

# Journal of School Choice

## International Research and Reform

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/wjsc20>

## How Useful are Ofsted Inspection judgements for Informing Secondary School Choice?

Christian Bokhove, John Jerrim & Sam Sims

**To cite this article:** Christian Bokhove, John Jerrim & Sam Sims (2023) How Useful are Ofsted Inspection judgements for Informing Secondary School Choice?, *Journal of School Choice*, 17:1, 35-61, DOI: [10.1080/15582159.2023.2169813](https://doi.org/10.1080/15582159.2023.2169813)

**To link to this article:** <https://doi.org/10.1080/15582159.2023.2169813>



© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 27 Feb 2023.



Submit your article to this journal [↗](#)



Article views: 668



View related articles [↗](#)



View Crossmark data [↗](#)

## How Useful are Ofsted Inspection judgements for Informing Secondary School Choice?

Christian Bokhove<sup>a</sup>, John Jerrim <sup>b</sup> and Sam Sims<sup>c</sup>

<sup>a</sup>Southampton Education School, University of Southampton, Southampton, UK; <sup>b</sup>University College, UCL Social Research Institute, London, UK; <sup>c</sup>University College, UCL Centre for Education Policy and Equalising Opportunities, London, UK

### ABSTRACT

School inspections are a key component of the accountability system in many education systems, including England. The judgments and reports produced through these inspections are widely used by parents when they are choosing a school for their children. But should they be? This paper presents new evidence on this issue. We illustrate how parents selecting secondary schools using Ofsted judgments will often be basing their decision on dated information. Indeed, half the time, this will be based on a period in which the school had a different headteacher. We find there are almost no differences in future academic, behavioral, school leadership and parental satisfaction outcomes between schools rated as good, requiring improvement and inadequate in the inspection data available to parents at the point of school selection. That is, parents who choose a “good” secondary school for their child will not leave with appreciably better outcomes than a parent who selects an “inadequate” school. The one exception to this is an Outstanding judgment, which does predict future academic outcomes – though only if the inspection was conducted within the last five years. We thus advise parents that – besides choices involving Outstanding schools – Ofsted judgments are of limited use to them in selecting a school.

### KEYWORDS

Ofsted; school-inspection; school-choice; accountability

## Introduction

In recent years, many western countries have made moves toward a market-based education system. Some have introduced controversial voucher systems, while others have created quasi-markets, where parents are able to choose between “competing” schools (OECD, 2017). Indeed, in England – the empirical setting for this paper – recent data from the Programme for International Student Assessment (PISA) demonstrates how almost 90% of secondary headteachers report they are in competition with at least two other schools for prospective pupils (Sibieta & Jerrim, 2021). Such competition is meant to empower young people and their parents as the “consumers” of the education

**CONTACT** John Jerrim  [J.Jerrim@ucl.ac.uk](mailto:J.Jerrim@ucl.ac.uk)  UCL Social Research Institute, 55-59 Gordon Square, Bloomsbury WC1H 0AL, London

Author list is alphabetical. All joint first authors. December 2022

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

that schools provide. If the education provided is sub-standard, then parents – or at least prospective parents – can choose an alternative school (Bradley, Crouchley, Millington, & Taylor, 2000; Burgess, Greaves, & Vignoles, 2019; Sibieta, 2006). In theory, market forces should thus raise education standards as a result (Jabbar et al., 2019).

However, for any market to work efficiently, consumers (parents and their offspring when it comes to school choice) must be well-informed (Allen & Burgess, 2010). More precisely, they need to have access to accurate, timely and easy to understand information about the quality of nearby schools. In practice, parents take a range of factors into account when choosing where they send their children (YouGov, 2017). Informally, parents are likely to have discussions with other parents, some of whom will already have children attending local schools (Ball, Bowe, & Gewirtz, 1995; Ball & Vincent, 1998). More formally, a range of codified performance indicators are also available including historic measures of school performance, levels of staff and student absenteeism, and characteristics about school intake. Together, it is hoped such information will help parents make rational, well-informed decisions with respect to their offspring's schooling.

One of the most widely used resources for school selections in England are Office for Standards in Education, Children's Services and Skills (Ofsted) inspection outcomes and reports (YouGov, 2017). Ofsted inspects England's educational institutions and publishes reports on its findings to monitor and improve the quality of education that young people receive. These reports provide a unique insight into England's schools, based on visits to schools by inspectors. Headline results from Ofsted inspections are widely reported in the media, are used by schools in their marketing to prospective parents/pupils and even take a prominent position amongst the literature provided by estate agents to those looking to purchase a house. Focus groups from Ofsted suggest that around one-in-five parents "*read all the relevant Ofsted reports before choosing which schools to apply to*" and highlight cases where the information from Ofsted "*was decisive*" (Ofsted, 2019, p. 6). Indeed, the Ofsted website<sup>1</sup> even notes how: "*Inspection provides important information to parents, carers, learners and employers about the quality of education, training and care. These groups should be able to make informed choices based on the information published in inspection reports.*" The information provided by school inspections hence plays a key role in school choice in England.

In this paper, we discuss the pros and cons of using Ofsted inspection outcomes as a basis for school choice. The research is closely related to Leckie and Goldstein (2009), who investigate the limitations of using school league tables to help parents choose a school. Like Leckie and Goldstein, our empirical analysis focuses on the time delays involved in this process. Consider, for example, parents who chose a secondary school for their eldest child in October 2013. This child began secondary school in September 2014

and took their General Certificate of Secondary Education (GCSEs) in May/June 2019. However, the most recent Ofsted inspection available to the parents in October 2013 would most likely be two or more years old (National Audit Office [NAO], 2018). Thus, we are interested in answering the question: is an Ofsted report informative enough about the quality of that same school during the period – three to eight years later – when the child would attend the school?

There has been previous work exploring the usefulness of Ofsted inspection judgments. Ofsted (2017) have conducted their own research on the reliability of judgments, finding that, on most occasions, independent inspection teams reach the same decision when conducting short inspections. Other research has investigated the predictive validity of inspection judgments. Hussain (2012) investigate whether Ofsted judgments predict pupil survey responses capturing various aspects of teaching quality, finding that “*even after controlling for things like test scores and the socioeconomic background of pupils, inspection ratings appear to have substantial power in predicting underlying quality*” (p. 12). He thus advises that “*parents who are looking for a good school ought to place at least some weight on inspection ratings*” (p. 12). In contrast, Von Stumm et al. (2021) find Ofsted ratings of secondary schools account for just 1% of variation in pupil achievement at age 16 and correlate only 0.03 with wellbeing/engagement measures. They thus state their “*findings call into question the usefulness of Ofsted ratings as a guide for parents who are looking to make an informed choice for their children’s secondary school.*”

A related strand of the literature has examined whether Ofsted inspections can act as an early warning system for school failure by gathering “on the ground” information that would not otherwise be available to parents. Research by FFT Education Datalab (2015) found Ofsted judgments were not a leading indicator of future examination performance and that there is “*no evidence that Ofsted judgements reflect schools on the cusp of change*” (p. 15). Likewise, research by Hutchinson (2016) considered whether inspection ratings are good at identifying schools that go on to suffer a deterioration in national examination scores. She notes how “*notable proportions of ‘good’ and ‘outstanding’ schools are not down-graded, despite a substantial [subsequent] deterioration in their academic performance*” (p. 7). The one exception to this pattern of findings comes from Hussain (2015), who finds primary schools that receive a poor inspection rating experience a gain in test scores in the future. Of course, Hussain’s (2015) findings may not hold for secondary schools, which tend to be larger, teach more advanced material to pupils and prepare young people for important high-stakes examinations. Nevertheless, the aforementioned studies highlight the dual (and potentially conflicting) role of Ofsted inspections in 1) measuring the quality of schools to inform school choice and 2) triggering improvement in underperforming schools.

## The limitations of inspection outcomes for school-choice

Inspection outcomes potentially provide unique information to parents when they are choosing a school, acting as an independent measure of school quality. When inspectors visit schools, they get an impression of what life is like “on the ground,” which they then communicate to parents through their reports. Such valuable qualitative information may not be available to parents through other sources, such as exam results. Inspection outcomes also partially reflect those who engage with the school on a day-to-day basis (parents, pupils, teachers), all of whom complete a survey when an inspection takes place. Yet the limitations of inspection outcomes for informing school choice are less well-known, or at least less widely discussed. In this section, we document some of the most important issues, many of which are further explored in our empirical analysis.

### *Time-lag*

Ofsted reports are backward looking, capturing a picture of a school at one point in time. This is potentially valuable information for parents with children *currently* attending the school. Yet for parents making forward-looking schooling decisions, this time-lag is a notable limitation. Even if a school is inspected just before parents make their choice, it will still be around a year before their child joins and six years before they leave (in reference to secondary schools). In reality, schools are not inspected every year. Between 2012 and 2020 schools with an “outstanding” judgment were exempt from routine inspections. Schools previously judged as “good” are subject to a short inspection every four years. Meanwhile, “inadequate” schools or those “requiring improvement” receive another full inspection within 2.5 years. Hence the most recent Ofsted reports available to parents at the point of school selection may already be several years old, particularly if the school received a “good” or “outstanding” grade.

### *Siblings*

Building upon the argument above, parents with more than one child usually send them to the same school. This is in-part due to the logistical challenges and financial/time constraints associated with sending siblings to different schools. As a result, many schools preferentially offer places to siblings of existing students (Burgess, Greaves, & Vignoles, 2020). Yet this also has implications for the use of data in school choice. Assuming that parents have strong preferences for sending all their children to the same school, the choice made for the eldest child will also be the default choice for their

younger sibling(s). This means that – to be informative for schooling decisions of multi-child families – inspection outcomes need to be predictive of the academic environment even further into the future i.e. until the youngest child has left the school.

### ***Inspection as measurement of quality versus inspection as a driver of quality***

The main aim of Ofsted inspections are to provide an independent, holistic assessment of school quality – and to identify (and support) areas that need improvement. Providing prospective parents with useful information is a potential by-product of this process. Indeed, in an ideal world, one would want inadequate and schools requiring improvement to see a marked improvement as quickly as possible. Yet, were that to happen, then the information provided by the inspection would soon become obsolete – and of little use to parents when selecting schools. In other words, there is a trade-off between inspections stimulating school improvement and providing useful information to parents when they are selecting schools – they cannot easily do both. Moreover, if Ofsted inspection grades are only weak predictors of future school outcomes, this may be due to the school-improvement process working well (i.e. inspections improving underperforming schools).

### ***Comparability of data***

To choose between competing options, individuals require comparable information to weigh up the pros and cons. This poses a challenge to the use of Ofsted inspections in the selection of schools. First, inspection regimes change over time, with greater focus on some areas than others at different points. For example, the Ofsted inspection framework introduced in 2019 put less emphasis upon test scores and more on curriculum quality. Yet, given the difference across schools in the timing of the latest available inspection report at the point of school choice, the establishments parents are choosing between may have been judged according to different criteria. Second, schools are evaluated by different inspectors and inspection teams. Although they are trained professionals who all attempt to apply consistent standards (see Ofsted, 2017) what constitutes good practice may – in places – be open to interpretation. It is thus unlikely that there is complete consistency in inspection judgments (and their written reports) across inspectors. This limits the comparability of information available to parents when they are selecting a school. See Bokhove, Jerrim, and Sims (2023) for evidence on this issue with respect to Ofsted inspections in England.

### **The available choice set**

In practice, parents will have some constraints on their choice of school due to their location. They may have only two or three schools within their local catchment area or within a commutable distance (Burgess, Briggs, McConnell, & Slater, 2006). Such constraints will mean there is limited variation in Ofsted grades amongst the schools that parents may realistically apply to. Indeed, all schools realistically available to families may have the same Ofsted grade. Ofsted judgments will hence be of little use to parents if they do not vary sufficiently amongst the schools that are realistically available to them. Although parents may take Ofsted grades into account when making house buying decisions, this is often done well in advance of when they will start applying to schools. Thus, when house purchases are made based upon *current* Ofsted grades, parents are relying on these predicting the quality of the school (and the future outcomes of its pupils) even further into the future.

### **Added value over other information?**

Ofsted judgments and reports summarize many available sources of information to provide a holistic judgment. For example, inspectors take into account publicly available data on e.g. attendance rates. However, there is evidence to suggest that parents also take account of these other sources of information when choosing schools (Burgess, Greaves, Vignoles, & Wilson, 2015). In asking *How useful are Ofsted inspection judgments for informing school choice?* we might also question whether Ofsted inspection judgments “add-value” (in terms of aiding school choice) above and beyond other sources of data available to parents.

### **Methods and data**

Our empirical analysis is based around the following hypothetical scenario.<sup>2</sup> Parents with two children – born exactly two years apart – are choosing a secondary school for their eldest offspring in October 2013.<sup>3</sup> Their eldest child will begin secondary school in September 2014 and take their GCSEs in May/June 2019. For every secondary school in the country,<sup>4</sup> we then ask two questions: (a) how old is the most recent inspection report available to our family at the time they make a choice about secondary school (b) how well does the most recent Ofsted judgment predict the school environment and pupils outcomes during the period in which their children would actually attend this school? To implement this approach, we draw on the data sources outlined below.

### **Ofsted data**

Information from the Department for Education's "School Performance Tables" spine<sup>5</sup> is used to generate a list of all secondary schools open in the 2013/14 academic year (when our hypothetical family was selecting a secondary school). Independent and special schools are removed, leaving a total of 2,538 schools. This information is merged with information published by Ofsted on all inspections between 2005 and 2015.<sup>6</sup> The most recent school inspection before October 2013 is taken as the information source our hypothetical family used when choosing a secondary school.<sup>7</sup> We focus on overall effectiveness judgments which classify schools into four groups:

- Outstanding (n = 628)
- Good (n = 1,207)
- Requires improvement<sup>8</sup> (n = 576)
- Inadequate (n = 127)

In additional analysis we have focused on the following Ofsted sub-judgments instead and found our substantive findings to remain unchanged.

### **Achievement outcomes**

We utilize two school-level achievement metrics. The first is school-average "Attainment 8" scores, capturing *absolute* performance of pupils in GCSE examinations. The second is the school-average "Progress 8" score. This captures the average academic *progress* pupils make during secondary education (Prior, Jerrim, Thomson, & Leckie, 2021). We focus on Attainment 8 and Progress 8 outcomes for the 2018/19 academic year,<sup>9</sup> which is when their eldest child will sit their GCSE examinations.

### **Parent view data**

When a school is inspected, parents are invited to complete a questionnaire (known as "Parent View"). This includes 11 questions using a four-point scale (strongly agree to strongly disagree) such as "*My child is taught well at this school,*" "*This school responds well to any concerns I raise*" and "*My child is happy at this school.*" School-level summary information has been published by Ofsted since January 2018. Response rates are relatively low (the median response rate across secondary schools is around 20%) meaning these data are best considered a convenience sample. We nevertheless use this information (averaged over 2018 and 2019) to consider the relationship between Ofsted judgments and a broader set of outcomes. Further details are provided below.



### **Parental satisfaction with school**

The Parent View data is used to capture parental satisfaction with schools in two ways. First, a principal components analysis is conducted upon the 11 questions, with the first component (which accounts for around 85% of the total variation) then used as an overall measure of parental satisfaction with the school (Appendix B presents loadings for each item). This results in a “parental satisfaction with school” scale that we standardize to mean zero and standard deviation one, meaning results are presented as effect sizes (Appendix B presents the distribution of this scale).<sup>10</sup> Second, at the end of the survey parents were asked whether they would “*recommend this school to another parent*” (yes/no). We use this as an alternative measure of parental school satisfaction.

### **Pupil absenteeism and behavior**

Prior research has found that general levels of behavior within a school are worse when pupil absenteeism is high (Malcolm, Wilson, Davidson, & Kirk, 2003). Hence, as a proxy for behavior, we consider the percent of absent sessions across the 2015–2019 academic years (i.e. during the eldest child’s time at school). We again draw on the “Parent View” data for an alternative measure. Specifically, we take the percentage of parents who strongly agree that “*This school makes sure its pupils are well behaved*” as a measure of parental satisfaction with behavior at the school during the time that their child attends.

### **Headteachers and satisfaction with leadership**

To illustrate the timeliness of inspection data for school choice, we consider whether the headteacher during the relevant inspection is the same as the headteacher when the children from our hypothetical family attend. This is important as the headteacher is a key figure within England’s schools. According to data from PISA 2015, over 90% of headteachers in England report having considerable responsibility over hiring and firing teachers, allocating budgets and determining disciplinary policies, with over 85% saying they hold considerable responsibility for establishing teachers’ salaries and determining assessment policies.<sup>11</sup>

To begin, we construct a panel dataset of the headteacher of each school between 2010/11 and 2018/19. As there is no routinely published longitudinal database of schools’ headteachers, we combine information from various Freedom of Information requests.<sup>12</sup> We then designate the “head at inspection” as the most recently named headteacher before the inspection took place. If the most recent inspection occurred before August 2010 – the earliest point where

we have comprehensive information available – then the headteacher recorded in August 2010 is used. This means we will to some extent *overestimate* the extent that the headteacher leading the school when children from our hypothetical family attends is the same as the headteacher as when the relevant Ofsted inspection took place (i.e. our estimates of change in headteacher will be conservative). Using this information, we derive a binary variable for the 2013/14 academic year onwards. This is coded one if the surname of the headteacher in that academic year was the same as the surname of the headteacher at the point of the relevant inspection, and zero if different.<sup>13</sup>

Data from Parental View is also used to gauge parental satisfaction with the leadership of school when the eldest child from our hypothetical family was working toward their GCSEs (2018/19). This is measured by the percent of parents who strongly agree that “*This school is well led and managed.*”

### **Analytical approach**

We begin by presenting descriptive statistics summarizing the timeliness of Ofsted inspections for informing school decisions. Specifically, we document the number of days between the secondary school decision being made and the date of the most recent prior inspection:

$$Length_t = Decision_t - Most$$

Where:

$Decision_t$  = Approximate date when secondary school decision is made (assumed to be October 2013).

$Most^t$  = Date of most recent inspection prior to October 2013.

t = time t.

Next, we turn to school leadership. This is one important input into inspection ratings. Yet, with respect to school choice, information from inspections about leadership and management are (arguably) only relevant if the leader (headteacher) does not change. We hence investigate whether the headteacher at the most recent inspection prior to the school selection decision is the same headteacher leading the school when children from our hypothetical family join and leave the school.

A set of Ordinary Least Squares (OLS) models will then be estimated:

$$O_j = \alpha + \beta.R_j + \gamma.C_j + \delta.D_j + \theta.A_j + \tau.I_j + \varepsilon_j$$

Where:

$O_j$  = One of our outcomes of interest (e.g. standardized Attainment 8 scores).

$R_j$  = Most recent Ofsted judgment before October 2013 (reference = “good”).

$C_j$  = School background characteristics (e.g. admissions policy, gender composition).

$D_j$  = Demographic composition of recent intakes (e.g. percent FSM eligible, average Key Stage 2 scores).

$A_j$  = School’s historical GCSE performance prior to the schooling decision (e.g. capped GCSE points in 2013).

$I_j$  = Historical school absence rates.

$j$  = school  $j$ .

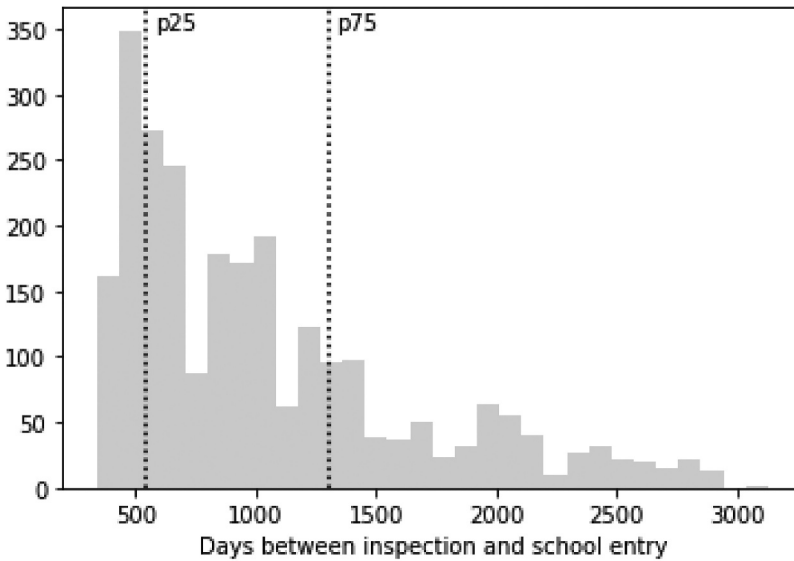
$\varepsilon_j$  = Random error term.

The  $\beta$  parameter captures the association between the Ofsted judgment used by our hypothetical family and the school environment/outcomes when their eldest child attends. This coefficient is estimated across five specifications M0 to M4 ([Appendix A](#) presents a full list of covariates included in each).

To begin, the Ofsted grade available to parents ( $R_j$ ) will be the only covariate (specification M0). The  $\beta$  parameter will thus capture the predictive power of Ofsted inspection grades alone. These estimates will be affected by selection – families with certain characteristics (e.g. those with higher incomes) are more likely to send their offspring to certain types of schools (e.g. those with outstanding Ofsted grades). One would therefore observe an association between the Ofsted judgments used in school choice and future outcomes simply due to who selects into such schools, rather than the judgments being predictive per se.

Model specifications M1 and M2 (our preferred specification) add background characteristics of the school ( $C_j$ ) and their demographic intakes ( $D_j$ ) to the model. As we are only able to control for observable characteristics of schools, these models may only reduce – rather than completely remove – the aforementioned problem of selection. Hence estimates from our preferred model specification (M2) may still provide an *upper bound* on the usefulness of Ofsted judgments to parents in making school choices.

Models M3 and M4 add historical GCSE performance ( $A_j$ ) and historical school absence rates ( $I_j$ ) to the model. These partially capture school “quality” and various aspects of the academic environment that parents are interested in (e.g. quality of teaching, disciplinary standards). Yet such information will also be available to parents when making schooling decisions. Hence the  $\beta$  parameters from these final two models will establish the added value of parents using Ofsted inspection judgments to inform their school decisions, over and above an array of other publicly accessible data.



**Figure 1.** The length of time between the latest inspection at the point of school choice and school entry. Notes: Figure illustrates the distribution of the length of time between the most recent inspection data available to parents when selecting a school and the date of secondary school entry for the focal child (October 1<sup>st</sup> 2014). P25 and p75 refer to the 25<sup>th</sup> and 75<sup>th</sup> percentile of the distribution.

## Results

### *How timely is the information available?*

Figure 1 illustrates the length of time between the inspection judgment available to parents when selecting schools and when the eldest child from our illustrative family begins secondary education. There are two key points. First, the mean (median) length of time is 1,040 (885) days. At the point the eldest child starts secondary school, the relevant inspection data is on average almost three years old. Second, the distribution is positively skewed, with significant inequality in the timeliness of information available. For instance, the data is at least 3.5 years old for around a quarter of schools (p75) and for 10% of schools the information is at least 5.5 years old. Of course, by the time the younger child from our illustrative family attends the school, this information will be even more dated. Hence, if parents use Ofsted judgments to pick a school, they are relying on inspection outcomes capturing school environment/quality a long period into the future.

Table 1 and Figure 2 shows how this varies by school characteristics. Information from inspections are particularly dated for certain school types. For instance, the gap between the most recent inspection and school entry of the eldest child from our hypothetical family is (on average) around four years for schools in the top GCSE quartile and those previously rated as 'outstanding', while for selective schools inspection results are on average five years old.

**Table 1.** The average length of time between the most recent Ofsted inspection used in secondary school choice and school entry. Variation by school characteristics.

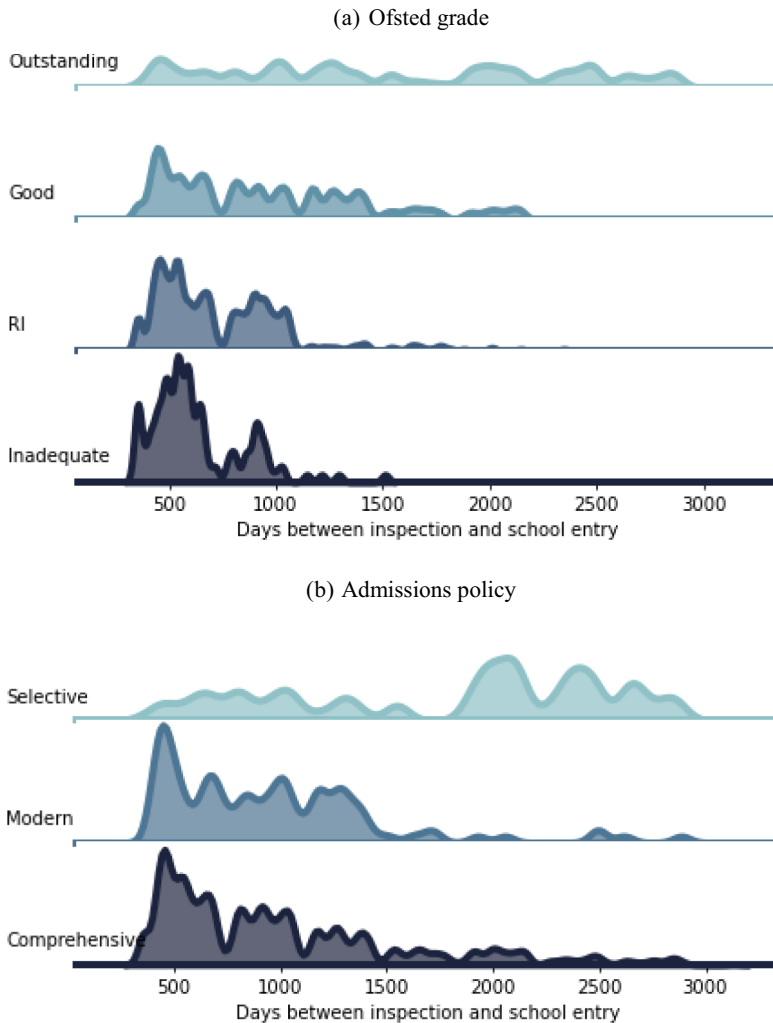
School characteristic	Observations	Mean (days)
<b>Most recent Ofsted rating</b>		
Outstanding	628	1,546
Good	1,207	965
RI	576	736
Inadequate	127	622
<b>FSM quartile</b>		
Low FSM	626	1,295
FSM Q2	613	1,014
FSM Q3	614	929
High FSM	613	901
<b>% EAL (2013)</b>		
Low EAL	637	995
EAL Q2	596	1,046
EAL Q3	601	1,040
High EAL	612	1,110
<b>% SEN (2013)</b>		
Low SEN	642	1,222
SEN Q2	633	1,072
SEN Q3	628	972
High SEN	635	890
<b>GCSE quartile (2013)</b>		
Low GCSE	650	793
GCSE Q2	622	904
GCSE Q3	664	1,020
High GCSE	577	1,483
<b>Admissions Policy</b>		
Comprehensive	2,282	994
Modern	103	956
Selective	153	1,780
<b>School type</b>		
Academy Converter	510	632
Academy Sponsor Led	255	753
Community School	788	1,104
Foundation School	551	1,268
Voluntary Aided School	375	1,308
Other	59	1,099

Notes: Figures refer to the average number of days between the inspection used in school choice and when the eldest born (focal) child entered the school. RI = Requiring Improvement

This highlights how the inspection information available to parents about different schools is unlikely to be comparable. It will be more dated for some than others, with the inspection potentially conducted in different ways and using different frameworks.

To provide a concrete example of what might change, [Table 2](#) documents the “survival” of headteachers over time. Specifically, between 2012/13 and 2020/21, [Table 2](#) presents the percent of headteachers leading the school who were also the headteacher when the relevant inspection took place.

For most of the time that our hypothetical children attend the school, the headteacher will be different to the one whose management/leadership was judged in the relevant inspection. Around 30% of headteachers will have already left when



**Figure 2.** The distribution of the length of time between the most recent Ofsted inspection used in secondary school choice and school entry. (a) Ofsted grade (b) Admissions policy Notes: See Table 1 for sample sizes. RI = Requiring Improvement

parents are making their schooling choices, around half when their children start at the school, and almost 90% by the time their youngest child leaves. Moreover, Figure 3 demonstrates how headteachers whose school received an “inadequate” judgment will almost certainly change by the time our hypothetical children attend.

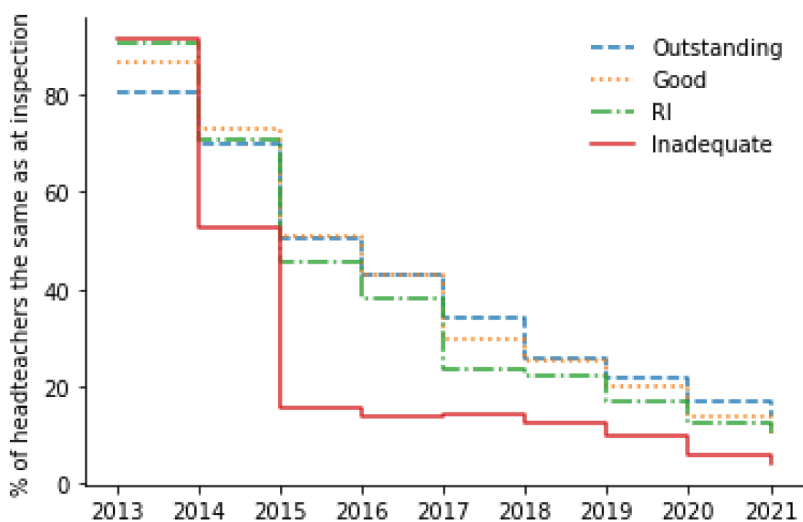
***Do the inspection judgments available to parents predict future outcomes?***

Even though inspection data may be dated, it may still be useful to parents selecting schools if it predicts school environment and outcomes some

**Table 2.** The percentage of schools with the same headteacher as at the most recent inspection.

Academic Year	% with the same headteacher as in the inspection	Number of observations
2012/13	86%	2,161
2013/14 (schooling decision made)	71%	2,161
2014/15 (eldest/focal child school entry)	49%	2,158
2015/16	41%	2,158
2016/17 (younger sibling starts school)	29%	432
2017/18	24%	2,137
2018/19 (eldest/focal child sits GCSEs and leaves school)	20%	2,087
2019/20	14%	2,058
2020/21 (younger sibling leaves school)	11%	2,025

Figures refer to the percentage of schools with the same headteacher in a given academic year as was in place at approximately the time of the previous inspection.



**Figure 3.** The “survival” of headteachers by Ofsted grade. Note: As the most recent inspection may have taken place before 2013 (potentially many years before) the curves do not start at 100%. RI = Requiring Improvement

distance into the future. Table 3 thus explores the link between Ofsted judgments used for school selection in 2013 and examination outcomes in 2019.

In the unconditional estimates (M0) a strong association is observed. For instance, average Attainment 8 scores are around one standard deviation higher in outstanding schools than in good schools, while in “inadequate” and “requiring improvement” schools they are around half a standard deviation lower. This could, of course, be driven by selection; if well-heeled families with already high-achieving children disproportionately choose good and outstanding schools, then the strong unconditional associations observed in model M0 is a self-fulfilling prophecy. We hence turn to estimates from model M2 (our preferred specification) which control for differences in school characteristics and pupil intakes.

**Table 3.** The association between Ofsted inspection judgment at time of school selection and academic outcomes in 2019.

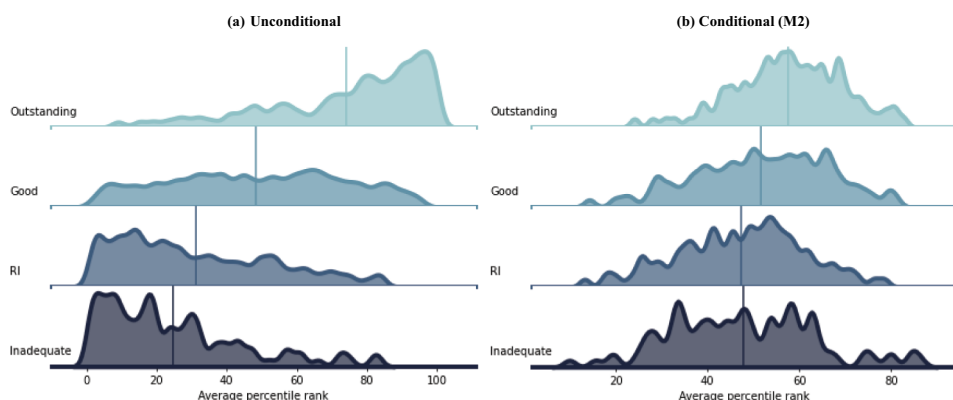
Model	Number of observations	Outstanding		RI		Inadequate	
		Effect size	SE	Effect size	SE	Effect size	SE
<b>(a) Attainment 8 (standardized)</b>							
M0	2,469	0.98*	0.04	-0.46*	0.04	-0.65*	0.08
M1	2,469	0.56*	0.03	-0.34*	0.03	-0.54*	0.06
M2	2,466	0.17*	0.02	-0.06*	0.02	-0.03	0.04
M3	2,444	0.11*	0.02	-0.01	0.02	0.05	0.04
M4	2,444	0.11*	0.02	-0.02	0.02	0.00	0.04
<b>(b) Progress 8 (standardized)</b>							
M0	2,464	0.76*	0.04	-0.42*	0.05	-0.48*	0.09
M1	2,464	0.55*	0.04	-0.34*	0.04	-0.40*	0.08
M2	2,461	0.31*	0.04	-0.13*	0.04	-0.05	0.07
M3	2,439	0.19*	0.04	-0.04	0.04	0.08	0.08
M4	2,439	0.19*	0.04	-0.05	0.04	-0.01	0.08

Notes: Figures refer to effect size differences compared to “good” schools as the reference category. RI = Requiring Improvement. Model M0 does not include any controls. M1 adds controls for school characteristics. M2 adds controls for the demographic intake of schools. Measures of historic performance in GCSEs is added in model M3, while M4 additionally controls for historic levels of school absence. See [Appendix A](#) for full list of control variables. \* indicates difference from the “good” category statistically significant at the five percent level.

There is now a stark difference in the results. The predictive power of Ofsted judgments available to our hypothetical parents and future school outcomes is weak. There is little difference in Attainment 8 scores between good, requiring improvement and inadequate schools, with effect sizes below 0.1 standard deviations. Although effect sizes below 0.1 are common in education research (Kraft, 2020), it is nevertheless clear that the predictive power of Ofsted grades observed in M0 is largely due to the selection of families into different schools. There is little to suggest that parents who choose a “good” secondary school for their child will leave with better academic outcomes than parents who selected an “inadequate” school. The one exception to this pattern of zero/small differences is with respect to the “outstanding” grade. While the association between an outstanding Ofsted judgment and A8/P8 scores is greatly reduced between M0 and M2, a moderately sized association remains (e.g. a 0.17 effect size difference in A8 scores between outstanding and good schools in model M2, down from 0.98 in model M0).

Figure 4 provides a graphical representation of these results. Models M0 and M2 have been estimated using quantile regression to illustrate the unconditional (M0) and conditional (M2) distribution of Attainment 8 scores by the Ofsted grade available to parents. The latter refers to the predicted distribution of attainment 8 scores for a mixed-sex, comprehensive academy converter school, with an average proportion of EAL, SEN and FSM pupils, with Key Stage 2 scores of the pupils around the national average. Estimates are presented in terms of percentile ranks, with a value of 1 indicating that a school is in the *bottom* 1% of the Attainment 8 distribution and a value of 99 indicating the top 1%. The solid vertical lines refer to the center points.





**Figure 4.** The predicted school-level distribution of attainment 8 scores by the Ofsted judgment available to aid school choice. (a) Unconditional (b) Conditional (M2) Note: Unconditional distributions based upon model M0 (no controls). Conditional estimates based upon model M2, controlling for school characteristics and demographic intake of pupils. Estimates have been produced using quantile regression. Conditional estimates refer to predicted distributions for a mixed-sex, comprehensive academy converter school with an average proportion of EAL, SEN and FSM pupils, and with Key Stage 2 scores of their intake at the national average. Vertical lines indicate the average percentile rank. RI = Requiring Improvement

There are clear differences in the Attainment 8 distribution by Ofsted grade in the unconditional results. While the distribution for inadequate and requiring improvement schools are reasonably similar, there is a notably wider spread (and higher average) for schools rated as good. Meanwhile, the Attainment 8 scores for outstanding schools has a strong negative skew, with little overlap of the distribution with those rated as inadequate. These differences are reflected by the average percentile ranks: the average inadequate school is in 25<sup>th</sup> position, compared to 31<sup>st</sup> for those requiring improvement, 48<sup>th</sup> for good schools and 74<sup>th</sup> for outstanding.

A rather different picture emerges in the conditional results. Consistent with the results for model M2 presented in Table 3, virtually no differences can be observed between the good, requiring improvement and inadequate Attainment 8 distributions. Moreover, although the distribution for outstanding schools is slightly to the right of the others, the difference is much reduced from the unconditional estimates in panel (a). Specifically, the average *conditional* percentile ranks by Ofsted grade are 48<sup>th</sup> for inadequate schools, 47<sup>th</sup> for those requiring improvement, 52<sup>nd</sup> for good schools and 57<sup>th</sup> for outstanding. In other words, most of the difference in A8 outcomes by the Ofsted grade available to parents when they are selecting a school can be explained by differences in their intake and background characteristics. Figure 4 panel (b) thus illustrates how – besides Outstanding schools – Ofsted judgments are of limited value to parents in their selection of schools.

These largely null findings are supported – and, indeed, somewhat strengthened – by the results from models M3 and M4 (Table 3). Recall that these add further controls for other data freely and easily accessible to parents when they are choosing schools (e.g. historic performance in GCSEs). Estimates from these models hence reveal the “added-value” to parents of using Ofsted judgments in informing their choices, over and above other easily accessible measures. There continues to be no clear difference in A8/P8 scores between good, requiring improvement and inadequate schools, with only a small benefit of picking an outstanding school. In other words, using Ofsted judgments in addition to other freely available information does little to help parents to pick a “better” school (at least in terms of academic outcomes).

Table 4 replicates this analysis focusing on parental satisfaction with the school. The smaller sample is due to these data only being available for schools that were inspected in 2018 or 2019. Null results are even more apparent. Even in the unconditional estimates (M0) we largely fail to find any association between Ofsted grades and whether parents would recommend the school to others (panel a) and overall satisfaction levels (panel b). The addition of controls does little to change this result. Indeed, in M3/M4, although statistically not significant, parents who send their child to an outstanding school are three percentage points *less* likely to recommend it to others than parents who

**Table 4.** The association between Ofsted inspection judgment at time of school selection and parental satisfaction with school.

		Outstanding		RI		Inadequate	
Number of observations		% point difference	SE	% point difference	SE	% point difference	SE
<b>(a) % of parents who would recommend school to others</b>							
M0	1,159	-0.6%	1.5%	-1.2%	1.0%	-1.9%	2.4%
M1	1,159	-1.1%	1.5%	-0.5%	1.0%	-0.7%	2.4%
M2	1,159	-2.8%	1.5%	0.6%	1.0%	2.0%	2.3%
M3	1,149	-3.1**%	1.5%	0.9%	1.0%	2.4%	2.4%
M4	1,149	-3.1**%	1.5%	0.9%	1.0%	2.4%	2.4%
		Outstanding		RI		Inadequate	
Number of observations		Effect size	SE	Effect size	SE	Effect size	SE
<b>(b) Parental satisfaction scale</b>							
M0	1,159	0.20*	0.10	0.02	0.07	0.04	0.16
M1	1,159	0.16	0.10	0.06	0.07	0.12	0.16
M2	1,159	0.11	0.10	0.09	0.07	0.20	0.16
M3	1,149	0.09	0.10	0.09	0.07	0.21	0.16
M4	1,149	0.09	0.10	0.09	0.07	0.21	0.16

Notes: Figures in panel (a) refer to the percentage point difference in parents who say they would recommend the school compared to the reference (“good”) category. Figures in panel (b) refer to effect size differences compared to “good” schools as the reference category. RI = Requiring Improvement. Model M0 does not include any controls. M1 adds controls for school characteristics. M2 adds controls for the demographic intake of schools. Measures of historic performance in GCSEs is added in model M3, while M4 additionally controls for historic levels of school absence. See Appendix A for full list of control variables. \* indicates difference from the “good” category statistically significant at the five percent level.

**Table 5.** The association between Ofsted inspection judgment at time of school selection and school leadership/management.

	Number of observations	Outstanding		RI		Inadequate	
		% point difference	SE	% point difference	SE	% point difference	SE
<b>(a) Same headteacher when leaving school as at previous inspection</b>							
M0	1,159	2.9%	1.8%	0.1%	1.2%	-0.2%	2.9%
M1	1,159	2.6%	1.9%	0.5%	1.2%	1.3%	2.9%
M2	1,159	1.4%	1.9%	1.3%	1.2%	3.2%	2.9%
M3	1,149	1.0%	1.9%	1.4%	1.3%	3.3%	2.9%
M4	1,149	1.0%	1.9%	1.4%	1.3%	3.3%	3.0%
<b>(b) Parental satisfaction with school leadership</b>							
M0	1,159	2.9%	1.8%	0.1%	1.2%	-0.2%	2.9%
M1	1,159	2.6%	1.9%	0.5%	1.2%	1.3%	2.9%
M2	1,159	1.4%	1.9%	1.3%	1.2%	3.2%	2.9%
M3	1,149	1.0%	1.9%	1.4%	1.3%	3.3%	2.9%
M4	1,149	1.0%	1.9%	1.4%	1.3%	3.3%	3.0%

Notes: On average, 20% of schools have the same headteacher in 2019 as the headteacher leading the school in the inspection used for school choice. Average percent of parents who strongly agree that the school is well led and managed is 44%. Figures refer to percentage point differences compared to the reference ("good") category. RI = Requiring Improvement. Model M0 does not include any controls. M1 adds controls for school characteristics. M2 adds controls for the demographic intake of schools. Measures of historic performance in GCSEs is added in model M3, while M4 additionally controls for historic levels of school absence. See [Appendix A](#) for full list of control variables. \* indicates difference from the "good" category statistically significant at the five percent level.

**Table 6.** The association between Ofsted inspection judgment at time of school selection and pupil behavior.

	Number of observations	Outstanding		RI		Inadequate	
		% point difference	SE	% point difference	SE	% point difference	SE
<b>(a) Percentage of overall absence (authorized and unauthorized)</b>							
M0	2,537	-0.8%*	0.05%	0.5%*	0.05%	1.0%*	0.09%
M1	2,537	-0.6%*	0.05%	0.5%*	0.05%	0.9%*	0.09%
M2	2,534	-0.2%*	0.04%	0.2%*	0.04%	0.4%*	0.07%
M3	2,512	-0.1%*	0.04%	0.1%*	0.04%	0.4%*	0.08%
M4	2,512	-0.1%*	0.04%	0.1%	0.04%	0.2%*	0.08%
<b>(b) Percent of parents who strongly agree that the school makes sure pupils are well-behaved</b>							
M0	1,159	6.2%*	1.6%	-0.2%	1.1%	-0.5%	2.6%
M1	1,159	5.4%*	1.6%	0.4%	1.1%	0.5%	2.5%
M2	1,159	4.3%*	1.6%	1.3%	1.1%	2.4%	2.5%
M3	1,149	3.5%*	1.6%	1.5%	1.1%	2.9%	2.6%
M4	1,149	3.6%*	1.6%	1.5%	1.1%	2.9%	2.6%

Notes: Average level of absence across schools is 5.4%. Average percent who strongly agree that the school makes sure pupils are well-behaved is 38%. Figures in panel (a) refer to the percentage point difference in the absence rate compared to the reference ("good") category. Figures in panel (b) refer to the percentage point difference in the percent of parents who strongly agree that the school makes sure pupils are well-behaved compared to the reference ("good") category. RI = Requiring Improvement. Model M0 does not include any controls. M1 adds controls for school characteristics. M2 adds controls for the demographic intake of schools. Measures of historic performance in GCSEs is added in model M3, while M4 additionally controls for historic levels of school absence. See [Appendix A](#) for full list of control variables. \* indicates difference from the "good" category statistically significant at the five percent level.

selected a good school. Such a finding could reflect parents who send their children to “Outstanding” schools being more demanding about the standard of education provided, which may in-turn be linked to their socio-economic status.

Table 5 turns to school leadership outcomes. Panel (a) focuses on whether the headteacher has changed since the last inspection by 2019 and panel (b) overall parental satisfaction with school leadership. Null results again emerge, with most estimates small and not reaching statistical significance. The one potential exception is that headteachers are somewhat more likely to have left “inadequate” schools by 2019, though the limited sample size means the estimated coefficients do not reach statistical significance at conventional levels.

Table 6 concludes with results for pupil absenteeism (panel a) and parental views of pupil behavior (panel b). From panel (a), we do observe some association between Ofsted judgments available to parents when they are selecting schools and future levels of absenteeism, although most differences are modest. In particular, outstanding schools have a 0.2 percentage point lower absence rate than good schools, while in schools requiring improvement and inadequate, the absence rate is 0.2 and 0.4 percentage points higher. (For context, the average absence rate across all schools is 5.4%). These differences across Ofsted grades are further reduced in models M3 and M4 with the introduction of controls for other sources of information available to parents.

The results from panel (b) – capturing differences in the percentage of parents who strongly agree that schools make sure pupils are well-behaved – provide little evidence of differences in behavior between good, requiring improvement and inadequate schools. On the other hand, parents whose children attend an “outstanding” school are moderately more positive about behavior than those whose children attend “good” schools (differences between “outstanding,” “requires improvement” and “inadequate” schools do not reach statistical significance at conventional levels). Together with the results from panel (a), our interpretation is that Ofsted ratings available to parents when they are choosing schools are only modestly predictive of behavior at schools in the future.

### ***Are more recent inspections of greater use to parents than more dated ones?***

Table 7 presents results from a set of regression models stratified by when the most recent inspection took place. For instance, the top row refers to inspections that were conducted up to two years before the eldest child from our hypothetical family entered secondary school (i.e. the inspection was conducted between September 2012 and October 2013). In contrast, the bottom row refers to inspections that occurred more than five years previously (i.e. it occurred before September 2009). If more recent inspection judgments are

**Table 7.** The association between Ofsted inspection judgment at time of school selection and future outcomes. Separate estimates by length since last inspection.

Length before child starts school	N	Outstanding		RI		Inadequate	
		Effect size	SE	Effect size	SE	Effect size	SE
<b>(a) Attainment 8</b>							
Up to 2 years	988	0.19*	0.04	-0.08*	0.03	-0.11*	0.05
2 to 3 years	614	0.17*	0.05	-0.06	0.04	0.14	0.09
3 to 4 years	370	0.14*	0.05	-0.02	0.09	-0.20	0.23
4 to 5 years	145	0.14	0.08	-0.03	0.12	-0.51	0.41
More than 5 years	349	0.03	0.04	-0.08	0.14	-	-
<b>(b) Progress 8</b>							
Up to 2 years	985	0.34*	0.08	-0.13*	0.06	-0.18*	0.09
2 to 3 years	612	0.33*	0.10	-0.13	0.08	0.29	0.17
3 to 4 years	370	0.24*	0.09	-0.12	0.17	-0.62	0.45
4 to 5 years	145	0.26	0.15	-0.14	0.22	-0.61	0.76
More than 5 years	349	0.13	0.08	-0.37	0.26	-	-
<b>(c) Absence levels</b>							
Up to 2 years	1,027	-0.3%*	0.09%	0.1%	0.06%	0.3%*	0.10%
2 to 3 years	629	-0.4%*	0.08%	0.1%	0.07%	0.4%*	0.14%
3 to 4 years	379	-0.1%	0.09%	0.4%*	0.17%	-0.3%	0.44%
4 to 5 years	147	-0.1%	0.16%	0.6%*	0.23%	0.3%	0.80%
More than 5 years	352	-0.3%*	0.10%	1.2%*	0.31%	-	-

Notes: Estimates based upon model M2, which controls for school background characteristics and the demographic intake of schools. RI = Requiring Improvement. See Appendix A for full list of control variables. Figures refer to effect size differences compared to "good" schools as the reference category. \* indicates difference from the "good" category statistically significant at the five percent level.

indeed more useful for parents when choosing schools than older ones, then one would expect the estimated effect sizes to decline in absolute magnitude when moving down the rows. These estimates are all based upon our preferred model specification (M2), with a focus upon the relationship between Ofsted judgments and school performance measures (panels a and b) and school absence levels (panel c).<sup>14</sup>

There is little clear pattern for the inadequate and requires improvement judgments; effect sizes are not clearly higher or lower for more recent or more distant inspections (although sample sizes for these categories are small for most recent inspections being conducted more than three years ago). However, there is some suggestion that more recently awarded outstanding judgments may be somewhat more predictive than those awarded some time ago. Take Progress 8 scores (panel b) for instance. Schools rated as outstanding within the last two years achieved Progress 8 scores 0.34 standard deviations higher than schools rated as good during the same period. The difference becomes notably smaller when the good/outstanding grades were based on inspections that took place more than five years prior to when the eldest child started secondary school (0.13) and is not statistically significant at conventional levels. A broadly similar pattern can also be observed for the Attainment 8 and school absence outcomes. Thus, building upon the findings above, an outstanding judgment may be of some use to parents when they are selecting schools, but only when the inspection has been conducted relatively recently.

## Discussion

In theory, the reasons why Ofsted judgments are so widely used by parents when making schooling decisions are sound. Ofsted provides an independent, expert view of the school judged according to a common framework, providing a unique perspective of what life is like “on the ground.” Yet, in practice, their value to inform school choice is not so clear. As we have set out, there may be sizable time-lags between when schools were last inspected, when parents make their schooling decisions and when young people attend them. Such time-lags are further compounded by parents having strong incentives to send their offspring to the same school. Much about a school can change in the interim – including the composition of teaching staff and school leaders. There are also questions as to whether Ofsted judgments “add-value” to parents selecting schools over other freely available information (e.g. prior performance in national examinations).

The main contribution of this paper – in addition to clearly spelling out such issues – is to provide empirical support for how much parents can learn from an Ofsted report when choosing schools. In particular, we document the length of such time-lags, key factors that change in the interim (e.g. the headteacher) and whether inspection judgments available to parents when choosing schools can predict the academic environment and attainment when their offspring actually attend. For our hypothetical family, our analysis illustrates how inspection judgments are, on average, three years old by the time the school they select starts to impact their eldest-born (and eight-years-old by the time that their eldest leaves). This average value masks substantial heterogeneity in the timeliness of information available to parents about different schools, with those with certain characteristics (e.g. selective grammar schools) more heavily affected than others. It is hence likely that, for most of the time that a family’s children attend the school, it will be led by someone other than the headteacher named within the relevant inspection report. Moreover, the written inspection reports may also provide different information, given changes made to inspection frameworks and activities over time.

Our empirical analysis then demonstrates how Ofsted inspection judgments available to parents when they are choosing a school are weak predictors of the future school environment and outcomes. Indeed, GCSE grades, pupil absenteeism and parental satisfaction are very similar across parents who choose good, requiring improvement and inadequate schools, once differences in their pupil intakes have been controlled. Moreover, although an “outstanding” judgment does seem to provide some future predictive power, the “added-value” over other pieces of freely available information is limited, and only applies when the Ofsted inspection was relatively recent. Overall, these results lead us to conclude that – despite their widespread use in school selection – inspection judgments are not actually that informative for parents.

This conclusion should be interpreted considering the limitations of this work. First, our empirical analysis has focused on secondary school choice. Although most of our arguments apply to primary school choice as well, empirical evidence has not been presented on this matter. Second, some of the outcomes we have explored (e.g. parental satisfaction with schools) are limited in terms of response rates and only being available for a sub-sample of schools (those inspected in 2018 or 2019). Third, our statistical models have controlled for a limited array of observable characteristics. To the extent that there are other aspects of schools' intakes that we have not been able to control, our preferred model specification (M2) is likely to provide an upper-bound on the future predictive power of Ofsted grades. Fourth, the inspections we considered were all conducted under previous iterations of the Ofsted inspection framework, which was revised in 2019. It is hence possible that different results could emerge in the future. Yet as we have argued, this simply reflects the problems parents face when selecting schools – they must make their decisions using potentially dated information with limited comparability across their choice set. Fifth, for simplicity, we have not explicitly considered catchment areas and the choice set available to parents based upon where they live.

Notwithstanding these limitations, we believe our findings have some important implications for parents and policymakers. In our view, they do not illustrate problems with Ofsted inspections per se. In line with what many would expect from a school inspectorate, empirical evidence suggests that Ofsted inspections resulting in an “inadequate” judgment improve the outcomes of such schools (Hussain, 2017). Given this, we should not expect or even want Ofsted judgments to predict what such a school will be like in the medium term. Thus, the fact that once inadequate and requiring improvement schools become largely indistinguishable from their good counterparts may be a sign of success. Moreover, Ofsted judgments are likely to be of interest to parents whose children are *currently* attending a school, providing them with important information about what it is (and is not) currently doing well. Rather, our results provide a cautionary tale about using Ofsted judgments to inform school choice. Our advice to parents is to not place too much emphasis on them. While they may act as a catalyst for thinking about differences between schools, and perhaps some insight into a school's ethos, they are not going to provide much information about the academic environment and the outcomes of pupils during the period when their children will be going there.

## Notes

1. <https://www.gov.uk/government/publications/education-inspection-framework/education-inspection-framework>
2. By choosing this hypothetical scenario we realize that a case for generalizability needs to be made, which we do in the conclusion.

3. We have chosen this date so that our scenario avoids – as far as possible – complications surrounding the COVID pandemic, while still being relatively recent.
4. For simplicity, we abstract away from “catchment areas” or feasible choice sets available to parents. Our analysis therefore all conducted at the school level.
5. <https://www.compare-school-performance.service.gov.uk/download-data>)
6. <https://www.gov.uk/government/statistical-data-sets/monthly-management-information-ofsted-school-inspections-outcomes>
7. Most Ofsted reports are published 38 working days after the end of the inspection (<https://www.gov.uk/guidance/inspecting-schools-guide-for-maintained-and-academy-schools>), though there may be delays if schools challenge the findings.
8. Before September 2012, the “satisfactory” Ofsted grade was renamed “requires improvement”. We do not make a distinction between these two terms, and thus use “requires improvement” for consistency.
9. The younger child in our hypothetical family would take their GCSEs even later (2021). We are however unable to investigate GCSE outcomes later than 2019 due to the COVID-19 pandemic.
10. Figures are used from the “strongly agree” category for each question.
11. Authors’ calculations using the PISA 2015 database.
12. These requests contain the named headteacher in August 2010, December 2011, August 2012, March 2013, May 2014, October 2015, May 2016, September 2017, March 2018 and January 2019.
13. There is a small amount of missing information for some academic years where school headteacher could not be linked.
14. We focus upon these outcomes due to their larger sample sizes.

## Disclosure statement

John Jerrim is currently on part-time secondment at Ofsted. This research was conducted independently of that role, and solely through his academic position at UCL.

## ORCID

John Jerrim  <http://orcid.org/0000-0001-5705-7954>

## References

- Allen, R., & Burgess, S. (2010). *The future of competition and accountability in education*. University of Bristol: 2020 Public Services Trust at the RSA.
- Ball, S. J., Bowe, R., & Gewirtz, S. (1995). Circuits of schooling: A sociological exploration of parental choice of school in social class contexts. *The Sociological Review*, 43(1), 52–78. doi:10.1111/j.1467-954X.1995.tb02478.x
- Ball, S. J., & Vincent, C. (1998). ‘I Heard It on the Grapevine’: ‘hot’ knowledge and school choice. *British Journal of Sociology of Education*, 19(3), 377–400. doi:10.1080/0142569980190307
- Bokhove, C., Jerrim, J., & Sims, S. (2023). Are some school inspectors more lenient than others? QSS working papers <https://www.ucl.ac.uk/ioe/departments-and-centres/centres/quantitative-social-science/working-papers>



- Bradley, S., Crouchley, R., Millington, J., & Taylor, J. (2000). Testing for quasi-market forces in secondary education. *Oxford Bulletin of Economics and Statistics*, 62(3), 357–390. doi:10.1111/1468-0084.00176
- Burgess, S., Briggs, A., McConnell, B., & Slater, H. (2006). *School choice in England: Background facts*. University of Bristol: Centre for Market and Public Organisation.
- Burgess, S., Greaves, E., & Vignoles, A. (2019). School choice in England: Evidence from national administrative data. *Oxford Review of Education*, 45(5), 690–710. doi:10.1080/03054985.2019.1604332
- Burgess, S., Greaves, E., & Vignoles, A. (2020). *School places: A fair choice? School choice, inequality and options for reform of school admissions in England*. London: Sutton Trust.
- Burgess, S., Greaves, E., Vignoles, A., & Wilson, D. (2015). What Parents Want: School Preferences and School Choice. *Economic Journal*, 125(587), 1262–1289. doi:10.1111/eoj.12153
- Education Datalab, F. F. T. (2015). *Floors, tables and coasters: Shifting the education furniture in England's secondary schools*. FFT education. Accessed 11 Jan 2022: <https://mk0ffteducation79fru.kinstacdn.com/wp-content/uploads/2016/02/2015-Educationfurniture-04.pdf>
- Hussain, I. (2012). School inspections: Can we trust Ofsted reports? *LSE Centrepiece Winter 2011/12*. Accessed 23 Aug 2021 from <https://cep.lse.ac.uk/pubs/download/cp358.pdf>
- Hussain, I. (2015). Subjective performance evaluation in the public sector: Evidence from school inspections. *Journal of Human Resources*, 50(1), 189–221. doi:10.1353/jhr.2015.0006
- Hussain, I. (2017). *Do consumers respond to short-term innovations in school productivity? Evidence from the housing market and parents school choices*. Accessed 10 Sep 2021 from <https://www.sole-jole.org/assets/docs/16502.pdf>
- Hutchinson, J. (2016). *School inspection in England: Is there room to improve?* Education policy institute. Accessed 23 Aug 2021 from <https://epi.org.uk/wp-content/uploads/2018/01/school-inspection-in-england-web.pdf>
- Jabbar, H., Fong, C. J., Germain, E., Li, D., Sanchez, J., Sun, W.-L., & Devall, M. (2019). The competitive effects of school choice on student achievement: A systematic review. *Educational Policy*. 10.1177/0895904819874756
- Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, 49(4), 241–253. doi:10.3102/0013189X20912798
- Leckie, G., & Goldstein, H. (2009). The limitations of using school league tables to inform school choice. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(4), 835–851. doi:10.1111/j.1467-985X.2009.00597.x
- Malcolm, H., Wilson, V., Davidson, J., & Kirk, S. (2003). *Absence from school: A study of its causes and effects in seven LEAs*. DfES Research Report DfES RR 424. <http://www.education.gov.uk/research/data/uploadfiles/RR424.pdf>
- National Audit Office [NAO]. (2018). *Ofsted's inspection of schools*. National Audit Office. Accessed 23 Aug 2021 from <https://www.nao.org.uk/wp-content/uploads/2018/05/Ofsted-inspection-of-schools.pdf>
- OECD. (2017). *School choice and school vouchers: An OECD perspective*. Accessed 24 Aug 2021 from <https://www.oecd.org/education/School-choice-and-school-vouchers-an-OECD-perspective.pdf>
- Ofsted. (2017). *Do two inspectors inspecting the same school make consistent decisions?* Ofsted research report 170004. Accessed 23 Aug 2021 from <https://www.gov.uk/government/publications/do-two-inspectors-inspecting-the-same-school-make-consistent-decisions>
- Ofsted. (2019). *Retaining the current grading system in education: Some arguments and evidence*. Ofsted research report 190012. Accessed 24 Aug 2021 from [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/936220/Retaining\\_the\\_current\\_grading\\_system\\_-\\_arguments\\_and\\_evidence\\_290419.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/936220/Retaining_the_current_grading_system_-_arguments_and_evidence_290419.pdf)

- Prior, L., Jerrim, J., Thomson, D., & Leckie, G. (2021). A review and evaluation of secondary school accountability in England: Statistical strengths, weaknesses, and challenges for 'Progress 8.' *Review of Education*, 9(3). doi:10.1002/rev3.3299
- Sibieta, L. (2006). Choice and competition in education markets. *Economic Review*, 24(2). <https://ifs.org.uk/publications/3867>. Accessed 24 Aug 2021 from.
- Sibieta, L., & Jerrim, J. (2021). *A comparison of school institutions and policies across the UK*. Education Policy Institute. Accessed 23 Aug 2021 from <https://epi.org.uk/wp-content/uploads/2021/04/EPI-UK-Institutions-Comparisons-2021.pdf>
- Von Stumm, S., Smith-Woolley, E., Cheesman, R., Pingault, J.-B., Asbury, K., Dale, P. S., . . . Plomin, R. (2021). School quality ratings are weak predictors of students' achievement and well-being. *Journal of Child Psychology and Psychiatry*, 62(3), 339–348. doi:10.1111/jcpp.13276
- YouGov. (2017). *Annual Parents Survey 2017*. Accessed 23 Aug 2021 from [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/698913/Ofsted\\_Annual\\_Parents\\_Survey\\_2017.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/698913/Ofsted_Annual_Parents_Survey_2017.pdf)

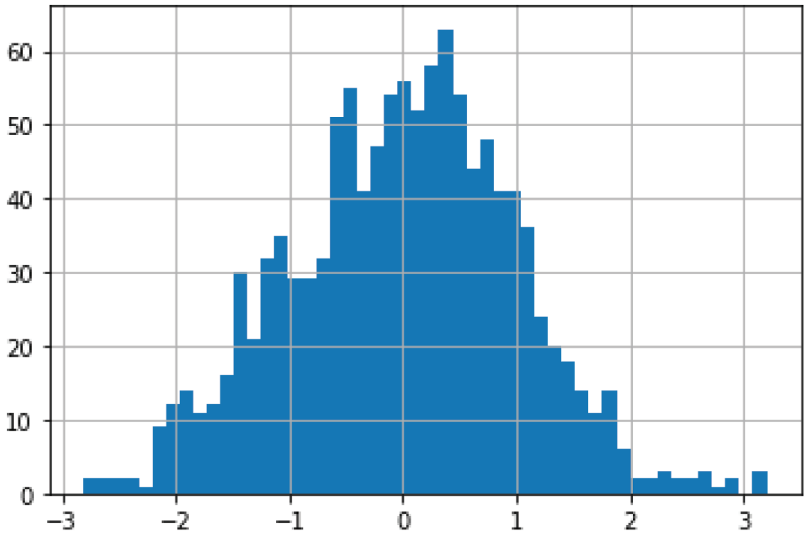
## Appendix A. Full details on controls included in each model specification

	M0	M1	M2	M3	M4
Admissions policy	-	Y	Y	Y	Y
School type	-	Y	Y	Y	Y
School gender	-	Y	Y	Y	Y
% children with English as Additional Language (averaged across 2015–2019)	-	-	Y	Y	Y
% children with Special Educational Needs (averaged across 2015–2019)	-	-	Y	Y	Y
% children ever eligible for Free school Meals (averaged across 2015–2019)	-	-	Y	Y	Y
Key stage 2 average point score of prior cohorts (who entered secondary school in 2011,2012 and 2013)	-	-	Y	Y	Y
Key stage 2 average point score for the cohort in question (who entered secondary school in 2014)	-	-	Y	Y	Y
% of pupils who achieved 5 A*-C GCSEs in 2014	-	-	-	Y	Y
Capped average GCSE points score in 2013	-	-	-	Y	Y
Capped GCSE progress scores in 2013	-	-	-	Y	Y
% absences in 2013	-	-	-	-	Y

## Appendix B. Additional details about the parental satisfaction with school scale

Table B1. Loadings from the principal components analysis for the parental satisfaction with school scale

	Component										
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th
Q1. My child is happy at this school	0.31	-0.32	-0.31	-0.20	0.09	0.03	0.15	0.35	0.64	0.20	-0.25
Q2. My child feels safe at this school.	0.30	-0.38	-0.21	-0.35	0.03	0.21	0.18	-0.04	-0.33	-0.01	0.63
Q3. My child makes good progress at this school.	0.30	0.27	-0.39	-0.02	-0.03	-0.43	-0.02	0.10	-0.46	0.49	-0.20
Q4. My child is well looked after at this school.	0.31	-0.30	-0.10	0.02	0.18	-0.11	0.07	-0.22	-0.27	-0.61	-0.51
Q5. My child is taught well at this school.	0.30	0.42	-0.20	-0.06	-0.06	-0.42	-0.16	-0.17	0.39	-0.40	0.38
Q6. My child receives appropriate homework for their age.	0.29	0.47	0.00	-0.15	0.49	0.58	-0.27	0.05	-0.04	0.00	-0.09
Q7. This school makes sure its pupils are well behaved.	0.30	0.12	0.36	-0.44	-0.51	0.14	0.11	-0.44	0.08	0.18	-0.23
Q8. This school deals effectively with bullying.	0.29	0.04	0.69	-0.04	0.28	-0.33	0.28	0.40	-0.06	-0.02	0.10
Q9. This school is well led and managed.	0.30	-0.19	0.13	0.18	-0.46	0.10	-0.64	0.42	-0.10	-0.11	0.02
Q10. This school responds well to any concerns I raise.	0.30	-0.30	0.16	0.45	0.29	-0.07	-0.24	-0.51	0.16	0.37	0.12
Q11. I receive valuable information from the school about my child's progress.	0.30	0.21	-0.10	0.61	-0.28	0.32	0.54	0.06	0.01	-0.07	0.05



**Figure B1.** The distribution of the parental school satisfaction scale.