Not Just Test Scores: Parents' Demand Response to School Quality Information

Iftikhar Hussain[1]

University of Sussex

iftikhar.hussain@sussex.ac.uk

December 2013

Abstract

There is scant evidence on the effects of providing school quality information, other than test scores, on parents' school choice decisions. This paper investigates the causal effects of a novel measure of quality, school inspection ratings. Using variation in the timing of inspections, I demonstrate that a school's market share, measured by total enrollment, responds to the top and bottom ratings. Next, using data on parents' ranked preferences over local schools, and exploiting the gradual rollout of a policy reform which led to major simplifications in the presentation style of the reports, the paper estimates a random utility model. The results show that there is a strong response to all ratings, not just those at the extreme, suggesting that families discriminate between the majority of schools located in the middle of the quality distribution. These effects are very large relative to parents' response to school test scores.

JEL I2; D8; D1

Key Words: school choice; information; school inspections; school quality

[1] For helpful discussions and comments I would like to thank Kehinde Ajayi, Richard Dickens, Julien Grenet, Edwin Leuven, Steve Machin, Doug Staiger, and seminar and conference participants at the IZA-Bonn, NIESR-London, University of Kent, University of Sussex, the CEPR Public Policy Symposium-Lugano, Educational Governance and Finance Workshop-Trondheim, Matching in Practice workshop- Paris, and the 2013 Society of Labor Economists annual meeting.

1. Introduction

Whether parents care about and respond to school quality information is an active and important research area for at least two reasons. First, it is commonly recognised in the literature that the effectiveness of demand side pressure in raising schools' performance depends partly on how parents respond to information on school quality (e.g. Hastings et. al, 2009). Second, and arguably less emphasized in the current literature, is the fact that releasing information on certain dimensions of school quality and subsequent parental response may have important consequences for the degree of sorting and stratification among schools and neighbourhoods.[2] The evidence suggests that parents respond to (relatively) easily observable characteristics such as a school's racial and SES (socioeconomic status) composition, as well as published information on, for example, the school's test score performance.[3] Thus releasing test score information, for example, has the potential to exacerbate stratification. This would be the case if parents care about schools' SES composition but observe this measure imperfectly. Disclosing test scores may yield more information on SES, assuming test scores are sufficiently correlated with SES (Mizala et. al, 2007), and thus enable greater sorting, even if parents do not value test scores *per se*.

Recent studies have investigated the effects of disclosing one measure of school effectiveness, a school's test score *value added*, but the balance of evidence suggests modest or negligible market response (Fiva and Kirkeboen, 2011, Imberman and

[2] Theoretical models emphasize sorting on the basis of income in the presence of peer effects (e.g. Nechyba, 2000, and Epple and Romano, 2003).

[3] The limited research that exists in this area has tended to focus on test scores – or report cards derived from test scores – as proxies for school quality. See, for example, Figlio and Lucas (2004), Hastings and Weinstein (2008), Burgess et. al (2009). One limitation of using test scores as a measure of quality is that they may reflect students' social background rather than quality per se.

Lovenheim, 2013, Mizala and Urquiola, 2013). One interpretation of these findings is that parents pay substantially more attention to peer quality than to school effectiveness (Rothstein, 2006).

This study focuses on the demand response to a novel measure of quality, school inspection ratings, produced by independent evaluators. As explained in detail below, these ratings reflect soft (in-class observation of teaching) as well as hard (test scores) measures of a school's performance. Arguably, these ratings better capture the multifaceted nature of education production which parents, as well as policymakers, likely care about – such as richness of instruction and curriculum, enjoyment of lessons and classroom behavior – than test scores alone. [4]

In addition, as shown below, although there is some correlation between evaluator ratings and both measures of SES as well as school test rank, this correlation is far from perfect. For example, a substantial fraction of schools with high concentrations of low SES students receive the best ratings. Thus under certain circumstances, publicising these ratings has the potential to *reduce* the impact of forces of stratification. Whether this is indeed the case depends on the informativeness of ratings and to what extent parents are prepared to trade-off better SES composition (and higher test rank) for a school with a higher rating.

I exploit a natural experiment to evaluate the causal effect of the ratings on parents' school choice decisions. The setting is the English public (state) school system, where parents also have relatively easy access to test score information. The estimated effect of the inspection rating is thus over and above any reaction to test score information. Another feature of the study is that ratings are made available in the form of inspection reports,

[4] See, for example, Schneider and Buckley (2002).

freely available in the public realm via the internet. Thus the effects identified in this study are less susceptible to concerns about saliency and suggestion which may arise in a field experiment setting where information on school quality is presented directly to participants.[5]

The empirical strategy exploits variation in the timing of inspections (the inspection body does not consider it cost-effective to inspect every year). This institutional feature yields comparisons between early and late inspected schools. I demonstrate that timing of inspections is exogenous.

Using school-level data for a panel of all schools in England I first investigate the effect of ratings on enrollment. The results show that enrollment at primary schools declines by around 4 percent the year after a fail rating (the worst rating, on a scale of 1 to 4). The effect of receiving the top rating ('Outstanding') boosts enrollment by around 3 percent, rising to 6 percent in communities with fewer capacity constraints (i.e. municipal areas experiencing especially rapid declines in total enrollment numbers).

The main part of the paper is concerned with estimating consumer demand using data on parents' ranked preferences for primary schools from a London borough. Importantly, using the borough's assignment rules, I am able to determine the individual choice set each family faces.[6] This analysis yields three key findings. First, there is a strong parental response to inspection ratings. For example, the results imply a rise of 50 percent in the distance parents are willing to travel in order to attend a school with a one

[5] The research design employed in this study identifies the total effect of the inspection reports on parents' decision making. This includes the direct effect as well indirect effects, which may, for example, be mediated through social networks.

[6] As described below, the borough uses a form of the deferred acceptance algorithm where distance between home and school determines whether one family receives priority over another. See Pathak and Sonmez (2011) for an overview of school choice assignment mechanisms in England.

unit improvement in the inspection rating (i.e. a Good-rated school versus one rated Satisfactory).

In order to overcome the concern that inspection ratings may be correlated with omitted variables such as a school's reputation, I evaluate the effects of a policy reform which substantially simplified the presentation style of the inspection reports.[7] This reform is rolled out gradually due to the fact that at the time of the reform, inspections are on a three-year cycle. This gradual rollout helps identify the additional effect of simplifying the way in which the inspection outcomes are presented to consumers. The results from this analysis imply that the new, simplified inspection reporting system has large effects on parents' choices.

The second key finding is that the estimated impact of school test scores on parental demand shrinks dramatically once inspection ratings are included in the regression model. One interpretation of this result is that when measures of school quality such as inspection ratings are available, parents place much smaller weight on test scores than previously thought (e.g. Hastings et al, 2009, and Burgess et. al, 2009).

Finally, the demand estimates also demonstrate that families of non-free lunch status children are willing to accept a rise of around one to two decile points in the percent of students eligible for free lunch for a one unit improvement in the inspection rating. This result suggests that school inspection reports and ratings may help reduce stratification by income in the English schooling system.

The remainder of this paper is laid out as follows. Section 2 describes the context for this study. Section 3 reports the school-level analysis of effects of inspection ratings on

[7] Reports for schools inspected from 2005/06 onwards have a headline grade reported at the beginning of the report; prior to this, ratings had to be deciphered in the text, with no highlighting of the overall inspector assessment.

enrollment. Section 4 lays out the data and empirical strategy for the student-level school choice model. Section 5 presents the results of this analysis and section 6 concludes.

2. Context

See Hussain (2011) for details of the English school inspection system. I exploit the reform implemented from September 2005 onwards which led to inspection reports being produced in a substantially simplified format. Prior to this reform, reports were very dense, with no clear indication of the overall inspection rating either in numerical format or highlighted within the text of the document. Following the reform, reports had a headline numeric rating (from 1 to 4) at the front of the main inspection report, with plain text explaining the range of ratings and their meanings. (For examples, see Appendix xx.) [TO BE COMPLETED]

3. The effect of inspection ratings on enrollment

3.1 Empirical strategy

The empirical challenge in identifying the effect of a Fail inspection rating, say, on total school enrollment is that poorly performing schools may be contracting even in the absence of disclosure of the inspection rating. Thus simple regression analysis, even with a panel of schools, may yield severely biased estimates of the true effect of inspection ratings.

This study exploits variation in the timing of inspections to identify the causal effect of inspection ratings. This allows for a comparison of enrollment outcomes for

early and late inspected schools. Figure 1 illustrates the main idea. The figure depicts two sets of schools which both receive Outstanding ratings, the best of the four possible ratings, in 2006 or 2008.[8] Schools inspected and rated Outstanding in 2006 are the treatment group, whilst schools inspected and rated Outstanding in 2008 are the control group. The outcome, enrollment, is measured in the post-treatment period, 2007. A comparison of outcomes for these two groups then yields the effect of receiving the Outstanding rating.

Importantly, the evidence suggests that timing of inspections is exogenous. Over the period covered by this analysis, schools are typically inspected once in a given inspection cycle.[9] Inspection cycles last between three and five years. For a given school the timing of inspection within a cycle is a function of the timing of its inspection in the previous cycle. I.e. schools inspected early in previous inspection cycles are also inspected early in later inspection cycles. The descriptive statistics in Table 1 shed some further light on this.

Panel A of Table 1 shows baseline characteristics for Schools inspected and rated Outstanding in 2006 or 2008. The first row demonstrates that, on average, schools inspected in 2006 were inspected in 2000 in the previous inspection cycle, and schools inspected in 2008 were previously inspected in 2003.[10] This evidence (and in Panel B, discussed below) supports the idea that inspectors use an exogenous rule to determine timing of inspections.[11] Furthermore, Panel A shows that these two groups of schools appear to be comparable on a broad set of observable characteristics: there are no are not

[8] 2006 refers to the academic year 2005/06 and 2008 to 2007/08.
[9] In the most recent years (i.e. in inspection cycles after the 2006-2008 cycle used in the current analysis), the inspection regime switched to one where timing or frequency of inspections is determined partly by past performance.
[10] The gap in the timing of inspections for the two sets of schools is 3 years in the previous inspection round, whilst the gap in the current inspection round is 2 years. This is a consequence of the fact that the previous inspection cycle took place over 5 years, whilst the current one is 3 years, running from 2006 to 2008.
[11] Hussain (2012) shows that even *within* a year, schools inspected in the early part of the academic year were inspected somewhat earlier in the previous inspection cycle than schools inspected later in the academic year.

statistically significant differences in the inspection rating in the previous inspection round; the proportion of students eligible for free lunch, the proportion of students who are white British; and total enrollment. There are statistically significant differences between the two groups in prior test scores. However, as discussed in Hussain (2012), inspectors appear to put substantial weight on test score performance and some of the high (low) performance prior to receiving a good (poor) rating from inspectors likely reflects good (bad) luck. This interpretation is supported by Appendix Table 1 which shows trends in test scores for the two sets of schools before and after inspection. For both sets of Outstanding schools, test scores peak in the year before inspection.[12] In the regression analysis below I demonstrate that the estimates (for the Outstanding as well as the Fail treatment) are robust to these differential trends in test scores.

Panel B paints a very similar picture for the Fail category of schools: the timing of early and late failed schools can be explained by the timing of inspections in the previous round; the two groups are balanced on all pre-treatment covariates except test scores; when we compare test scores in the year before inspection for both groups (Appendix Table 1) the two groups appear to be very similar.[13]

This empirical strategy is implemented using difference-in-differences models.[14] For example, in order to estimate the effect of a school receiving an Outstanding rating, I select on those schools rated Outstanding in 2006 or 2008. The unit of observation is the school, and the treatment effect is identified by comparing the change between 2005 and

[12] Appendix Table 1 shows that when we compare test performance in the year before inspection for both sets of schools, there is no difference between the treatment and control groups. Thus, for schools receiving an Outstanding rating in 2008, the percent of students attaining competency in 2007 is 87.5 percent, almost identical to the 2004 mean for schools rated Outstanding in 2005. Similar conclusions hold for the Fail rating.

[13] One additional noteworthy point in Table 1 is that there is no evidence to suggest that inspectors bring forward the inspection for fail schools: the average year of previous inspection is almost identical for Outstanding and Fail schools inspected in 2006 as well as for those inspected in 2008.

[14] Although baseline differences in enrollment levels between early and late inspected schools are not statistically significant, DID models are employed to account for any remaining differences.

2007 in the log of enrollment for early and late inspected schools. Specifically, the following DID model is estimated:

$$y_{st} = X'_{st}\beta + \gamma D_{st} + \delta.post_{07} + \lambda.EarlyInspected_s + u_{st}, \qquad (1)$$

where y_{st} is log enrollment at school *s* in year *t*. The treatment dummy, D_{st} is switched on in the post period, i.e. in 2007, for schools rated Outstanding in 2006, and set to zero otherwise. The parameter of interest, γ, is the average effect of treatment on the treated. [15] X_{st} is a vector time-varying school characteristics, $post_{07}$ is the post dummy, switched on in 2007, and $EarlyInspected_s$ is a dummy switched on for schools inspected in 2006. The residual, u_{st}, is assumed to be uncorrelated with D_{st}, conditional on the other covariates. In some regression results below, school fixed effects are also included.

The key identifying assumption is that in the absence of an Outstanding rating in 2006 for the early inspected schools, the trend in enrollment between 2005 and 2007 for these schools would have been that observed for schools rated Outstanding in 2008. I can probe this assumption by exploring whether this common trends assumption holds in the pre-treatment period.

3.2 Enrollment Results

Table 2, Panel A shows results for the effect of an Outstanding rating on enrollment one year after disclosure. The first row ('2007 x early inspected') reports estimates of the

[15] In the current setting the average effect of the treatment on the treated is the effect of publicly disclosing schools which inspectors believe to be the best. Another treatment effect, the average treatment effect of an Outstanding rating (the effect of declaring a school at random to be Outstanding), is unlikely to be policy relevant.

treatment effect whilst the second row ('2007') in the table corresponds to the 'post' dummy in equation (1) above. Column 1 reports the basic DID result, without any school fixed effects or time-varying controls. This suggests that the effect of an Outstanding rating is to raise enrollment by 2.55 percent, which is statistically significant at the 1 percent level. Columns 2 and 3, which add school fixed effects, test score performance and other time-varying controls, leave this basic estimate virtually unchanged.[16]

Columns 1 to 3 of Panel B, Table 2 report the effect of a Fail inspection. These results suggest that on average, a school shrinks by 4.39 percent the year after being declared a Fail school. In addition, it is worth noting that for Fail schools the coefficient on the 'post' dummy – identified off changes in enrollment experienced by the control group – is large (-5.27 percent) and statistically significant. This suggests that fail-type schools experience relatively large declines even in the absence of being publicly disclosed as Fail schools; the treatment leads to further decline in student numbers.[17]

These estimates of Outstanding and Fail ratings may be underestimates of the underlying demand response if good schools cannot expand in the short run due to lack of physical space or if there is a lack of good alternatives to Fail schools.[18] Enrollment effects may be larger when such capacity constraints are relaxed. In order to do explore this, I undertake separate analysis for schools located in areas experiencing relatively low population growth. Over the period 2005 to 2007 the median primary school enrollment growth rate for English Local Authorities (the local public school jurisdiction) was *minus*

[16] The other time-varying controls are the proportion of students eligible for free lunch and the proportion of white British students.
[17] This result also demonstrates that a simple school fixed effect analysis would lead to severely upward biased (in absolute terms) estimates of the effect of a Fail treatment.
[18] Although Besley and Machin (2010) have shown that principals at high performing schools may be rewarded in the labour market, anecdotal evidence suggests that incentives for public schools to expand in England remain weak and are often resisted.

3.5 percent.[19] The fact that the inspection ratings have *any* effect on enrollment (as opposed to demand) may in part be a function of this demographic feature: if Local Authorities respond slowly to declines in the local student population then this may lead to some spare capacity in the system, leading to greater responsiveness in allocation of school seats to parental preferences.

Column 4 of Panel A (Panel B), Table 2 reports results for the effect of an Outstanding (Fail) rating for schools located in Local Authorities where enrollment declines by more than that for the median Local Authority between 2005 and 2007. Column 5 shows estimates for schools in Local Authorities where enrollment growth is in the bottom quartile over this period (growth at the 25th percentile Local Authority is -5.5 percent).

The results in Panel A show that the response to an Outstanding rating is substantially larger in those jurisdictions experiencing especially large falls in enrollment: the positive enrollment effect is estimated to be around 4 percent and 6 percent in columns 4 and 5, respectively. These results suggest that response to an Outstanding rating is strongest where there is greater spare capacity. Conversely, there would appear to be substantial pent up demand in those jurisdictions where spare capacity is much more limited.

The results for the Fail rating in columns 4 and 5 of Panel B, on the other hand, do not point to substantial variation in treatment effect by enrollment growth in the local area. This may be in part because these schools experience very large falls in enrollment even in the absence of the Fail treatment. As the 'post' dummy indicates, enrollment at the control set of schools falls by around 7 percent. The effect of a Fail rating is to increase this by

[19] The two-year average of the primary school student population growth rate in England over the period 2003 to 2007 ranges between -3 percent and -4 percent.

another 4 percentage points, leading to a total decline in enrollment of 11 percent for the treatment group.

Table 3 reports results form a falsification exercise. The question addressed is whether there is any evidence of a 'treatment effect' in the two years before disclosure of the reports. For example, for the Outstanding rating, as before, schools rated Outstanding in 2006 or 2008 are selected for the analysis. This time, however, enrollment data for the regression analysis are taken from 2003 and 2005. The first row in Table 3 shows that there is a small and statistically insignificant difference in the rise in enrollment for the early inspected (2006) schools versus the late inspected (2008) schools between 2003 and 2005. These results hold across all specifications, both for the Outstanding treatment (Panel A) and the Fail treatment (Panel B). This evidence lends credibility to the DID assumption that in the absence of treatment, trends in enrollment would have been same for the treatment and control groups.

Appendix Table A2 reports results for the ratings 'Good', or grade 2, and 'Satisfactory', or grade 3. These results suggest that there is no enrollment response to disclosing a school to be Good or Satisfactory: all the estimated treatment effects are close to zero, statistically insignificant and relatively precisely estimated. These results suggest that, at least in the very short term, there is no enrollment response to these two middle ratings.

4. Student-level analysis: data and empirical strategy

A limitation of the school-level analysis above is that it does not shed light on how parents' response to the inspection ratings varies with the availability of possible alternatives. In

this section I estimate a conditional logit model using individual-level school choice data in order to better understand the tradeoffs consumers face.

4.1 Data and descriptive statistics

The data are from a London borough, and consist of parents' ranked preferences for primary schools from applications made in the fall of 2006, 2007 and 2008. The school-level information on geographical location, test score performance, percent free lunch and inspection ratings are sourced from administrative data as described in section 4. 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.[20] This latter piece of information 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. For secular schools facing excess demand, children living closest to the school are given priority. For religious schools, spots in the school are allocated on the basis of religious affiliation.

In order to determine whether a given secular school facing excess demand is in a child's choice set we need to determine whether 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 can be determined using information on the child without special needs or a sibling, living furthest away from the school who was allocated a place. For religious schools, whether the school is in the

[20] 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.

child's choice set cannot be determined without information on religious affiliation. Consequently, religious schools and students who apply to a religious school are dropped from the analysis.[21]

Table 4 shows mean characteristics of the first choice school (column 1), the three nearest to the child's home which are available to her (column 2) and all schools available to the child in the borough (column 3). A school is deemed to be unavailable if the child does not have a sibling at the school and she resides outside the cutoff distance for the school. On average, the first choice school is 1.1 km from the child's home and its distance rank (over all available schools) is 2.9. Compared to the mean for the nearest three schools in applicants' choice sets, 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.[22] Appendix Table 3 shows the characteristics of the applicants. 58 percent of students are non-white British; 42 percent have an older sibling in one of the four ranked primary schools; and on average, applicants have 30 schools available to them to choose from in the borough.

Finally, Table 5 shows the distribution of the 2006 inspection ratings by the school's SES composition (proxied by the percent of free lunch students, Panel A) and school test rank (Panel B, test rank shown in reverse order). These clearly show that although there is some correlation between the quality of the school as measured by SES or test rank and inspection ratings, this is far from unity. For example, Panel A shows that around 50 (30) percent of schools ranked in the third and fourth (fifth) SES quintiles are

[21] Special needs students are also dropped from the sample. See the data appendix for details.
[22] Measures of school test performance and the proportion of students eligible for free lunch are taken from the academic year prior to the academic year in which parents make their applications. Inspection ratings are the latest available from the academic year before application or earlier since schools are not inspected every year.

rated Good or Outstanding by the inspectors. Similarly high proportions of schools ranked in the lower half of the test score distribution (Panel B) receive one of the two best ratings.

4.2 Conditional logit model and identification strategy

The approach to the empirical analysis is the standard conditional logit model (McFadden, 1974). Parents of student i are assumed to choose from the available set of schools, $j \in \{1, 2, .., J\}$, in order to maximize utility,

$$U_{ij} = \sum_r \delta_r D_{jr} + x'_{ij}\beta + e_{ij}.$$

The deterministic part of utility is represented by school j's inspection rating, captured by dummies for each rating r, D_{jr}, and $x'_{ij}\beta$, where x'_{ij} represents a vector of school characteristics including test scores, the percentage of students eligible for free lunch and total enrollment, as well as distance from i's home. The error term e_{ij} 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

$$\Pr(Y_i = j \mid x'_{ij}) = \frac{\exp\left(\delta_1 D_{j1} + \delta_3 D_{j3} + \delta_4 D_{j4} + x'_{ij}\beta\right)}{\sum_l \exp\left(\delta_1 D_{l1} + \delta_3 D_{l3} + \delta_4 D_{l4} + x'_{il}\beta\right)} \qquad (1),$$

where D_{j1}, D_{j3} and D_{j4} represent dummies for inspection ratings Outstanding, Satisfactory and Fail, respectively; the base (omitted) category is Good. In attempting to identify the effect of inspections ratings in the above model, omitted variable bias may be a potential

concern. For example, demand for a school rated Outstanding may be relatively high, conditional on observable school characteristics, even in the absence of the inspection rating.

In order to address such concerns, I also estimate the *additional* effect on demand of simplifying the presentation style of the post-September 2005 reports. The empirical strategy exploits the fact that new style reports are introduced gradually, rather than all at once, over the inspection cycle starting in 2005/06.[23] Thus families submitting their applications in fall 2006, for example, will have access to old style reports for some schools and new style reports for other schools. The model now includes a dummy for each rating as well as a rating * new-style-report interaction term. The latter variable identifies the additional effect of a new style (i.e. simplified) report. Under this setup, the probability that student *i* chooses school *j* is given by

$$\Pr(Y_i = j \mid x'_{ij}) = \frac{\exp\left(\sum_{r=1,3,4} \delta_r D_{jr} + \lambda New_j + \sum_{r=1,3,4} \gamma_r D_{jr} * New_j + x'_{ij}\beta\right)}{\sum_l \exp\left(\sum_{r=1,3,4} \delta_r D_{lr} + \lambda New_l + \sum_{r=1,3,4} \gamma_r D_{lr} * New_l + x'_{il}\beta\right)} \qquad (2),$$

where $D_{jr} * New_j$ represents the interaction between the inspection rating dummy for school *j* and whether the rating is reported in a new style report.

To gain some intuition for this approach, consider, for example, a family with two nearby schools, one rated Outstanding in the old style report and the other rated Outstanding in the new style report. Both schools are excellent schools; the only difference is that the information on one is more transparent than for the other. The

[23] This gradual phase-in of new style reports is a natural consequence of the fact that schools are not inspected every year.

coefficient on the Outstanding dummy, δ_1, represents the demand for a high quality school, which may be a consequence of both receiving an outstanding rating in the old style inspection report, as well as correlated unobservables, such as reputation of the school. The coefficient γ_1 on the interaction term, $D_{j1} * New_j$, is the parameter of interest and it identifies the additional effect of simplifying the reports on consumer demand. In the robustness analysis below, I am able to rule out other possible explanations, including the possibility that families react more to more recent reports.

Table 5 highlights the variation in the school-level data which permits identification of the parameters in model (2). The first column in Table 5 shows that by the end of 2005/06, the first year in which the simplified reports are produced, 13 schools had new style reports. By 2007/08, nearly all schools are rated under the new style reporting system.

5. School choice results

5.1 Main results

Column 1 of Table 6 reports results from the basic conditional logit model of first choice school without any indicators for inspection ratings. This in effect reproduces the traditional choice model in the literature, where choice depends on distance between home and school, the school's performance on test scores, the percent of students eligible for free lunch and the ethnic composition of the student body (percent white British), as well as the latter variable interacted with applicant's own ethnic status. 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.

Column 2 of Table 6 reports estimates from the model incorporating schools' inspection ratings. The three inspection dummies included in the model are Outstanding, Satisfactory and Fail and the omitted category is for the Good rating, which lies between Outstanding and Satisfactory. The results in column 2 demonstrate that parents make a sharp distinction between all four categories.

The value families place on school characteristics such as ratings can be measured in terms of the tradeoffs they are willing to make with respect to extra travel distance. [24] A coefficient of -0.61 for the Satisfactory rating, significant at the 1% level, implies a willingness to travel an extra 0.5km (assuming an initial travel distance of 1km) in order to attend a Good-rated school instead of a Satisfactory school located 1km from home. This is a large effect when compared to the average distance for families' first choice school, around 1km. The coefficient for the Outstanding rating suggests some small positive preference for these schools relative to Good schools, but the estimate is not statistically significant. The effect for the Fail rating implies that there is strong aversion to these schools.

As the discussion in section 5.2 highlighted, the estimated effects of ratings reported in column 2 may be subject to omitted variables bias if factors such as school reputation are not adequately captured by test scores and the free lunch measure of student SES. Column 3 exploits the gradual rollout of the simplified reports in order to identify the differential effect of the new style reports on consumer demand. The interaction between Satisfactory, say, and a new style report represents the additional response to the

[24] 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.

more salient simpler style reports. As explained earlier, if the coefficient on the old style Satisfactory rating represents parents' response to the differences in, for example, school reputation between Good and Satisfactory schools, rather than the ratings *per se*, then there should be no additional effect of a new style rating, *unless* parents are responding to the information in this type of report.

A key finding from the results in column 3 is that for the Outstanding and Satisfactory ratings, simplification has substantial effects on demand. The results suggest that the effects of a Satisfactory rating are driven mostly by the new style reporting system. Although a statistically significant coefficient of -0.33 suggests that there is lower demand for schools rated Satisfactory in the old style reports relative to schools rated Good (also in the old style reports), a new style Satisfactory rating results in additional disutility of -0.41.[25] Similarly, there are relatively large positive demand effects of a school receiving an Outstanding rating under the simplified reporting regime.

For the Fail rating however, there are no additional effects from the new style reports (the coefficient on the Fail interaction term is relatively small and statistically insignificant). This is consistent with the hypothesis that a Fail rating is 'big news' even when reported under the old regime and that further simplification of the reporting style has little effect on demand.

One final important finding in Table 6 relates to the effects of a school's test score performance on parents' choices. The results suggest that once inspection ratings are included in the model, the effect of test scores on choice is much diminished: the estimated coefficient on the school's test decile in columns 2 and 3 is fully two-thirds

[25] Note that the coefficient on the 'New style report' dummy, which represents the additional effect of simplifying the reports for schools rated Good (the omitted category), is close to zero. Hence we may compare the coefficient on the Satisfactory dummy with Satisfactory interacted with the new style dummy in order to infer the additional effect of simplifying the Satisfactory rating.

smaller than that in column 1. The coefficients of the distance variables (as well as the proportion of students on free lunch measure) exhibit very little change. Thus the implied marginal willingness to pay for test scores, in terms of distance travelled, is substantially smaller once we condition for inspection ratings.

This last result suggests that once we take account of aspects of school quality captured by the inspection ratings, parents appear to place much less weight on test scores than is implied by standard revealed preference analysis of parents' schooling decisions (e.g. Hastings et al, 2009, Burgess et. al, 2009). This finding has two possible (not necessarily mutually exclusive) implications. First, when parents have ready access to school quality information such as that made available in the inspection reports, the role of test scores is substantially diminished. Second, the effect of test scores may be overstated in prior studies which typically do not take account of such school quality measures.

5.2 Robustness checks

One alternative explanation for the effect of the new style inspection ratings is that these ratings are more recent and hence the large and significant effects for some of the interaction terms reported above simply reflect the larger response to a more up-to-date signal of quality. Furthermore, the most recent ratings may be more newsworthy and hence more salient to parents.[26] In such cases, the effect captured in the results of Table 6 is a causal response to the new style ratings, but this is not because of simplification of the

[26] For example, the latest inspection findings may be reported in the local press. Anecdotal evidence suggests that this does take place, especially when schools receive the worst – Fail – or best – Outstanding – outcomes.

new style reports, but rather, reflects the larger weight families attach to the most recent ratings.

Yet another interpretation of the results reported above is that the response to more recent (new style) ratings simply reflects changes in school quality over time. For example, a school currently rated Satisfactory in an old style report may have improved in the (relatively long) intervening period, whereas a school rated Satisfactory in the very recent past, and hence receiving a new style inspection report, may have changed little since the inspection. In this case the inspection rating understates actual quality for the former school, but accurately captures the status of quality for the latter school. Such changes in quality over time would then lead to the pattern of results observed in Table 6.

Tables 7 and 8 address these two issues. Consider first the idea that there is a larger consumer response to more up-to-date information. Column 2 in Table 7 reports results from a conditional logit model which now also includes a second set of interaction terms: the rating x new style report interaction term further interacted with time (years) since the new inspection took place. For example, for families applying in fall 2008 the three-way interaction term 'Satisfactory x new style report x years since new style inspection' for a school rated Satisfactory in the academic year 2005/06 is set to 3 (since the inspection took place up to 3 years prior to the application).[27]

The results in column 2 Table 7 show that the triple interaction term is statistically significant for only the Good rating.[28] For this case the results suggest that the demand boost from receiving a Good rating in the new style reports *increases* with the number of years since inspection. For the Outstanding rating, although this effect is not statistically

[27] There is no three-way interaction term for the Fail category since schools rated Fail in one year are re-inspected in the following year and graduate out of the Fail category.
[28] For ease of comparison, column 1 reproduces the results from the final column of Table 6.

significant, the coefficient estimates again suggest that demand rises with years elapsed since inspection. Possible explanations for a rising response over time include learning over time, say through social networks, as well as adjustment costs (e.g. if older siblings are enrolled in less desirable schools).

Table 8 assesses whether the differential response to old and new style inspection reports is a result of changes in school quality over time. For this exercise, the analysis focuses on those schools which do not experience a change in their rating between the old and new style inspections. Appendix Table 4 shows that there are 19 schools for which the rating between the old and new style inspections do not change.[29] Of these 19, two are rated Outstanding but exhibit no variation in the availability of old style and new style ratings in any of the three application years. Thus the analysis in Table 8 focuses on schools which were either rated Satisfactory in both the previous and current inspection cycle or rated Good in both cycles.[30]

For the regressions reported in Table 8, the 'No change in rating' dummy is switched on for the 17 schools experiencing the Satisfactory-to-Satisfactory or Good-to-Good transition. Definitions for the other variables are as before. For example, the 'No change x Satisfactory' interaction is switched on for the no change in rating schools which are rated Satisfactory whilst the omitted category consists of those schools experiencing no change in rating and are rated Good.

The coefficient on the 'No change in rating' in column 2 of Table 8 shows that there is a positive, though insignificant, demand response to a Good rating presented in an

[29] For example, for applications in fall 2007, there are 10 Good schools (column 2) of which 5 are rated Good in the new style reports. The remaining 5 will be rated Good in their subsequent inspections (as shown in column 3).

[30] To see why Outstanding schools are excluded, note that for families applying in fall 2006 there are two schools rated Outstanding in the new style reports but no schools rated Outstanding in the old style reports. Thus, there is no control group for new style Outstanding rating to identify the differential effect of simplifying the inspection reports.

old style report; relative to these schools, for the no-change-Satisfactory (old style) schools there is large and significant negative effect. The key finding in Table 8 is that there are large negative (positive) effects for a new style Satisfactory (Good) rating relative to the same rating in the old style report (rows 3 and 4 in column 2, respectively. Thus the key finding from the main set of results (column 3, Table 6) survives this robustness test.

5.3 Heterogeneous effects

Poverty status

Table 9 stratifies the sample by eligibility of free lunch, an indicator of poverty status. Column 1 suggests that disadvantaged families are responsive to school quality as captured by the inspection ratings: the coefficients on the Satisfactory and Fail ratings are relatively large and highly statistically significant. There is some evidence that poorer families respond to the simplified reports (column 3): the coefficient on the Satisfactory x new-style-report interaction term is -0.19, although the estimates are noisy due to the much smaller sample size for this group.

The response for non-poor families (i.e. those where the child is not eligible for free lunch) are between 50 percent and 100 percent larger than for disadvantaged families (column 2). The non-poor are also highly responsive to the simplified reports (column 4). Response to the simplified reports is substantially larger for the non-poor.

Finally, columns 2 and 4 suggest that non-poor families are prepared to trade-off higher ratings for worse SES composition (as measured by the coefficient on the school's percent free lunch decile).

Presence of older sibling

Table 10 stratifies the sample by whether there is an older sibling attending one of the schools listed by the parents on their choice form. These results suggest larger effects of both inspection ratings as well as schools' test score performance for parents reporting no older child attending one of the listed schools than for families which do.

As new information regarding inspection ratings and test score performance arrives, there are two possible explanations why these two sets of parents might react differently. First, those families with children already enrolled in a primary school likely face higher costs of selecting an alternative school for a second child than parents without an older child already enrolled in a primary school. Thus the latter group will appear to be more sensitive to school quality characteristics such as test scores and inspection ratings. A second explanation is that parents with a child already enrolled in a school may have better information on the quality of that school and may update their priors to a lesser extent than the second group of parents in response to new signals of quality from inspectors and test score results.

Mixed logit model

Table 12..[To be completed]

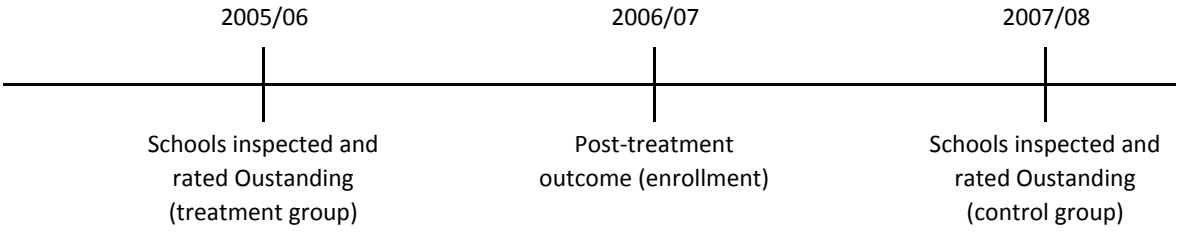
6. Conclusion

Whether providing parents with information on school quality other than test scores affects their school choices remains an open question. This paper attempts to close this gap in the literature by exploiting inspection ratings provided by independent assessors. The first set of results in this study demonstrate that schools do expand and contract in response to positive and negative ratings, respectively. But, as noted above, incentives for public schools to expand are weak and the results show that the effects of the ratings on enrollment are only discernible for ratings at the extreme. For the vast majority of schools in the middle of the quality distribution there is little consequence as measured by the enrollment outcome. This finding may reflect muted parental response, but it may also be a consequence of the limited choices available to parents in the English public schooling system.

The second part of the paper suggests that the latter explanation best fits the facts. This analysis investigates underlying demand by focusing on parents' ranked preferences over local schools. The results suggest that there is a strong response to all ratings, not just those at the extreme. In particular, simplifying the way in which information is presented in the reports appears to generate a large response. A robust finding is that the new, simplified style of reports helps families differentiate between the good and less good in the middle part of the school quality distribution.

References

Figure 1: Example time line showing treatment and control groups for evaluating the effect of an 'Outstanding' inspection rating on school enrollment



Note: This time line depicts schools rated Outstanding in 2005/06 and 2007/08. The post-treatment outcome is enrollment in 2006/07. See text for further details.

Table 1: Descriptive Statistics for Schools by Inspection Year and Inspection Rating

	Inspected 2006 ('treatment' group)	Inspected 2008 ('control' group)	p-value for t-test of difference in means
Panel A **Grade in 2006 or 2008 inspection: Outstanding (= Grade 1)**			
Previous inspection year	2000.4 0.1	2003.6 0.1	0.00
Previous inspection rating (range: 1-4)	1.75 0.07	1.62 0.05	0.12
% of students attaining Mathematics and English competency, age 11, 2005	88.0 0.8	83.4 0.9	0.00
% students entitled to free school meal, 2005	19.2 1.7	18.8 1.4	0.86
% students white British, 2005	73.9 2.9	78.5 2.1	0.19
Total enrolment	295.1 12.6	309.6 9.7	0.35
Number of schools	130	172	

	Inspected 2006 ('treatment' group)	Inspected 2008 ('control' group)	p-value for t-test of difference in means
Panel B **Grade in 2006 or 2008 inspection: Fail (= Grade 4)**			
Previous inspection year	2000.3 0.1	2003.5 0.1	0.00
Previous inspection rating (range: 1-4)	2.35 0.05	2.24 0.06	0.20
% of students attaining Mathematics and English competency, age 11, 2005	61.0 1.3	64.7 1.4	0.05
% students entitled to free school meal, 2005	29.1 1.8	29.1 1.7	0.99
% students white British, 2005	78.2 2.5	76.4 2.7	0.62
Total enrolment	293.4 10.5	308.0 10.5	0.33
Number of schools	122	109	

Notes: Standard errors in brackets. English and Mathematics competency at age 11 defined as percentage of students achieving level 4 on Key Stage 2 test. Samples consist of schools receiving an Oustanding (Panel A) or Fail (Pabel B) rating in 2006 or 2008 from a full sample of all regular (i.e. excluding schools serving special needs students exclusively), community primary schools (i.e. excluding religious schools) serving students aged 5 to 11.

Table 2: The Effect of Inspection Ratings on Enrollment

(Outcome: log enrolment; schools inspected in 2006 or 2008)

	(1)	(2)	(3)	(4)	(5)
	Basic DID	DID with school fixed effects			
		Basic FE	Full set of controls	Local growth in student pop. below national median	Local growth in student pop. below bottom quartile
Inspection grade: Outstanding					
2007 x early inspected	0.0255**	0.0255**	0.0252**	0.0398**	0.0593*
	(0.0075)	(0.0075)	(0.0074)	(0.0102)	(0.0228)
2007	-0.0054	-0.0054	-0.0055	-0.0234**	-0.0191
	(0.0053)	(0.0053)	(0.0051)	(0.0072)	(0.0122)
School FE	No	Yes	Yes	Yes	Yes
Full set of controls	No	No	Yes	Yes	Yes
Observations	604	604	604	328	126
Number of schools	302	302	302	164	63
R-squared	0.005	0.043	0.048	0.089	0.123
Inspection grade: Fail					
2007 x early inspected	-0.0433**	-0.0433**	-0.0439**	-0.0439**	-0.0411
	(0.0127)	(0.0127)	(0.0127)	(0.0161)	(0.0294)
2007	-0.0552**	-0.0552**	-0.0557**	-0.0649**	-0.0729**
	(0.0094)	(0.0094)	(0.0097)	(0.0112)	(0.0262)
School FE	No	Yes	Yes	Yes	Yes
Full set of controls	No	No	Yes	Yes	Yes
Observations	464	464	464	268	100
Number of schools	232	232	232	134	50
R-squared	0.016	0.416	0.421	0.515	0.599

Notes: Standard errors clustered at the school level; * and ** indicate significance at the 5% and 1% levels, respectively. Coefficients and standard errors multiplied by 100. Models estimated using enrollment data from 2005 and 2007. Panel A reports results for schools rated Oustanding in the 2006 or 2008 inspection; Panel B reports results for schools rated Fail in the 2006 or 2008 inspection. Column (4) shows estimates for schools located in Local Authorities (LA) where enrollment declines by more than that for the median LA between 2005 and 2007 (growth at median LA is -3.5 percent); column (5) shows estimates for schools in LAs where enrolment declines by more than that for the 25th percentile LA (growth at the 25th percentile LA is -5.5 percent). Schools with missing enrollment data from either of 2005 or 2007 are dropped. Time-varying controls are within-local authority percentiles on the average over the previous two years for: the school's English and Mathematics performance; the proportion of students receiving a free school meal; and the proportion of white British students. Missing dummies are included for the proportion of students receiving a free school meal and the proportion of white British students.

Table 3: Effect of Inspection Ratings on Enrolment in Pre-Treatment Years (Faslification Test)

(Outcome: log enrolment; schools inspected in 2006 or 2008)

	(1)	(2)	(3)	(4)	(5)
	Basic DID	DID with school fixed effects			
		Basic FE	Full set of controls	Local growth in student pop. below national median	Local growth in student pop. below bottom quartile
Inspection grade: Outstanding					
2005 x early inspected	0.0087	0.0087	0.0079	0.0142	0.0143
	(0.0075)	(0.0075)	(0.0075)	(0.0107)	(0.0135)
2005	-0.0047	-0.0047	-0.0043	-0.0196**	-0.0246**
	(0.0052)	(0.0052)	(0.0052)	(0.0066)	(0.0086)
School FE	No	Yes	Yes	Yes	Yes
Full set of controls	No	No	Yes	Yes	Yes
Observations	578	578	578	316	152
Number of schools		289	289	158	76
R-squared	0.003	0.005	0.009	0.053	0.115
Inspection grade: Fail					
2005 x early inspected	-0.0092	-0.0092	-0.0083	-0.0107	0.0189
	(0.0138)	(0.0138)	(0.0142)	(0.0218)	(0.0397)
2005	-0.0415**	-0.0415**	-0.0428**	-0.0546**	-0.1226**
	(0.0097)	(0.0097)	(0.0096)	(0.0137)	(0.0290)
School FE	No	Yes	Yes	Yes	Yes
Full set of controls	No	No	Yes	Yes	Yes
Observations	428	428	428	204	82
Number of schools		214	214	102	41
R-squared	0.021	0.177	0.181	0.264	0.415

Notes: Standard errors clustered at the school level; * and ** indicate significance at the 5% and 1% levels, respectively. Models estimated using enrollment data from 2003 and 2005. Panel A reports results for schools rated Oustanding in the 2006 or 2008 inspection; Panel B reports results for schools rated Fail in the 2006 or 2008 inspection. '2005 x early inspected' dummy switched on in 2005 for schools inspected in 2006; '2005' switched off in 2003 and on in 2005. Column (4) shows estimates for schools located in Local Authorities (LA) where enrollment declines by more than that for the median LA between 2003 and 2005 (growth at median LA is -2.7 percent); column (5) shows estimates for schools in LAs where enrolment declines by more than that for the 25th percentile LA (growth at the 25th percentile LA is -4.4 percent). Schools with missing enrolment data from either of 2003 or 2005 are dropped. See also notes in previous table.

Table 4: Summary statistics for first choice school and schools in the choice set

	1st choice school	Nearest 3 available schools	All available schools in borough
Distance from home (km)	1.09	0.92	5.25
	(1.38)	(0.77)	(3.11)
Distance rank	2.91	2.00	15.75
	(3.85)	(0.82)	(8.84)
Latest inspection rating (range:1-4)	2.11	2.32	2.50
	(0.76)	(0.79)	(0.78)
English and Mathematics decile	5.59	4.78	4.04
	(2.70)	(2.90)	(2.59)
% Eligible free lunch decile	5.54	6.44	7.39
	(2.44)	(2.38)	(1.97)
% White British	45.0	42.5	42.7
	(26.4)	(25.0)	(23.9)
Log enrollment	5.82	5.75	5.72
	(0.40)	(0.43)	(0.44)
Observations	6,467	19,401	196,907

Notes: Standard deviations in parentheses . Data from applications made in the fall of 2006, 2007 and 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). Latest inspection ratings range from 1 (Outstanding) to 4 (Fail); from academic year prior to application or earlier. 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. Enrollment equals number of full-time equivalent students. Applicants who missed the application deadline are excluded. See data appendix for further details.

Table 5: Distribution of inspection ratings

(Each cell records number of schools)

Panel A

School % free lunch rank (quintile)	Inspection rating: Out-standing	Good	Satis-factory	Fail
1	2	9	0	0
2	3	4	2	1
3	1	4	6	0
4	2	3	5	0
5	1	2	7	0

Panel B

School test rank (quintile)	Inspection rating: Out-standing	Good	Satis-factory	Fail
5	1	7	0	0
4	5	6	0	0
3	1	4	5	0
2	1	3	5	1
1	1	2	7	0

Notes: Inspection ratings are latest available for each school, 2006. School's percent free lunch rank calculated as average for 2005 and 2006; quintile rank is calculated from the borough distribution. School test rank calculated using 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 2005 and 2006 . Total of 52 secular schools; 3 have missing test score data in Panel B.

Table 6: Rollout of new style inspection reports

	Latest inspection ratings for schools at the end of academic year:		
	2005/06	2006/07	2007/08
Outstanding	9	10	6
o/w Outstanding, new style	2	6	5
Good	22	24	24
o/w Good, new style	5	14	22
Satisfactory	20	16	19
o/w Satisfactory, new style	6	12	18
Fail	1	2	2
o/w Fail, new style	0	2	2
Total number of schools	52	52	51

Notes: Table shows the distribution of inspection ratings for all secular schools in the borough at the end of academic year 2005/06 (column 1), 2006/07 (column 2) and 2007/08 (column 3). Over this period old style reports are gradually replaced by new style ones. Total number of secular schools is 54 in 2005/06 (two new schools are inspected in 2006/07); this total falls to 52 and 51 in the following two years as three schools are forced to merge with other schools.

Table 7: The effect of inspection ratings on school choice: conditional logit estimates

(Outcome: first choice school)

	(1)	(2)	(3)
Outstanding		0.062	-0.054
		(0.041)	(0.062)
Satisfactory		-0.614***	-0.328***
		(0.042)	(0.061)
Fail		-1.071***	-0.958***
		(0.103)	(0.165)
New style report			0.019
			(0.053)
Outstanding x new style report			0.255***
			(0.077)
Satisfactory x new style report			-0.410***
			(0.064)
Fail x new style report			-0.168
			(0.203)
Distance	-1.739***	-1.714***	-1.707***
	(0.025)	(0.025)	(0.025)
Distance squared	0.083***	0.082***	0.081***
	(0.003)	(0.003)	(0.003)
English and Maths decile	0.135***	0.053***	0.054***
	(0.007)	(0.009)	(0.009)
% Eligible free lunch decile	-0.245***	-0.264***	-0.274***
	(0.009)	(0.009)	(0.009)
% White British	-0.011***	-0.011***	-0.011***
	(0.001)	(0.001)	(0.002)
% White British x applicant white British	0.019***	0.020***	0.020***
	(0.002)	(0.002)	(0.002)
Observations	196,907	196,907	196,907
Log-likelihood	-10,133.85	-9,983.40	-9,954.00

Notes: Standard errors in parentheses; ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. Inspection ratings range from Outstanding (rating = 1), Good (=2), Satisfactory (=3) and Fail (=4); Good is the omitted category. 'New style report' dummy turned on if inspection rating available at the time of application is from 2005/06 or later. See Table 4 and main text for definitions of other variables. Missing dummies included for school's English and Mathematics decile.

Table 8: Effect of ratings by years since inspection

(Outcome: first choice school)

	(1)	(2)
Outstanding	-0.054	0.011
	(0.062)	(0.067)
Satisfactory	-0.328***	-0.263***
	(0.061)	(0.066)
Fail	-0.958***	-0.895***
	(0.165)	(0.167)
New style report	0.019	-0.006
	(0.053)	(0.054)
Outstanding x new style report	0.255***	0.190
	(0.077)	(0.138)
Satisfactory x new style report	-0.410***	-0.392***
	(0.064)	(0.092)
Fail x new style report	-0.168	-0.126
	(0.203)	(0.203)
Outstanding x new style report		0.040
x years since new style inspection		(0.064)
Good x new style report		0.077***
x years since new style inspection		(0.029)
Satisfactory x new style report		0.011
x years since new style inspection		(0.046)
Distance	-1.707***	-1.706***
	(0.025)	(0.025)
Distance squared	0.081***	0.081***
	(0.003)	(0.003)
English and Maths decile	0.054***	0.056***
	(0.009)	(0.009)
% Eligible free lunch decile	-0.274***	-0.276***
	(0.009)	(0.009)
% White British	-0.011***	-0.011***
	(0.002)	(0.002)
% White British x applicant white	0.020***	0.020***
British	(0.002)	(0.002)
Observations	196,907	196,907

Notes: Standard errors in parentheses; ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. See notes to previous table and main text.

Table 9: Effect for schools receiving the same rating in the old style and new style report

(Outcome: first choice school)

	(1)	(2)
No change in rating	0.295***	0.123
	(0.043)	(0.083)
No change x Satisfactory	-0.603***	-0.433***
	(0.060)	(0.078)
No change x New style report		0.209**
		(0.085)
No change x New style x Satisfactory		-0.325***
		(0.091)
Distance	-1.726***	-1.726***
	(0.025)	(0.025)
Distance squared	0.082***	0.082***
	(0.003)	(0.003)
English and Maths decile	0.113***	0.110***
	(0.007)	(0.007)
% Eligible free lunch decile	-0.240***	-0.245***
	(0.009)	(0.009)
% White British	-0.011***	-0.011***
	(0.001)	(0.001)
% White British x applicant white British	0.019***	0.019***
	(0.002)	(0.002)
Observations	196,907	196,907

Notes: Standard errors in parentheses; ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. See notes to Table 7 and main text.

Table 10: Heterogeneous effects - poverty status

(Outcome: first choice school)

	(1)	(2)	(3)	(4)
	Student eligible for free lunch?			
	YES	NO	YES	NO
Outstanding	0.014	0.085*	0.037	-0.076
	(0.097)	(0.046)	(0.128)	(0.071)
Satisfactory	-0.329***	-0.709***	-0.223**	-0.420***
	(0.089)	(0.048)	(0.110)	(0.074)
Fail	-0.741***	-1.144***	-0.722**	-1.009***
	(0.193)	(0.123)	(0.310)	(0.196)
New style			-0.045	0.062
			(0.102)	(0.062)
Outstanding x new style report			-0.026	0.312***
			(0.177)	(0.088)
Satisfactory x new style report			-0.185	-0.400***
			(0.120)	(0.078)
Fail x new style report			-0.037	-0.214
			(0.368)	(0.244)
Distance	-1.694***	-1.724***	-1.690***	-1.719***
	(0.056)	(0.028)	(0.056)	(0.028)
Distance squared	0.096***	0.079***	0.096***	0.079***
	(0.007)	(0.003)	(0.007)	(0.003)
English and Maths decile	0.068***	0.045***	0.064***	0.048***
	(0.019)	(0.010)	(0.020)	(0.010)
% Eligible free lunch decile	-0.057***	-0.316***	-0.066***	-0.324***
	(0.021)	(0.010)	(0.022)	(0.011)
% White British	-0.015***	-0.010***	-0.016***	-0.011***
	(0.003)	(0.002)	(0.003)	(0.002)
% White British	0.027***	0.018***	0.028***	0.018***
x applicant white British	(0.005)	(0.003)	(0.005)	(0.003)
Observations	36,409	160,498	36,409	160,498

Notes: Standard errors in parentheses; ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. See notes to Table 7 and main text.

Table 11: Estimates for families with an older child enrolled in primary school

(Outcome: first choice school)

	Older sibling in primary school?	
	No	Yes
Outstanding	-0.039	-0.083
	(0.082)	(0.094)
Satisfactory	-0.387***	-0.276***
	(0.085)	(0.088)
Fail	-1.027***	-0.870***
	(0.224)	(0.246)
New style report	-0.017	0.072
	(0.071)	(0.079)
Outstanding x new style report	0.380***	0.082
	(0.103)	(0.119)
Satisfactory x new style report	-0.474***	-0.314***
	(0.090)	(0.093)
Fail x new style report	-0.398	0.031
	(0.288)	(0.290)
Distance	-1.741***	-1.687***
	(0.034)	(0.036)
Distance squared	0.077***	0.088***
	(0.004)	(0.004)
English and Maths decile	0.076***	0.024*
	(0.012)	(0.013)
% Eligible free lunch decile	-0.308***	-0.233***
	(0.013)	(0.014)
% White British	-0.010***	-0.013***
	(0.002)	(0.002)
% White British	0.020***	0.022***
x applicant white British	(0.003)	(0.003)
Observations	113,951	82,956

Notes: Standard errors in parentheses; ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 12: Mixed logit estimates

(Outcome: first choice school)

Mean	
Outstanding	-0.089
	(0.060)
Satisfactory	-0.756***
	(0.058)
Fail	-1.252***
	(0.116)
Distance	-2.320***
	(0.052)
English and Maths decile	0.067***
	(0.010)
Standard deviation	
Outstanding	0.829***
	(0.173)
Satisfactory	0.956***
	(0.143)
Fail	0.070
	(0.310)
Distance	1.171***
	(0.037)
English and Maths decile	0.033
	(0.025)
Observations	196,907
Log-likelihood	-9,619.61