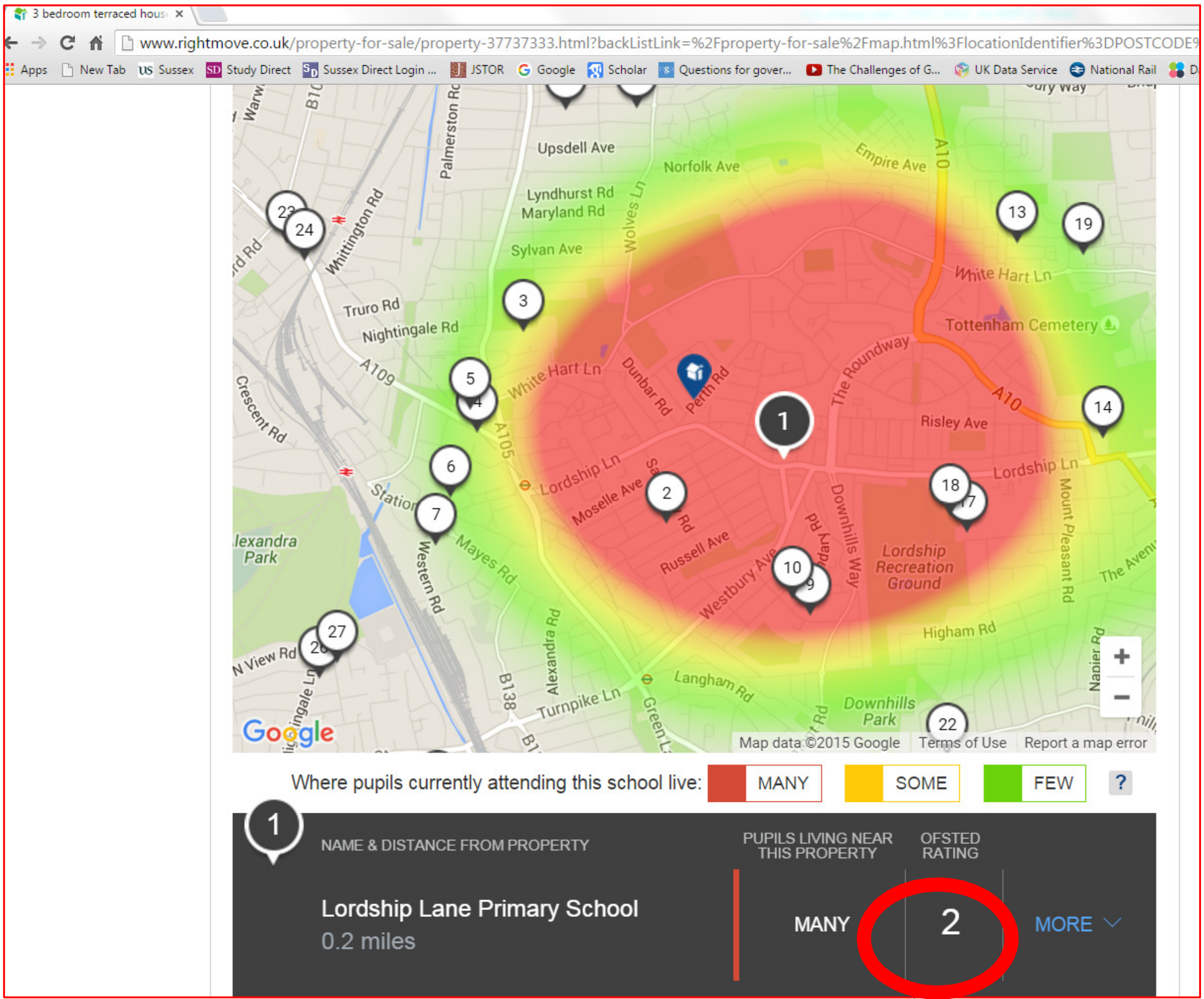
7. Conclusion

References



Price, P
Hedonic price function
pd
pc
pb
pa
qa
qb
qc
qd
School quality, Q

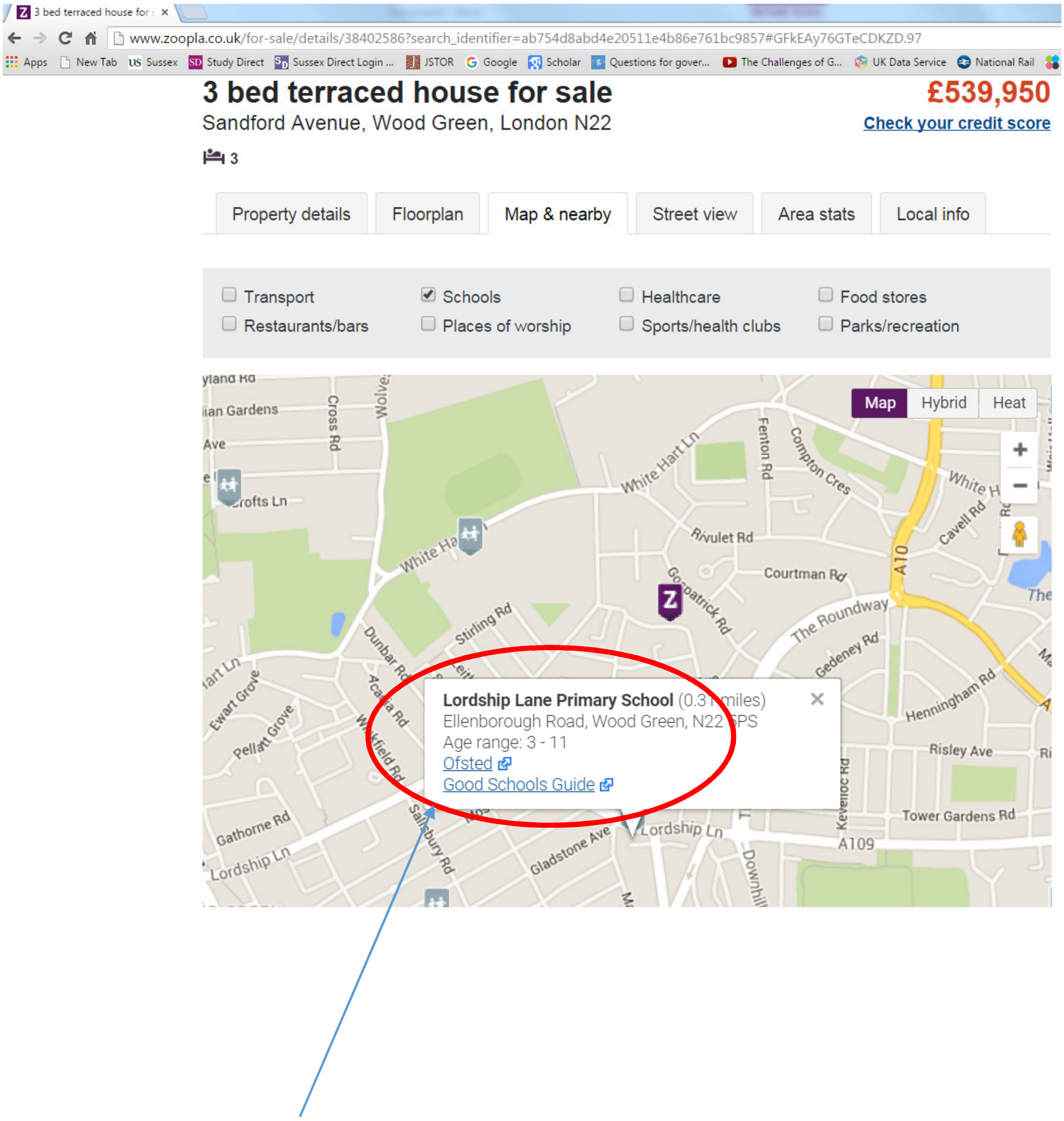
Figure 1: Housing Market Websites and School Inspection Ratings (Rightmove and Zoopla)



School Ofsted rating

(house for sale indicated by blue icon in map, nearest school identified as black circle with '1' in the map)

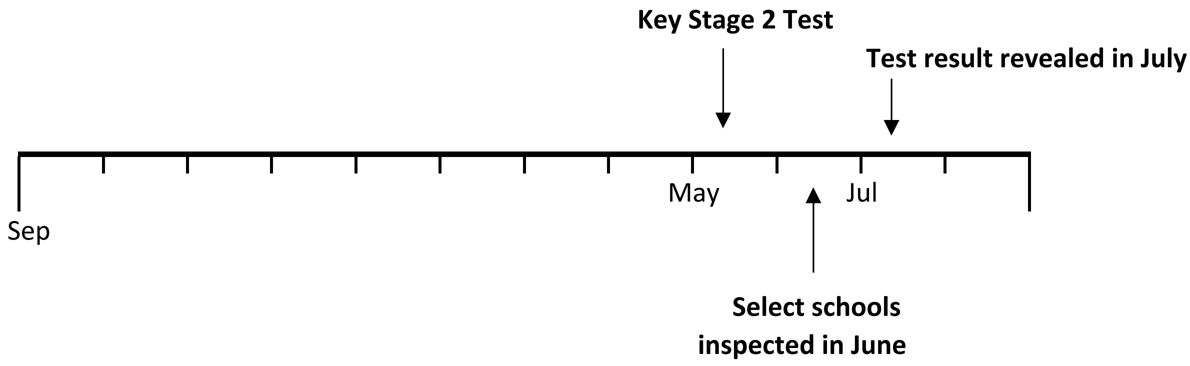
Figure 1 (cont.)



Link to school's Ofsted report

(house for sale indicated by 'Z' icon in map)

Figure 2: Timing of the 'Key Stage 2' Test and Results



Note: KS2 test takes place in second week of May.. See text for full details

Figure 3: Illustration of Empirical Strategy

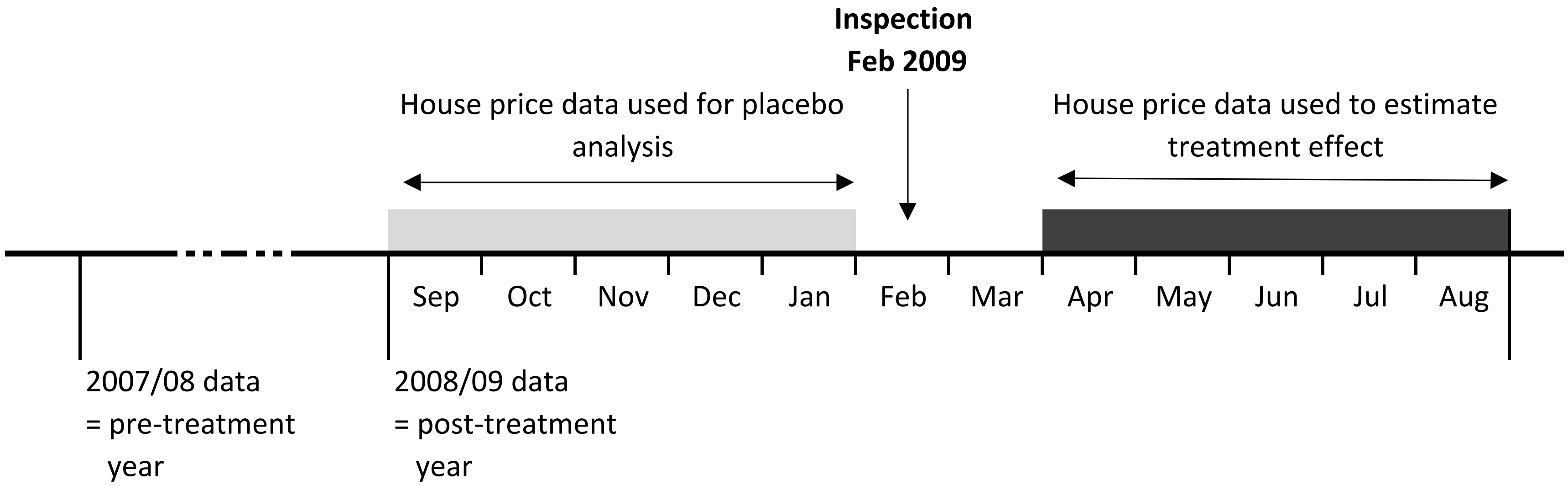


Table 1 Descriptive Statistics

Change in inspection rating	-0.19 (0.79)
Test score: percent students attaining performance threshold	80.6 (12.0)
Value added	100.2 (1.1)
Percent students free lunch	15.5 (14.7)
Percent students minority background	18.8 (24.4)
Size school (full-time equivalent students)	251.5 (120.3)
# schools	8,287

Notes: Schools inspected in one of 2006, 2007 or 2008; change in rating represents change between current inspection and previous inspection (typically three to six years earlier). School's test score performance calculated using the mean of percent of students attaining the Level 4 threshold on the age-11 Key Stage 2 English and Mathematics tests. Published Department for Education data used for value added measure; 100 corresponds to expected progress (between age-7 Key Stage 1 and age-11 Key Stage 2) and 101 (99) corresponds to one term's more (less) progress than expected. School-level variables are from the year before inspection.

Table 2 Current and Prior Inspection Ratings

	Current inspection rating		
	1	2	3
Prior rating			
1	1,266	956	86
2	1,399	2,252	413
3	329	1,068	518

Notes: Current inspection is from 2006-2008. Schools receiving a fail rating are excluded from the analysis. Total number of schools: 8,287. See text for further details.

Table 3 Timing of Inspections: Current and Prior Year of Inspection

	Current year of inspection		
Prior year of inspection	2006	2007	2008
1995	0.2		
1996	1.7	0.1	0.0
1997	0.2	0.3	0.1
1998	0.0	0.1	0.1
1999	0.0	0.1	0.5
2000	61.1	2.3	0.2
2001	34.2	28.9	0.1
2002	0.1	44.5	4.8
2003	0.0	22.2	25.1
2004	2.2	0.8	52.3
2005	0.0	0.6	16.7
Total	100.0	100.0	100.0
# schools	2185	3184	2918

Notes: Cells report percentage of schools in each of 2006, 2007 or 2008. Each school is inspected once in the period 2006-2008, so number of schools equals number of inspections. The highlighted boxes show that schools inspected early (late) in the previopus cycle are likely inspected early (late) in the subsequent cycle.

Table 4 Relationship Between Change in Inspection Ratings and Change in Recent School Performance

(Outcome variable: Inspection rating)

	(1)	(2)	(3)
	2006 to 2008 inspections		2008 inspections
School test score percentile	0.0124***	0.0124***	0.0086***
	(0.0006)	(0.0006)	(0.0014)
Value added percentile			0.0054***
			(0.0007)
Percent students free		0.0020	0.0018
lunch, percentile		(0.0012)	(0.0027)
Percent minority students,		0.0011*	0.0017
percentile		(0.0006)	(0.0012)
Log size school		-0.1241	0.0969
(full-time equivalent students)		(0.0813)	(0.1972)
School fixed effects	Yes	Yes	Yes
Observations	15,958	15,958	3,802
# schools	7,979	7,979	1,901
R-squared	0.7078	0.7086	0.7392

Notes: Standard errors reported in parentheses; *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. S.e.'s clustered at the school level. All regressions include year of inspection dummies. Outcome is from current inspection year (one of 2006, 2007 or 2008 for column 1; 2008 for columns 2 and 3) and the previous inspection year. Schools with missing prior inspection and/or school test score percentile are dropped. Missing dummies included for percent students free lunch and percent minority students. Schools' national test score percentile rank calculated using published data on prior two years' mean for percent of students reaching the Level 4 threshold on the age-11 Key Stage 2 test. School value added percentile calculated using published data on prior year's mean English and Mathematics value added scores. Value added available from 2003 onwards only and so can only be included for 2008 inspection analysis. See text for further details.

Table 5 Preliminary Evidence on Changes in Inspection Ratings and Mean Reversion in Test Scores

(Each cell shows gain or loss in test score percentile)

	Change in inspection rating				
	-2	-1	0	1	2
Panel A: all schools					
Gain / loss in test score percentile:					
Between inspections (tests just before each insp)	-12.67	-6.28	1.19	10.47	20.46
	(0.99)	(0.39)	(0.29)	(0.55)	(2.39)
Tests in year of latest inspection vs. just before	0.16	0.16	-1.67	-2.13	1.47
	(1.20)	(0.45)	(0.33)	(0.59)	(2.17)
Tests one year after latest insp vs. just before	1.59	0.35	-1.68	-2.84	0.69
	(1.22)	(0.48)	(0.36)	(0.62)	(2.12)
Tests one and two years after latest insp vs. just before	1.65	0.21	-1.64	-3.46	1.52
	(1.14)	(0.43)	(0.33)	(0.57)	(1.97)
N (# schools)	317	2,364	3,903	1,311	84
Panel B: June inspected schools only					
Gain / loss in test score percentile:					
Between inspections (tests just before each insp)	-13.18	-6.22	1.59	9.61	14.08
	(3.89)	(1.06)	(0.90)	(1.60)	(5.20)
Tests in year of latest inspection vs. just before	-3.59	-1.63	-2.08	-0.82	-5.04
	(3.70)	(1.26)	(0.95)	(1.81)	(3.85)
N (# schools)	29	291	447	160	12

Notes:

Table 6 Inspection Ratings and Forecasts of the Change in Test Performance
(Outcome: Change in test score percentile rank; 'post' test taken before inspection, results revealed after inspection)

	(1)	(2)	(3)	(4)
Changes (between current and previous inspection) in:				
Inspection rating		3.07***	3.06***	2.70***
		(0.949)	(0.947)	(0.966)
School test percentile	-0.335***	-0.372***	-0.375***	-0.164***
	(0.038)	(0.040)	(0.040)	(0.046)
Percent free lunch percentile			-0.127	0.040
			(0.098)	(0.096)
Percent minorites percentile			0.003	0.035
			(0.050)	(0.055)
Log size of school			-3.054	1.785
			(4.936)	(5.109)
Further control variables, in levels	NO	NO	NO	YES
Observations	939	939	939	939
R-squared	0.242	0.252	0.255	0.317

Notes:

Table 7 Inspection Ratings and House Prices: Basic Results

(Outcome variable: log house prices)

	(1)	(2)	(3)	(4) Placebo 1	(5) Placebo 2
Change in rating * Post year	0.00382*	0.00515***	0.00481**	0.00003	-0.00125
	(0.00218)	(0.00198)	(0.00197)	(0.00186)	(0.00215)
School test score percentile			0.00007	-0.00003	0.00009
			(0.00008)	(0.00007)	(0.00008)
School value added percentile			0.00007	0.00001	-0.00005
			(0.00006)	(0.00006)	(0.00007)
Percent students free lunch, percentile			-0.00023	-0.00022	-0.00002
			(0.00016)	(0.00015)	(0.00018)
Property characteristics	No	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes
Observations	486,221	486,221	486,221	498,974	503,785
# schools	8,287	8,287	8,287	8,416	8,368
R-squared	0.61984	0.74914	0.74915	0.74991	0.75421

Notes: Standard errors reported in parentheses; *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. S.e.'s clustered at the school level. All regressions include year and month dummies. Change in inspection rating variable ranges from -2 to +2. All properties located within 500m radius of schools inspected between 2006 and 2008. Schools' national test score percentile rank calculated using published data on prior year's mean for percent of students reaching the Level 4 threshold on the age-11 Key Stage 2 test; value added percentile calculated using published data on prior year's mean English and Mathematics value added scores. Columns 3 to 5 also include school's percent minority students percentile rank. Columns 1 to 3: sample consists of transactions in the months straight after inspection (post = 1) and the year before inspection (post = 0). Column 4: transactions from the months immediately prior to the inspection (post = 1) and the year before inspection (post = 0). Column 5: transactions from the year before inspection (post = 1) and two years before inspection (post = 0). Property characteristics include dummies for whether the property is detached, semi-detached, terrace or flat; freehold or leasehold; and newbuild or not. See text for further details.

Table 8 Heterogenous Effects

(Outcome variable: log house prices)

	(1)	(2)	(3)
Change in rating * Post year	0.00486**	0.01469***	0.00067
	(0.00197)	(0.00426)	(0.00514)
Change in rating * Post year		-0.00017**	
* Prior free lunch percentile		(0.00008)	
Change in rating * Post year			0.00008
* Prior school test score percentile			(0.00008)
Property charachteristics	Yes	Yes	Yes
Time varying school controls	Yes	Yes	Yes
School FE	Yes	Yes	Yes
Observations	484,426	484,426	484,426
# schools	8,204	8,204	8,204
R-squared	0.74939	0.74940	0.74939

Notes: Standard errors reported in parentheses; *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. S.e.'s clustered at the school level. All schools inspected between 2006 and 2008. Schools' prior free lunch percentile rank calculated using data from 2004 and 2005. Schools' prior test score percentile rank calculated using age-11 English and Mathematics Key Satge 2 performance data from 2004 and 2005. Scools with missing prior free lunch or test score data dropped. All regressions include year and month dummies. For full set of property and school controls, see notes to prior table and main text.

Table 9 Medium-Term Effects

(Outcome variable: log house prices)

	(1)	(2)	(3)	(4)	(5) Placebo
Change in rating * Post year	0.00392**	0.00482***	0.00478***	0.00410**	0.00088
	(0.00191)	(0.00171)	(0.00172)	(0.00175)	(0.00160)
Change in rating * Post year				0.00159	
* Second year				(0.00162)	
School test score percentile			0.00005	0.00005	0.00010
			(0.00005)	(0.00005)	(0.00007)
School value added percentile			0.00003	0.00003	-0.00004
			(0.00004)	(0.00004)	(0.00006)
Percent students free			-0.00020*	-0.00020*	-0.00013
lunch, percentile			(0.00011)	(0.00011)	(0.00013)
Property characteristics	No	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes
Observations	833,600	833,600	833,600	833,600	749,850
# schools	9,552	9,552	9,552	9,552	9,414
R-squared	0.61259	0.74310	0.74311	0.74311	0.74998

Notes: Standard errors reported in parentheses; *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. S.e.'s clustered at the school level. All regressions include year and month dummies. 'Second year' dummy turned on for property transactions two years after the inspection. Columns 1 to 4: sample consists of transactions one and two years after inspection (post = 1) and the year before inspection (post = 0). Column 5: one year before inspection (post = 1) and two years before inspection (post = 0). Note that number of schools is larger in Table 3 than Table XXX because in the latter case, schools inspected very late in the academic year will not yield any property transactions for the post year (e.g. for July inspected schools) – see text and Figure 1 for details. See notes to Table XXX and text for further details of the regression models.

Table 10 Descriptive statistics, ranked preferences data

	Full sample	No sibling in primary school
Number preferences listed	2.32	2.71
	(1.25)	(1.20)
Number of available schools in borough (choice set)	26.4	26.4
	(1.08)	(1.07)
Older sibling in primary school	0.39	0
	(0.49)	
Free lunch status	0.24	0.22
	(0.42)	(0.42)
<u>First choice school:</u>		
Distance from home (km)	1.15	1.14
	(1.51)	(1.47)
Distance rank	2.92	2.86
	(3.81)	(3.64)
Year of inspection, 2006-2008 inspection cycle	2007.01	2007.01
	(0.76)	(0.76)
Change in rating (2006-2008 previous inspection)	-0.28	-0.28
	(0.96)	(0.96)
English and Mathematics decile	5.25	5.25
	(2.55)	(2.55)
% Students eligible for free lunch, decile	5.92	5.92
	(2.28)	(2.28)
Number of applicants	10,081	6,172

Notes: Standard deviations in parentheses . Data from applications made in the fall of 2004 to 2008. Distance measured in straight line from applicant's home to school. A school is 'available' if it is in the applicant's choice set (see main text). English and Mathematics performance measure corresponds to the proportion of students attaining the government attainment target (Level 4) for age-11 (Year 6) students on the official (Key Stage 2) English and Math test; averaged over the two academic years prior to application. Percent students eligible for free lunch also averaged over the two academic years prior to application. ‘Decile’ refers to the school’s position in the borough-level distribution of the performance measure. Applicants who missed the application deadline are excluded. See data appendix for further details.

Table 11 School Choice Results: Conditinal Logit Model

(Outcome: first choice school)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	No school fixed effects		Basic	Controls	Poverty Interaction	Sample incl. older sibling at a primary	Placebo (applications prior to treatment)
Change in rating * post year	0.053**	0.155***	0.104***	0.100***	0.109***	0.064**	0.013
	(0.024)	(0.026)	(0.036)	(0.036)	(0.037)	(0.027)	(0.029)
Change in rating * post year * free lunch status					0.066*		
					(0.040)		
Distance	-1.739***	-1.671***	-1.473***	-1.473***	-1.472***	-1.449***	-1.223***
	(0.023)	(0.024)	(0.025)	(0.025)	(0.025)	(0.019)	(0.019)
Distance-squared	0.084***	0.083***	0.073***	0.073***	0.073***	0.072***	0.064***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
School test score decile		0.167***		0.014	0.014	-0.001	0.006
		(0.007)		(0.015)	(0.015)	(0.012)	(0.015)
Percent students free lunch, decile		-0.245***		-0.049	-0.048	-0.030	-0.012
		(0.009)		(0.034)	(0.034)	(0.026)	(0.039)
School fixed effects	No	No	Yes	Yes	Yes	Yes	Yes
# students (applicants)	6,172	6,172	6,172	6,172	6,168	10,081	8,069
Observations	162,936	162,936	162,936	162,936	162,832	265,670	215,679

Table 12 School Enrolment

(Outcome: log(enrolment))

	(1) Basic	(2) Controls	(3) School SES interaction	(4) Placebo (enrolment prior to inspection)
Change in rating * post year	0.0164***	0.0164***	0.0123***	-0.00032
	(0.00101)	(0.00101)	(0.00208)	(0.00094)
Change in ratig * post year * percent students free lunch, decile			0.000868**	
			(0.000361)	
Test score decile		0.000665*	0.000541	0.000317
		(0.000348)	(0.000344)	(0.000416)
Percent students free lunch decile		-0.006781***	-0.005944***	-0.010167***
		(0.000739)	(0.000728)	(0.000979)
School fixed effects	Yes	Yes	Yes	Yes
Observations	46,736	46,736	46,712	35,232
Number of schools	11,747	11,747	11,747	11,747
R-squared	0.988	0.988	0.988	0.993