Parental Responses to Information About School Quality: Evidence from Linked Survey and Administrative Data*

Ellen Greaves Iftikhar Hussain Birgitta Rabe Imran Rasul†

November 2019

Abstract

Multiple inputs determine children's cognitive achievement. We study the interaction between family and school inputs by identifying the causal impact of information about school quality on parental time investment into children. Our setting is England, where credible information on school quality is provided by a nationwide school inspection regime. Schools are inspected at short notice, with school ratings using hard and soft information. As such soft information is not necessarily known to parents *ex ante*, inspection ratings provide news to parents that plausibly shifts inputs into their children. We study this using household panel data linked to administrative records on school performance and inspection ratings. We observe some households being interviewed pre school inspection (control), and others being interviewed post inspection (treatment). Treatment assignment is determined by a household's survey date relative to the school inspection date, and shown to be as good as random. We find that parents receiving good news over school quality significantly decrease time investment into their children (relative to parents that will later receive such good news). Hence on average, beliefs over school quality and parental inputs are substitutes. Our data and design allow us to provide insights on the distributional and test score impacts of the nationwide inspections regime, through multiple margins of endogenous response of parents and children. Our findings highlight the importance of accounting for interlinked private responses by families to policy inputs into education. *JEL Classification: I20, I24.*

*We gratefully acknowledge financial support from the ESRC through the Research Centre on Micro-Social Change (ES L009/53/1), Understanding Society (ES/H029745/1) and Centre for the Microeconomic Analysis of Public Policy (RES-544-28-0001). We thank Jerome Adda, Oriana Bandiera, Lorenzo Casaburi, Claire Crawford, Christian Dustmann, Alex Eble, Brian Jacob, Bentley Macleod, Emily Nix, Paolo Pinotti, Dina Pomeranz, Uta Schoenberg, Sami Stouli, Adam Szeidl, Guido Tabellini, Daniel Willhelm, Miguel Urquiola and numerous seminar and conference participants for valuable comments. All errors remain our own.

†Greaves: Bristol and IFS, ellen.greaves@bristol.ac.uk; Hussain: Sussex, iftikhar.hussain@sussex.ac.uk; Rabe: ISER and Essex, brabe@essex.ac.uk; Rasul: UCL and IFS, i.rasul@ucl.ac.uk.

1 Introduction

Family and school based inputs determine children's cognitive achievement. A vast literature has for decades studied family- and school-based inputs, yet despite it having long been recognized that family and school inputs can be substitutes or complements [Becker and Tomes 1976, Todd and Wolpin 2003], far less is known on interactions between family and school inputs.

We address the issue by identifying the causal impact of the exogenously release of information on school quality on parental inputs into children. Our study context is England, where a source of credible information on school quality is an established nationwide school inspection regime. We thus provide novel insights from a high-income context on the crowding in/out of private household inputs into children's attainment by beliefs held over school quality. Given the global roll out of school accountability regimes and widespread use of report cards [Figlio and Loeb 2011], this is a relevant issue for education systems around the world. By exploiting a nationwide regime, we document the impact information on school quality has on the distribution of parental inputs into children across households, and the ultimate impacts on children's high stakes exam scores.

English school inspections are conducted by the Office for Standards in Education (Ofsted). Schools are typically inspected every four years. Inspections occur at short notice: schools are told one or two days in advance, so there is little opportunity to game the system. Inspections are intense, lasting up to five days, and gather information from multiple sources including: (i) in-class observation of teaching; (ii) interviewing the school leadership team; (iii) reading students' books; (iv) speaking directly to parents. A school's assessment is based on hard performance data (test scores) and a wealth of qualitative evidence gathered by inspectors during their visit. Inspections thus place weight on dimensions of school quality that parents value and that correlate to school value added, and are based on soft information not necessarily known to parents *ex ante* [Jacob and Lefgren 2007]. There can thus be informational content in school inspection ratings, that plausibly shifts parental inputs into their children.[1]

Schools are given an inspection rating on a four-point scale. These are immediately disseminated to parents, and a full inspection report is quickly made available online. Parents respond to inspection ratings if there is new information in them relative to their prior beliefs. To construct prior beliefs we use a simple model to forecast a school's inspection rating based on publicly available information, including the school's past test score results. We use the model to define whether the inspection rating reveals good, bad or no news to parents about school quality.[2]

[1] There is a body of evidence highlighting parents have imperfect information over many aspects of human capital accumulation process of their children (Dizon-Ross 2019 overviews this literature). In the US, the No Child Left Behind Act 2002 required states to test students in reading and maths in grades 3 to 8, and in high school, building on a pre-existing system in which 45 states published report cards. There are state or district variations in NCLB provisions making it hard to draw implications for outcomes nationwide, and the system is based on the release of hard information: the UK system is uniform across the country and is based on hard and soft information.

[2] There are few papers that measure the news content of inspection outcomes: two notable exceptions are Rouse *et al.* [2007] and Feng *et al.* [2010] who build 'accountability shocks' in the context of NCLB.

To study the impact this news has on parental behavior, we exploit household panel data from the UK Household Longitudinal Survey (UKHLS), that records parental inputs into their children, as well as children's own inputs into their education. Uniquely, we are able to link this survey data to administrative records on school performance and inspection ratings. Our research design exploits the fact that: (i) school inspections can take place in any month during the academic year; (ii) household survey interviews can take place in any month. Hence in our linked household-school administrative data, we observe some households being interviewed prior to their school being inspected (the control group), and some being interviewed post inspection (the treated group). Treatment assignment is thus determined by the date a household is interviewed in the survey data relative to the date their school is inspected.

We provide a battery of evidence to suggest this treatment assignment is as good as random.

Our research design can then be summarized as follows. Consider the set of schools inspected in a given year t, and *hold constant* whether parents will receive good, bad or no news. The control group are households interviewed in survey year t but prior to the inspection actually taking place (holding constant that good, bad or no news will eventually be received by these parents). Treated households are *also* in schools inspected in the same year, but happen to be interviewed after the inspection takes place (and so know the news on school quality). Both sets of households are observed over time, and have children attending schools inspected in the same academic year. The key difference between them is that treated households *know* the inspection outcome and so hold posterior beliefs about school quality, while control households do not, and so hold prior beliefs about school quality.

The identifying assumptions needed to deliver causal impacts of information on school quality on parental inputs are: (i) there is no selection of schools by month of inspection; (ii) there is no selection of households by month of interview; (iii) there are no natural time trends in *changes* in parental input; (iv) there are no within school-year responses to inspections by schools. We provide evidence to underpin each assumption, drawing on multiple tests and data sources.

We develop a stylized framework to make precise the assumptions needed to back out whether beliefs about school quality and parental inputs are complements or substitutes from the heterogeneous treatment effect estimates. This is key to understanding: (i) the wedge between experimental and total policy effects of changing any school-based input [Todd and Wolpin 2003]; (ii) the distributional and test score impacts of the school inspection regime.

Our core result is that when parents receive good news about school quality they are significantly less likely to increase time inputs into their children (relative to parents in the control group who will receive the same news later in time). This implies that for the average household, beliefs about school quality and parental time investment are substitutes. In contrast, when parents receive bad news about school quality, their time input does not change.

We discuss two further insights our data and design provide on the nationwide inspections regime: (i) its distributional impacts across households; (ii) the impact it has on test scores

through multiple endogenous responses of parents and children.

The distributional impacts of school quality information depend on how good and bad news shocks relate to *ex ante* school quality. Given our forecasting model, we show that good and bad news shocks are evenly distributed across schools of different *ex ante* quality. Calibrating a simple model of parental investments shows given this distribution of news across schools, the impact of the information released by the inspection regime is to: (i) reduce the expected level of parental inputs by 14%; (ii) reduce across-school inequality in parental inputs by 15%. As parental inputs and beliefs about school quality are substitutes, the mechanism driving this is that parents with good news reduce inputs by more than parents receiving bad news, thus reducing inputs overall. Given the distribution of news across schools, parental inputs fall more in higher ranked schools, thus reducing across-school input inequality.

On the second issue of how households' multiple responses to information ultimately impact test scores, a key advantage of the UKHLS data is that a wide range of parental and child outcomes can be studied. We find that children's time inputs move in the opposite direction to the behavioral response of parents: when a household receives good news about school quality, children are significantly more likely to *increase* time spent on homework. In other words, children partly compensate for the loss of parental input by increasing their own time investment, so their effort is complementary to beliefs about school quality.

We then follow Todd and Wolpin [2003] and Pop-Eleches and Urquiola [2013] to set out a second framework making precise what can be inferred about the relative total products of these various input margins in producing test scores. Mapping this framework to an empirical specification, we estimate test score impacts of the school inspection regime using a similar research design as before, comparing end of academic year test scores between children in schools inspected early in the academic year, to those whose schools are inspected later in the academic year (but still prior to the exam period). We implement this by linking the administrative schools data with individual administrative data on test scores of $200,000$ children in nationwide high stakes exams taken at age 16.

We find the receipt of *good* news generated by school inspections early in the academic year significantly *lowers* test scores. Matching this to the earlier findings, this suggests that as good news causes parents to reduce their time input and children to increase their time input, children's own time investment into their homework has a lower total product in generating test scores than their parent's time investment.[3]

Our work provides novel insights for two literatures. As described above, while there is a voluminous literature studying parental, family, and school inputs into children's achievement, far less is known about interactions between these inputs. This is surprising because: (i) there

[3]Jacob [2005], Hanushek and Raymond [2005], Figlio and Loeb [2011] and Burgess *et al.* [2013] study test score impacts of school accountability regimes. Long run impacts of attending high rated schools on college attendance, completed four year degrees and earnings at age 25 have also been documented [Deming *et al.* 2016].

is long-standing literature in public economics on public-private crowd in/out, but this issue has been less studied in educational contexts; (ii) input interactions are at the heart of the rapidly growing literature on early (pre-school) childhood development [Cunha *et al.* 2010]. Our work adds to the small literature on family- and school-based input interactions, that has however been focused on how parents respond to specific school inputs, such as class size [Datar and Mason 2008, Fredriksson *et al.* 2016] or school resources [Houtenville and Conway 2008], or has been located in middle and lower-income countries [Pop-Eleches and Urquiola 2013, Das *et al.* 2013]. Studying a high income context can shed new light on this old issue because: (i) the information parents have on school quality will likely differ to lower-income contexts; (ii) the mechanism by which households sort into schools are also very different across these contexts.[4]

On parental responses to school accountability systems, the current literature largely focuses on 'extensive margin' school choice or house price responses as information on school quality is released [Figlio and Lucas 2004, Figlio and Loeb 2011, Hastings and Weinstein 2008, Hussain 2017]. In sharp contrast, this paper examines the 'intensive margin' of parental responses to school quality ratings for children that are already in school. These margins of impact are understudied, but affect a far larger cohort of parents (those with children in any school grade), than those facing an initial school choice problem.[5,6] Such policies can also reinforce/mitigate inequalities within and across schools and families, as we document.

The paper is organized as follows. Section 2 develops a stylized framework to understand how parental inputs vary with information about school quality. Section 3 describes our linked household and school administrative data. Section 4 presents our research design and identifying assumptions. Section 5 contains our core findings on parental responses to news, and calibrates the distributional impacts of the inspections regime. Section 6 examines the impact it has on test scores through multiple endogenous responses of parents and children. Section 7 concludes. The Appendix contains proofs and further results.

2 Conceptual Framework

2.1 Set-up

We present a simple model to highlight how signals of school quality impact parental time investment into their children. Parents are assumed to invest in one child, be uncertain over school

[4]Other studies have highlighted the role of parental time investments [Cunha *et al.* 2010, Del Boca *et al.* 2014, Fiorini and Keane 2014, Carneiro *et al.* 2015, Del Bono *et al.* 2016]. We build on the small literature examining parent-child input interactions in education [De Fraja *et al.* 2010, Del Boca *et al.* 2017, Caetano *et al.* 2019].

[5]A notable exception is Figlio and Kenny [2009] who find that positive information from school accountability regimes raises parental financial contributions to schools.

[6]We add to work examining parental responses to information: much of this relates to information over a child's ability [Dizon Ross 2019], population averaged returns to schooling [Jensen 2010, Hoxby and Avery 2012], or hard information on school test scores [Hastings and Weinstein 2008, Andrabi *et al.* 2017].

quality, and use Bayesian updating for their beliefs about school quality. We assume schools face short run adjustment costs and so do not immediately respond to information on school quality (an assumption validated below).[7]

Parents choose how to allocate their time between investments into their child (e) and leisure (l) subject to a time constraint: $1 = e + l$. School quality, s, is unknown and can be high (s_H) or low (s_L) where parental priors are prob$(s = s_H) = p_H$, and prob$(s = s_L) = p_L = 1 - p_H$. Parental utility is $U = u(e, l, s)$ where $u(.)$ is concave in each argument and we assume independence between the marginal utility of leisure and school quality ($u_{ls} = 0$). We make no assumption on the cross derivative $u_{es} \lesseqgtr 0$, so that parental input and beliefs about school quality can be substitutes, complements or independent. This framework makes precise the assumptions needed to identify $sign(u_{es})$. This depends on the crowding in/out of parental inputs as beliefs about school quality change, and has wider implications for how the school inspection regime impacts inequality of parental inputs within and across schools.

Parents choose e to maximize expected utility:

$$\max_e u(e, 1-e, s) = \max p_H u(e, 1-e, s_H) + (1-p_H)u(e, 1-e, s_L). \tag{1}$$

To see how parental investments vary in *prior beliefs* on school quality s, we totally differentiate the first order condition from (1) to obtain $\frac{de}{dp_H} = -\frac{1}{\Delta}[u_e(s_H) - u_e(s_L)]$, where $\Delta < 0$ (as shown in the Appendix). We thus have the following:

Result 1: If e and s are substitutes ($u_{es} < 0$), then $\frac{de}{dp_H} < 0$; if e and s are complements ($u_{es} > 0$), then $\frac{de}{dp_H} > 0$.

In the same school there will be across-family differences in parental effort e (given school quality s), variation driven by differences in: (i) prior beliefs p_H; (ii) preferences ($u_{es} \lesseqgtr 0$).

2.2 Signal of School Quality

The Ofsted school inspection regime provides an informative signal to parents of school quality. We model the signal as taking value $j \in \{H, L\}$ and assume it is symmetric so $prob(j = H|s = s_H) = prob(j = L|s = s_L) = q$. The probability the signal is correct is $q > .5$ (so better than a

[7] This assumption also matches our institutional setting. For example, in England schools make staffing decisions towards the end of each academic year. Evidence in favor of such short run adjustment costs has been documented for the US and the UK [Rouse *et al.* 2013, Hussain 2015].Of course in the longer term, school accountability systems might well impact teacher turnover [Feng *et al.* 2010, Figlio and Loeb 2011, Dizon-Ross 2018].

coin toss). Parents update their priors to the following posteriors after observing signal j:

$$
\begin{aligned}
prob(s = s_H | j = L) &= \frac{p_H(1-q)}{(1-p_H)q + p_H(1-q)}, \\
prob(s = s_L | j = L) &= \frac{(1-p_H)q}{(1-p_H)q + p_H(1-q)}, \\
prob(s = s_H | j = H) &= \frac{p_H q}{(1-p_H)(1-q) + p_H q}, \\
prob(s = s_L | j = H) &= \frac{(1-p_H)(1-q)}{(1-p_H)(1-q) + p_H q}.
\end{aligned} \tag{2}
$$

In response to good news ($j = H$), the change in parental input is,

$$
de|_{j=H} = K dp_H = K p_H [\frac{(1-p_H)(2q-1)}{(1-p_H)(1-q) + p_H q}], \tag{3}
$$

where $K = -\frac{1}{\Delta}[u_e(s_H) - u_e(s_L)]$. As $q > .5$ this yields the following:

Result 2A: If parents receive good news: (i) parental effort falls if e and s are substitutes (as $K < 0$); (ii) parental effort rises if e and s are complements (as $K > 0$).

In response to bad news, the change in parental input is,

$$
de|_{j=L} = K dp_H = K p_H [\frac{(1-p_H)(1-2q)}{(1-p_H)q + p_H(1-q)}]. \tag{4}
$$

Result 2B: If parents receive bad news: (i) parental effort rises if e and s are substitutes (as $K < 0$); (ii) parental effort falls if e and s are complements (as $K > 0$).

Mapping both results to our empirical research design, they correspond to differences between households with and without signals of school quality revealed by the inspections regime. A final result useful for interpreting our findings relates to the asymmetry of parental responses:

Result 3: In the knife-edge case of $p_H = 0.5$, parents respond symmetrically to good and bad news ($\left| de|_{j=H} \right| = \left| de|_{j=L} \right|$). If $p_H > 0.5$, parents respond more to bad news $\left| de|_{j=H} \right| < \left| de|_{j=L} \right|$, and if $p_H < 0.5$, parents respond more to good news $\left| de|_{j=H} \right| > \left| de|_{j=L} \right|$.

This helps us back out the parental priors in order to rationalize the findings.

The model makes precise how our study of school and family input interactions might differ from earlier work based on low- and middle-income contexts [Pop-Eleches and Urquiola 2013, Das *et al.* 2013]. In particular, parental priors p are likely to differ (hence the response to new information will differ) because the market for information on school quality is better developed, and because mechanisms enabling households to sort into schools differ across these contexts.

3 Context and Data

3.1 The Inspections Regime

School inspections are conducted by the Office for Standards in Education (Ofsted). The objectives of the regime are to [Johnson 2004]: (i) offer feedback to school principals and teachers; (ii) identify schools suffering serious weaknesses; (iii) provide information to parents to aid their decision-making. Under the Ofsted regime, schools are typically inspected once every few years. Inspections occur at short notice: schools are told one or two days in advance, so there is little opportunity for them to game the system. Inspections occur throughout the academic year (September through to July), and we exploit this continuous timing in our research design.[8]

Inspections are intense and gather information from multiple sources: during our sample period, they last up to five days and the components of information gathered are: (i) in-class observation of teaching; (ii) interviews with the school leadership team; (iii) inspecting students' books; (iv) speaking directly to parents. A school's rating is based on hard performance data (namely, test scores) and a wealth of qualitative evidence gathered by inspectors during their visit. Table A1 details Ofsted grade descriptors. These are complex, multi-dimensional and heavily based on qualitative information. Inspections place weight on dimensions of school quality that parents and educational stakeholders value, are correlated to measures of school value-added, and such soft information is not necessarily known to parents *ex ante* [Jacob and Lefgren 2007, Burgess *et al.* 2015, Beuermann *et al.* 2018]. There can thus be informational content in school ratings, constituting news to parents that shifts investments into their children.[9,10]

3.1.1 Ratings and Dissemination

The inspection awards schools a headline rating on a four-point scale: 4 (Outstanding), 3 (Good), 2 (Requires Improvement) and 1 (Inadequate/failing). This maps to the signal s of school quality in the framework above. These ratings are immediately disseminated to all parents via a letter, and a full inspection report is made available online within 3 to 4 weeks. Given the immediate and widespread dissemination of inspection ratings, there is near perfect compliance among treated households: once an inspection is conducted at their school, households will be informed about the headline school rating. Figure A1 provides an example of an Ofsted letter sent to parents. The letter is simple, concise, and clearly states the headline inspections rating.

[8] Schools have been subject to regular inspection by Ofsted in the English state education system since the early 1990s. In the pre-2005 inspection regime (before our study period), schools were inspected for a week every six years, with two months' notice.

[9] High-stakes nationwide exams are taken in England at ages 11 and 16. Exam scores are a key measure of performance used by the Department for Education and form the basis of school level exam league tables. Hard information on exam score outcomes and rankings is freely available online to parents.

[10] There is evidence of gaming of accountability regimes in US contexts where regimes are far more based on hard information [Jacob and Levitt 2003, Figlio and Loeb 2011].

In short, headline inspection ratings are highly salient to parents. This is corroborated by survey evidence: an annual survey of parents undertaken by YouGov since 2015 shows 90% of parents are aware of their child's school's inspection rating.[11]

Beyond this rating, four sub-components of activity are also rated by inspectors: achievement of pupils, quality of teaching, behavior and safety of pupils, and leadership and management. While these aspects might also mentioned in the letter, we do not use these sub-component ratings because only a selected group of parents are likely to be aware of such fine-grained inspection results. Part A of Table A1 shows the sub-components rated (and the dimensions considered in each); Part B shows grade descriptors by sub-component, so what schools need to achieve to be awarded any given grade. Clearly, this embodies a wealth of soft information that not easily available to parents *ex ante*.

As detailed below, there is dispersion across schools in overall ratings, and also within-school changes in ratings over time between one inspection and the next, that we exploit for our analysis.

3.2 Data

Our analysis is based on household survey panel data linked to administrative data on schools and school inspections. This data linkage is a novel aspect of our study and enables us to examine the impacts of the nationwide school inspection regime. By further linking our schools administrative data to administrative data on individual test scores, we shed light on the nationwide test score impacts of the inspections regime.

3.2.1 UK Household Longitudinal Survey

The UKHLS is a representative panel of around $40,000$ households tracked annually since 2009 [UKHLS 2018]. We use a restricted access version that identifies the school attended by each child in the household. The survey interviews annually all adults in the household aged 16 and over. We exploit three survey waves: 1, 3 and 5 (as these are the ones in which parental help with homework, our main outcome, is collected). The exact interview date is recorded in each wave.[12]

Parental behaviors such as time spent helping children with their homework are recorded, as are parental attitudes such as aspirations for their child attending university. We complement these with data on children's behaviors because the UKHLS contains a separate self-completed questionnaire for children aged 10-15. This records the young person's own time investment into their homework. We use this information to build a picture of how parents and children respond

[11] https://www.gov.uk/government/publications/annual-parents-survey; https://www.gov.uk/government/news/9-out-of-10-parents-know-the-ofsted-rating-of-their-childs-school-or-childcare-provider

[12] Survey wave 1 takes place between January 2009 and December 2010, thus (partly) covers academic years 2008/09, 2009/10, 2010/11; survey wave 3 takes place between January 2011 and December 2012, thus (partly) covers academic years 2010/11, 2011/12, 2012/13; survey wave 5 takes place between January 2013 and December 2014, thus (partly) covers academic years 2012/13, 2013/14, 2014/15.

to new information on school quality.

This key outcome we consider is parental time investment (e). This question is identically worded across waves as, "How often do you help your child/children with his/her/their homework?". Answers are given on a five-point Likert Scale (almost every day, at least once a week, at least once a month, less than once a month, never or hardly ever).[13]

The change in parental time investment, $\Delta e_t = e_t - e_{t-2}$, is measured between waves 1 and 3, and 3 and 5. $\Delta e_t = -1$ if the parent helps less frequently, 0 if equally frequent, and 1 if more frequently. By focusing on within-household changes, we remove cross sectional and time invariant components of school quality driving parental investments. To maintain sample size, we do not restrict children to be in same school across waves (although the majority of children are in the same school). Our working sample consists of children in the UKHLS whose school was inspected in the academic year of their UKHLS interview. As schools are inspected every three to five years, around a quarter of children attend a school inspected in the survey year. Table A2 details sample characteristics as we make each selection towards our working sample of 690 households.[14]

Panel A of Figure 1 shows parental time investment into children's homework, by survey wave. Time allocations across survey waves are relatively stable: almost half of parents report helping their child at least once per week; at the tails, 20-30% report helping almost every day, and 10-20% report never or hardly ever helping. Recall that Result 1 made precise that even within the same school, there will be across-family differences in parental effort e, variation driven by: (i) differences in prior beliefs about school quality, (ii) differences in preferences ($u_{es} \lesseqgtr 0$). Panel B then shows within household changes over time: (i) 18% of parents increase time investments; (ii) 43% keep constant their time investment; (iii) 39% of parents decrease their time investment.

3.2.2 Linked Administrative Schools Data

We link to three school-level administrative data sets: (i) Department for Education school performance tables: these provide longitudinal information on schools' academic performance; (ii) school census data: this provides characteristics of the student body and school type; (iii) Ofsted

[13] This question is asked separately of both parents if they have one or more child aged 10-15. Where responses are available for both parents, we choose the dominant parental helper, defined as the one helping more. When both parents are present, 81% of the time the dominant parent is mother. If there is more than one child aged 10-15, we restrict attention to those households where all children attend the same school. Del Boca *et al.* [2014] present evidence from a dynamic structural model of child development, that suggests maternal and paternal investments are equally productive.

[14] Table A2 shows that from the baseline sample of UKHLS households in England with children aged 10-15 (Column 1), there are few observable differences for households for whom the change in parental investment can be constructed (Column 2), and those that have a school-code needed to link to the administrative data (Column 3). The selection margin that reduces the sample is the need for the child's school to have been inspected in either survey wave 3 or 5: given schools are inspected every four years around a quarter of households also have their school inspected in the UKHLS data. Our working sample has similar characteristics of the household, mother and father to the earlier samples shown.

inspections data: this provides inspection outcomes and the exact date of inspection.[15]

The school performance tables cover academic years 2009/10 to 2013/14 (corresponding to survey waves 1 to 5) and provide hard information readily available to parents online. We access school census data for academic years 2008/09 to 2013/14.[16]

Ofsted data covers all inspections from September 2005 until December 2014: this provides exact dates for 63,337 inspections conducted in 23,778 schools. We are thus able to construct the trajectory of inspection ratings for a school over time, including from before parental inputs are measured in the UKHLS. Characteristics of inspected versus non-inspected schools in waves 3 and 5 are shown in Table A3. As expected, inspected schools are worse performing than non-inspected schools (as failing schools are subject to more regular inspection), but these differences are not large. To reiterate, our research design does *not* exploit across-school variation between inspected and non-inspected schools.

Panel C of Figure 1 shows inspection ratings by survey wave. The distribution of ratings is relatively stable over time: around 16% of schools receive an outstanding rating, 48% receive a good rating, 30% receive a rating of requires improvement, and 7% of schools are rated as failing/inadequate. Panel D shows within-school rating changes. The majority of schools change rating: (i) 28% of schools experience an improved rating, 46% of schools have an unchanged rating, and 26% of schools have a worse rating.

Finally, we link these schools data sets to administrative data on individual child test scores from the National Pupil Database (NPD). We use this to examine the test score impacts of news generated by the nationwide inspections regime.[17]

[15]The school identifier is collected in waves 1, 3, 5. To fill in missing identifiers, we use a two-step approach. First, households were also asked to provide consent to link their children's data to test score records in the National Pupil Database (NPD). The consent rate was 68%, and any consent bias should not impact our results as long as it is orthogonal to the selection into treatment and control groups based on the timing of the UKHLS interview relative to inspection date. Households are balanced on observables for those whose school identifier was in the UKHLS data and those for whom it was obtained through the UKHLS-NPD linkage. We further infer the school in wave t if the school in the preceding and subsequent waves is the same.

[16]The following school test score indicators are available: the percentage of pupils with 5 or more A*-C grades, the percentage with 5 or more A*-C grades including English and Maths, the percentage with 5 or more A*-G grades, the total average point score, the percentage of pupils making expected progress in English, and in Maths, and the percentage of English Baccalaureates. The schools census data contains information on school size (number of pupils), the percentage of pupils entitled to free school meals, the percentage of pupils for whom English is an additional language at home, the school type (academy, community, voluntary aided, controlled, foundation), whether it has a sixth form, any Christian or other religious denomination, and whether it is a mixed gender school.

[17]The NPD contains information on students attending schools and colleges in England. It combines high stakes and nationwide examination results with information on pupil and school characteristics.

4 Empirical Method

4.1 Defining Treatment and Control Households

School inspections take place in throughout the academic year (September to July). UKHLS interviews take place in all months. Hence in our linked household-school data, we observe some households being interviewed pre school inspection (the control group), and some being interviewed post inspection (the treated group).

Panel A of Figure 2 shows the timing of UKHLS interviews: these occur evenly over months. Panel B shows the timing of inspections. These are slightly shifted towards the first term of the academic year (September to December). Fewer inspections occur during holidays (December, April) or at the end of the academic year (July). These timing patterns of household interviews and school inspections ameliorate concerns over UKHLS enumerators or inspectors front/back-loading their effort, that could otherwise have led to measurement error in parental behavior or inspection ratings being correlated with month.

Our analysis is based on schools that are inspected at some point during the academic year (we never exploit differences between inspected and non-inspected schools). Exploiting the panel structure of the data, our outcome is the change in parental inputs for household i in school σ between period t and $t-2$, $\Delta Y_{i\sigma t}$. The treatment effect we capture is the difference in parental inputs over time between: (i) control households, whose children are in schools that will be inspected in year t but are surveyed prior to the inspection and school quality information being released; (ii) treated households, whose children are also in schools that will be inspected in year t but are interviewed after the inspection and school quality information has been released. Panel C shows month of interview for treated and control households. As expected, treated households are more likely to be interviewed in the UKHLS from March to August. 42% of sample households are controls; 58% are treated.

Treatment assignment is determined by the date at which households are surveyed in the UKHLS relative to the date of school inspection. Below we make precise the identifying assumptions our design requires, and provide a battery of evidence in support of them.

4.2 Measuring News

Parents should only respond to inspection ratings if there is new information, 'news', embodied in them. To construct prior beliefs we use a simple model to forecast a school's inspection rating based on publicly available information, including the school's previous inspection rating and test score results. As Ofsted inspectors attach some weight to prior test scores, there will be a predictable component to inspection ratings. We define news for school σ in time period t as,

$$news_{\sigma t} = \text{rating}_{\sigma t} - \text{predicted rating}_{\sigma t}. \tag{5}$$

If parents have access to additional information not observed by the econometrician (e.g. information from teachers, peers, or children), they will better predict the actual rating than our model, and $news_{\sigma t}$ overstates the information provided. On the other hand, if parents are unable to distinguish noise from the signal in volatile short-term test score movements, $news_{\sigma t}$ will understate the information provided by inspection ratings [Kane and Staiger 2002]. This kind of measurement error might be stronger in smaller schools, that we can check for (although in our context, children are aged 10-15 and are mostly in secondary schools with over 1000 pupils).[18]

To maximize the precision of the forecast, we use the sample of all secondary schools inspected during academic years overlapping with survey waves 1, 3 and 5 of the UKHLS. The sample covers 4,419 inspections conducted in 3,113 schools. As we have inspections data back to 2005, nearly all schools have a prior rating ($rating_{\sigma t-1}$). We use the following AR(1) specification for the forecasting model,

$$rating_{\sigma t} = \beta_0 rating_{\sigma t-1} + \beta_1 Z_{\sigma t} + \lambda_l + u_{\sigma t}, \tag{6}$$

where $rating_{\sigma t-1}$ is the previous Ofsted rating, $Z_{\sigma t}$ are school performance and school characteristics and λ_l are local education authority fixed effects.

Table A4 shows the results. Columns 1 to 6 estimate (6) by gradually expanding the set of covariates $Z_{\sigma t}$, capturing larger information sets being used by parents to predict inspection ratings. Column 1 estimates the naïve model, so ratings today only depend on the previous rating, while by Column 6 we additionally control for school characteristics and a rich set of past school performance indicators. Across specifications we find a persistence in ratings across inspection cycles: $\widehat{\beta}_0 \in [.268, .439]$ and is always highly significant. Throughout, $\widehat{\beta}_0 < 1$ so we reject a random walk forecasting model, and past exam performance is always statistically significant and economically meaningful in predicting inspection outcomes (consistent with the inspection framework giving weight to academic performance). These features of the forecasting model continue to hold when: (i) we restrict the sample to the same schools as in our analysis of parental inputs (Column 7); (ii) we estimate (6) using an ordered probit model (Column 8).

Most of the variation in ratings within local education authorities is unexplained: across OLS specifications, the $\bar{R}$-squared remains below .42. This suggests ratings are driven by soft information, not captured in publicly available schools data. This is in line with Ofsted inspections criteria that emphasizes the major weighting given to qualitative features of the school, reinforcing the notion that ratings can provide news to parents.

[18] The forecast is constructed at the school level: the UKHLS has no information at the household level of expected inspection ratings. Abdulkadiroglu *et al.* [2017] and Beuermann *et al.* [2018] overview recent work examining whether parents can tell what constitutes a good school.

We use the model prediction to define $news_{\sigma t} = f(rating_{\sigma t}, \widehat{rating_{\sigma t}})$ as follows:

$$news_{\sigma t} = \begin{cases} good_{\sigma t} \text{ if (actual rating} - rnd(\widehat{rating_{\sigma t}}) > 0 \\ none_{\sigma t} \text{ if (actual rating} - rnd(\widehat{rating_{\sigma t}}) = 0 \\ bad_{\sigma t} \text{ if (actual rating} - rnd(\widehat{rating_{\sigma t}}) < 0 \end{cases}, \quad (7)$$

where we use predicted ratings (rounded to the nearest integer). The (actual rating$-rnd(\widehat{rating_{\sigma t}})$) ranges from -2 to $+2$, with 18% if schools receiving bad news, 25% of schools receive good news, and 57% of schools receive no news (because (actual rating$-rnd(\widehat{rating_{\sigma t}})$)$= 0$).[19]

In relation to Results 2A and 2B, $news_{\sigma t}$ corresponds to dp_H, the change in belief that the school is of high quality. We have considerable variation in $news_{\sigma t}$ to identify parental preferences. We observe good and bad news being revealed to schools that had the highest rating (outstanding) in the previous cycle, and the same for schools that start with the lowest inspection rating. This is because over inspection cycles, a lot of hard information on school quality is revealed to parents, so that schools previously at the tails of the rating distribution can still be shocked up and down. In Section 6 we exploit this full variation in news across the schools to shed light on distributional impacts for parental inputs across schools.

4.3 Research Design

Figure 3 shows our research design, combining all the elements above. Treatment-control comparisons can be made across schools in which: (i) good news is received (top panel) so the key difference-in-difference (DD) is $\mathbf{E}[\Delta Y_{i\sigma t} - \Delta Y_{j\sigma' t} \mid good_{\sigma t}, good_{\sigma' t}]$, mapping to Result 2A; (ii) bad news is received (lower panel) so the key DD is $\mathbf{E}[\Delta Y_{i\sigma t} - \Delta Y_{j\sigma' t} \mid bad_{\sigma t}, bad_{\sigma' t}]$, mapping to Result 2B (there is a corresponding DD estimated for parents receiving no news that for expositional clarity we do not show on Figure 3). To reiterate, for each DD, we hold constant the news that will be received, and only exploit as good as random variation between treated and control households (those that have and do not have the same news on school quality).

We implement our research design by estimating the following specification:

$$\Delta Y_{i\sigma t} = \delta_0 + \beta_0 T_{i\sigma t} + \beta_1 [T_{i\sigma t}.good_{\sigma t}] + \beta_2 [T_{i\sigma t}.bad_{\sigma t}] + \delta_1 good_{\sigma t} + \delta_2 bad_{\sigma t} + \gamma_1 X_{i\sigma t} + \gamma_1 Z_{\sigma t} + \varepsilon_{i\sigma t}, \quad (8)$$

where $\Delta Y_{i\sigma t}$ is change in help with homework by parents i in school σ between t and $t-2$. $T_{i\sigma t}$ is a dummy equal to one for treated households, so those interviewed after an Ofsted inspection and zero otherwise. $good_{\sigma t}$, $bad_{\sigma t}$ are the news shocks received by households in school σ in year

[19] This OLS prediction model is subject to misclassification bias. This concern is partly ameliorated by rounding the predicted rating. In addition, we also present a robustness checks where: (i) the prediction model is based on the ordered probit from Column 8 in Table A4; (ii) we limit the sample where we only use schools rated good and satisfactory (so with a rating of two or three).

t. $X_{i\sigma t}$ are child- and family-level controls, and $Z_{\sigma t}$ are school-level controls.[20]

As $\Delta Y_{i\sigma t} \in \{-1, 0, 1\}$ we estimate (8) using an ordered probit model.[21] Finally, because $good_{\sigma t}$ and $bad_{\sigma t}$ are generated regressors we use bootstrap methods to derive standard errors, allowing them to be clustered at the local education authority level.[22]

By examining the change in parental input over time, $\Delta Y_{i\sigma t}$, we remove time invariant household and school factors driving parental inputs $(\alpha_i, \alpha_\sigma)$. This is important because the UKHLS does not contain multiple observations of children from different households in the same school, so we cannot condition on school fixed effects. Rather it provides a representative sample of children across schools, allowing us to evaluate the nationwide consequences of the inspections regime. Hence treatment and control children do not attend the same school, rather school fixed effects are differenced out in our design. We compare within the set of schools inspected in year t and condition on school characteristics and the actual news from the inspection rating. There will also be time trends in investment within the academic year, e.g. parents might help their child more closer to exams. However, these kinds of changes in parental input during the academic year are differenced out because households are surveyed in the same month each survey wave and exams take place in the same month each academic year.

As we condition on $news_{\sigma t}$, the central difference between treated and control households is that the former are aware of the actual inspection rating, while control households are not. Given the immediate and widespread dissemination of Ofsted ratings, there is near perfect compliance among treated households.

4.4 Identifying Assumptions

For a causal impact of news about school quality to be identified from the comparison of treated to control households, four assumptions are required: (i) no selection of schools by month of inspection; (ii) no selection of households by month of interview; (iii) no time trends in $\Delta Y_{i\sigma t}$; (iv) no within school-year responses to inspections by schools.

[20] $X_{i\sigma t}$ controls are as follows. The child-level controls are gender and age dummies; the family-level controls are household size, number of children in household, housing tenure (owned, rented, missing information), mother's ethnicity (white, non-white, missing), mother's highest education (5 binary indicators), and mother's marital status (married/cohabiting, single, missing). Where mother's information is missing, father's information is used. The school level controls $Z_{\sigma t}$ are school size and the proportion of children eligible for free school meals.

[21] In this model we define a latent variable $\Delta Y^*_{i\sigma t}$ for family i such that:

$$\Delta Y^*_{i\sigma t} = \beta_0 T_{i\sigma t} + \beta_1 [T_{i\sigma t}.pos_{\sigma t}] + \beta_2 [T_{i\sigma}.neg_{\sigma t}] + \delta_1 pos_{\sigma t} + \delta_2 neg_{\sigma t} + \gamma_1 X_{i\sigma t} + \gamma_1 Z_{\sigma t} + \varepsilon_{i\sigma t} = \boldsymbol{\beta}' \mathbf{X}_{i\sigma t} + \varepsilon_{i\sigma t},$$

where $\Delta Y_{i\sigma t} = -1$ if $-\infty < \Delta Y^*_{i\sigma t} \leq \mu_0$, $\Delta Y_{i\sigma t} = 0$ if $\mu_0 < \Delta Y^*_{i\sigma t} \leq \mu_1$, and $\Delta Y_{i\sigma t} = 1$ if $\mu_1 < \Delta Y^*_{i\sigma t} < \infty$.Assuming $\varepsilon_{i\sigma t} \sim N(0,1)$ yields the ordered probit model where, $\text{Prob}(\Delta Y_{i\sigma t} = j) = \phi(\mu_j - \boldsymbol{\beta}' \mathbf{X}_{i\sigma t}) - \phi(\mu_{j-1} - \boldsymbol{\beta}' \mathbf{X}_{i\sigma t})$.

[22] There are analytic correction methods to correct for generated regressors such as the two-step variance estimator derived in Murphy and Topel [1985], but we cannot use this because the first and second stages are at different units of observation (school and then child). Hence we proceed as follows: we append the school-level data used for the first stage and individual level data used for the second stage. We let the bootstrap program draw random samples across both data sets in each iteration. We then derive standard errors using bootstrap with 1000 replications.

We examine assumption (i) using two strategies. First, Panel A of Table 1 shows school characteristics by treatment and control (Figure A2 shows p-values on balance tests and normalized differences for a wider set of school outcomes). For the vast majority of characteristics there are no significant differences between groups. Second, we regress ratings on month of inspection. Table 2 shows the results: there is no statistically significant impact of month of inspection on rating once previous rating is controlled for. No month dummy is significant, and the joint F-test on month of inspection dummies does not reject the null ($p = .567$ in our preferred specification in Column 4). Linking back to Panel B of Figure 2 on inspection timing, we note that the September to December month dummies in Table 2 are also not jointly significant ($p = .945$). All results are robust to using an ordered probit model (Column 5). The results help rule out that Ofsted inspectors aim to reach 'rating targets' by the end of an academic year.

On assumption (ii), Panels B to E of Table 1 show balance between treatment and controls on characteristics of the household, child, mother and father (Figure A2 shows p-values on balance tests and normalized differences over a wider set of outcomes). We find no imbalances. Given that heterogenous treatment effects are central in our research design, we further show balance by type of news shock. Table A5 shows that even *conditional* on households receiving good news, no news, or bad news, there remains a high degree of balance between treatment and controls.

We present three strategies to underpin assumption (iii) of no time trends in $\Delta Y_{i\sigma t}$. First, UKHLS households are interviewed in the same month across surveys. Figure A3 shows the cumulative distribution in the absolute difference in interview date across waves. More than 75% of households at wave t are interviewed within 30 days of the date in wave $t-2$. Second, we later present a robustness check where we condition on month of interview (recall that Figure 2 showed variation in when treatment and controls are interviewed). Third, we construct a placebo check, taking schools to be inspected in year $t+1$ (so a year after survey waves 3 and 5) and assign next year's inspection date in the current year. This placebo check is presented in the Appendix.

On assumption (iv), that there are no within-year school responses to ratings, note that in English schools hiring decisions over teachers/assistants are made at the end of the academic year. However, schools might adjust on other margins in the short run. No data on fine-grained adjustments in secondary schools exists for England. However, to shed light on the issue we use the Millennium Cohort Study (MCS), a panel of children tracked since birth in 2000/1, that can be linked to a detailed survey of their teachers. We link the MCS and schools administrative data using school identifiers to examine fine-grained responses in school practices and organization, to good and bad news among schools inspected in academic years 2007/08 and 2008/09 (to just overlap with our main UKHLS sample). This analysis is in the Appendix, and documents little change in short run practices across a range of teaching practices, including homework set, the use of teaching assistants or supply teachers, time spent on numeracy/literacy, and ability grouping.[23]

[23]Hussain [2015] provides evidence of short run adjustment by schools labelled as failing by inspectors: they lengthen time devoted to instruction, change their instructional policies and practices and as a result, test scores

5 Results

Figure 4 presents evidence on how the raw unconditional $\Delta Y_{i\sigma t}$, the change in help with homework by parents i in school σ between t and $t-2$, varies with news. For each realization of $news_{\sigma t} \in \{good_{\sigma t}, none_{\sigma t}, bad_{\sigma t}\}$, we graph the raw difference between treated and control households in the proportion of parents with $\Delta Y_{i\sigma t} = -1, 0, 1$. The first set of bars show that conditional on good news, parents are much more likely to decrease help. This suggests beliefs about school quality and parental inputs are substitutes. The second set of bars reveals that there is little change in parental time investment when no news is revealed by the school inspection: $\Delta Y_{i\sigma t}$ is similar between treated and control households. The last set of bars show that in response to bad news, there are more heterogeneous parental responses, with many parents leaving inputs unchanged.

The last panel shows the net impact of receiving a positive rather than a negative news shock, corresponding to the DDD: $\mathbf{E}[\Delta Y_{i\sigma t} - \Delta Y_{j\sigma' t} \mid good_{\sigma t}, good_{\sigma' t}] - \mathbf{E}[\Delta Y_{i\sigma t} - \Delta Y_{j\sigma' t} \mid bad_{\sigma t},\ bad_{\sigma' t}]$. Parents are far more likely to decrease time investment in response to good news.

Overall the raw unconditional evidence suggests e and s are substitutes: parents engage in offsetting behavior with respect to their beliefs about school quality.

5.1 Regression Results

Table 3 presents our core results that estimate (8). The coefficients of interest are the DD estimate for good news, $\widehat{\beta}_0 + \widehat{\beta}_1$, and the DD estimate for bad news: $\widehat{\beta}_0 + \widehat{\beta}_2$. For completeness, we also show the DDD estimate $\widehat{\beta}_1 - \widehat{\beta}_2$, but unlike our core DD estimates, this exploits more than the variation induced by treatment assignment, also using differences in the news received. Hence this identifies a causal impact under stronger assumptions than those discussed earlier, and so this just forms a secondary focus of analysis.

Across Columns in Table3 we sequentially add in covariates ($X_{i\sigma t}$, $Z_{\sigma t}$). The estimates are stable across specifications suggesting there is not a high correlation between these child, parent, household and school characteristics and how parents respond to information on school quality.

The results show that when parents receive goods news about school quality, they are significantly less likely to increase time investment ($\widehat{\beta}_0 + \widehat{\beta}_1 < 0$). In contrast, when parents receive bad news about school quality, their time investment into their child does not change ($\widehat{\beta}_0 + \widehat{\beta}_2 = 0$). This implies that for the average family, there is substitutability between beliefs about school quality and parental time investment, so $u_{es} < 0$.[24]

improve. Recall that in our sample only 7% of schools are ranked as failing, and our core results are robust to dropping them.

[24] Table A6 shows all coefficients from (8). $\widehat{\beta}_0 = 0$ across specifications, so there is no effect on parental investment of being in a treated household that receives no informative signal from the school inspection ($news_{\sigma t} = none_{\sigma t}$). The full specification also shows $\widehat{\delta}_1 = \widehat{\delta}_2 = 0$, so that being in a school that receives good or bad news at some point in the academic year does not itself correlate to changes in parental investment. This further underpins identifying assumption (iii), that there are no time trends in $\Delta Y_{i\sigma t}$.

To help quantify impacts, we report averaged marginal effect estimates from our preferred specification in Column 4 of Table 3. The marginal effects measure how being treated with a given news shock changes the likelihood that parental investments increase, decrease, or stay the same. Figure 5 shows the same evidence graphically. For treated households receiving *good* news about school quality from Ofsted inspections, the probability their time investment: (i) increases, falls by 11pp; (ii) remains unchanged, falls by 8pp; (iii) decreases, rises by 20pp. For treated households that receive *bad* news about school quality, there are more muted responses in time investments, but the marginal effects are always of opposite sign to the reaction to good news.

The difference between responses to good and bad news are significantly different, as shown by the third row ($\widehat{\beta}_1 - \widehat{\beta}_2 < 0$). From Result 3 in the conceptual framework, as parents respond more to good news than to bad news suggests $p_H < 0.5$ (so $\left|de|_{j=H}\right| > \left|de|_{j=L}\right|$): the average family holds the prior their school is more likely to be low quality. This suggests parents hold pessimistic beliefs over school quality in an education market where parents typically make an explicit choice of school based on expected quality [Burgess *et al.* 2015, Agarwal and Somaini 2018, Beuermann *et al.* 2018].[25]

In the Appendix we present a battery of checks on the core result, showing it to be robust to alternative samples, controls and estimation methods, and examining the possibility schools with bad news strategically delay the release of such information (Tables A7 to A9).

For completeness we probe the data to examine heterogeneous responses to news. This is subject to the obvious caveat that given our sample size, we are not well powered so these results are merely suggestive. Heterogeneous responses to news can be driven by households having different prior beliefs, having different forecast models, or the cross derivative between beliefs about school quality and parental effort differing in their utility functions. These results are summarized in Figures A5A and A5B, that show marginal effect estimates from the ordered probit model. We find the differential response to good and bad news is driven by higher educated households, non-white households, those where the child is of higher birth order, for boys, and among children that are below median ability (as measured in administrative test score data).[26]

[25] In our sample, very few parents are observed responding to news from school inspections by changing the school their child attends. This is unsurprising given the large fixed costs of changing school for children aged 10-15.

[26] A common finding in the school accountability literature is that low-income families respond less to hard information on test scores – that might be because they place less weight on academic gains as they expect lower returns to education [Hastings and Weinsten 2008], or because it is more costly for them to act on their preferences. Del Boca *et al.* [2014] present evidence from a dynamic structural model of child development suggesting ambiguous impacts of household income on child development. The reason is that higher income often means greater labor supply and reductions in time investments into children. In line with our results, Autor *et al.* [2016] use data on opposite-sex siblings attending Florida public schools to document how boys benefit more from cumulative exposure to higher quality schools.

5.2 Distributional Impacts

Our results show that parents respond to news about school quality. A key consequence is that the inspection regime has distributional consequences depending on how news is allocated across schools. Our context and data provide an almost unique opportunity to understand such distributional consequences of a nationwide inspections regime. We proceed in three steps.

First, we establish how news correlates to the initial level of school ratings s by documenting prob($news_{\sigma t}|s = s_k$) for each value of $news_{\sigma t}$ and school quality k. Columns 1 to 3 of Table 5 show these descriptives, and then the equivalent regression adjusted probabilities in italicized braces. Column 1 shows that schools previously rated as outstanding are more likely than other schools to receive a positive news shock: prob($news_{\sigma t} = good_{\sigma t}|s = s_{\text{outstanding}}$)= .33. We note that prob($news_{\sigma t} = good_{\sigma t}|s = s_k$) is monotonically decreasing in initial school quality, s_k. This suggests some high quality schools that might have been expected to deteriorate given their observables since the last inspection do not, hence the good news to parents.

On prob($news_{\sigma t} = bad_{\sigma t}|s = s_k$), Column 3 shows that schools previously rated good or requiring improvement are most likely to receive bad news (with probabilities around .19), closely followed by the probability that an outstanding school receives bad news (.16). Schools previously rated inadequate are least likely to receive bad news (.05). This is in line with Hussain [2015] who finds that such schools respond to poor ratings and subsequently improve their performance.

Second, we examine how this translates into changes in the level and inequality of parental inputs. We first define a high level of parental input ($Y = 1$) if the frequency of help with children's homework is almost every day, or at least once a week (the top two frequencies from the Likert score), and $Y = 0$ otherwise. Column 4 then shows regression adjusted pre-treatment (period $t - 2$) levels of parental input. We find a slightly positive gradient of parental input with regard to school ratings: in outstanding schools, 85% of parents provide high levels of input into their children, and this falls to 73% in the lowest rated schools.

Third, we combine these probabilities with the regression estimates to calibrate implied impacts on: (i) expected parental inputs, $E[Y]$; (ii) pre-inspection between-school inequality in parental inputs between high and low quality schools (s_L, s_H), denoted Q; (iii) post-inspection treatment effects of school ratings information on $E[Y]$ and Q. We use the range as our measure of between-school inequality in parental inputs, Q. The expected parental input and inequality across schools are given by:

$$E[Y] = \sum\nolimits_k prob(s = s_k)E[Y|s = s_k], \tag{9}$$

$$Q = E[Y|s = s_H] - E[Y|s = s_L]. \tag{10}$$

The treatment (T) effects on these are as follows:

$$\frac{\partial E[Y]}{\partial T} = \sum\nolimits_k prob(s = s_k) \frac{\partial E[Y|s = s_k]}{\partial T} \tag{11}$$

$$= \sum\nolimits_k prob(s = s_k) \left\{ \sum\nolimits_j prob(j|s = s_k) \frac{\partial E[Y|s = s_k, j]}{\partial T} \right\}, \tag{12}$$

$$\frac{\partial Q}{\partial T} = \frac{\partial E[Y|s = s_H]}{\partial T} - \frac{\partial E[Y|s = s_L]}{\partial T} \tag{13}$$

$$= \sum\nolimits_j prob(j|s = s_H) \frac{\partial E[Y|j]}{\partial T} - \sum\nolimits_j prob(j|s = s_L) \frac{\partial E[Y|j]}{\partial T}, \tag{14}$$

where school quality $k = \{$outstanding, good, requires improvement, inadequate$\}$ and signal $j = \{$good, zero, bad$\}$.[27] Each element in (11) can be substituted in for either from the marginal effects in Table 4 ($\frac{\partial E[Y|s=s_k]}{\partial T}$) or the evidence in Table 5 ($prob(s = s_k)$, $prob(j|s = s_k)$).

Doing so yields the following calibration: pre-inspection, the expected parental input is $E[Y] = .81$ and given the distribution of news shocks across schools, this falls overall by 14% because of the information revealed by the school inspection regime ($\frac{\partial E[Y]}{\partial T} = -.11$). Given that parental inputs and beliefs about school quality are substitutes, the mechanism for this is that parents with good news reduce parental inputs by more than parents in schools that receive bad news.

Figure 6 summarizes the findings. The first four sets of bars show for schools of pre-inspection quality s_k: (i) the share of schools of this type; (ii) the unconditional treatment effect on parental inputs; (iii) the conditional treatment effect on parental inputs. For each type of school, parental inputs fall overall in response to the inspections regime (a result robust to conditioning on covariates). Aggregate parental inputs then fall (as shown in the fifth set of bars on $\frac{\partial E[Y]}{\partial T}$), and because inputs fall more in higher ranked schools than lower ranked schools, input inequality also falls with the inspection regime. Just comparing differences between outstanding and inadequate schools we find that $Q = .13$ and $\frac{\partial Q}{\partial T} = -.02$ so across-school inequality in parental inputs falls by 15% because of the information generated by the inspections regime.

6 Test Score Impact

We now build on our core results to understand the test score impacts of news on school quality. However, parental time is not the only margin of response to such news. The UKHLS data allows us to study a wide range of behavior, and so construct a holistic picture of how multiple margins of response to news combine to impact test scores for children.

The estimates for other margins of response are summarized in Figure 7, that shows marginal

[27] Our research design assumed the response to news was homogenous across schools, so $\frac{\partial E[Y|s=s_k,j]}{\partial T} = \frac{\partial E[Y|j]}{\partial T}$. Figure A6 explores this assumption by showing how the marginal impacts of news vary between: (i) schools with an earlier rating of outstanding/good; (ii) schools with an earlier rating of requires improvement/inadequate. These are found to be similar and so we maintain the assumption going forward.

impacts on $\Delta Y_{i\sigma t}$ in each case. Panel A considers the change in whether the child talks to their parent about important matters most days. We see these changes to mirror the time inputs of parents into children: in response to good news children are significantly less likely to talk to parents about important matters on most days. This highlights that parents do not seem to substitute one form of input into their child (time spent on homework) with another (time talking about important matters): rather *both* parent-child interactions are substitutes to beliefs over school quality.

Panel B shows changes in the amount of time children themselves report spending on their homework.[28] Children's time inputs move in the opposite direction to the behavioral response of parents: when a household receives good news about school quality, the child is significantly more likely to increase time spent on homework. In other words, children partly compensate for the loss of parental inputs by increasing their own time spent on homework, so their effort is complementary to beliefs about school quality. Children's responses to no news or bad news are imprecisely estimated (as was also the case for parental responses to such news).[29]

6.1 Framework

To understand how multiple response margins combine to impact test scores, we modify the Todd and Wolpin [2003] framework to our context. Period $t = 0$ signifies the school year prior to inspection; $t = 1$ and $t = 2$ signify the start and end of the academic year in which the inspection takes place. Achievement A_t is given by period specific test score production functions, $g_t(.)$:

$$\begin{aligned} A_1 &= g_0(F_0), \\ A_2 &= g_1(F_0, F_1, \Sigma_1), \end{aligned} \tag{15}$$

where F_t and Σ_t are family- and school-based inputs in period t. The household is assumed to target a level of school input $\bar{\Sigma}_1 = \theta(A_1)$ (say through an initial school choice) and the school in turn decides its supply of school-based inputs: $\Sigma_1 = \psi(A_1)$. $\Sigma_1 - \bar{\Sigma}_1$ is the deviation in school inputs received and targeted. On family inputs, we assume parents choose inputs after observing school inputs (matching our research design), so $F_1 = \varphi(A_1, \Sigma_1 - \bar{\Sigma}_1)$. *Ceterus paribus*, the

[28] The change in hours the child spends doing homework is derived from the response to following question in waves 2 and 4, "When you do homework on a week-day evening during term time, how many hours do you usually spend doing your homework?" ($N = 244$). We convert ΔY_{ist} into an increase, decrease or no change in time between waves.

[29] De Fraja *et al.* [2010] is among the few papers that also study the interplay between parental and child effort. Proxying child effort by their attitudes and parental effort by their interest in their child's education, they find the two margins to be complements.

achievement impact of a change in school input is:

$$\frac{\partial A_2}{\partial \Sigma_1} = \frac{\partial A_2}{\partial \left(\Sigma_1 - \bar{\Sigma}_1\right)} = \frac{\partial g_1}{\partial \Sigma_1}. \tag{16}$$

which depends on the production function, $g_1(.)$. However, given endogenous family responses, as Todd and Wolpin [2003] emphasize, total impacts are actually observed:

$$\frac{dA_2}{d\Sigma_1} = \frac{dA_2}{d\left(\Sigma_1 - \bar{\Sigma}_1\right)} = \frac{\partial}{\partial \Sigma_1} g_1(\Sigma_1, \varphi(A_1, \Sigma_1 - \bar{\Sigma}_1), F_0, \mu) = \frac{\partial g_1}{\partial \Sigma_1} + \frac{\partial g_1}{\partial F_1} \frac{\partial F_1}{\partial \left(\Sigma_1 - \bar{\Sigma}_1\right)}, \tag{17}$$

where the second term is the indirect behavioral response of families to school inputs. In our setting, there are multiple family inputs of parents and children, denoted F_1^p, F_1^c respectively, where we assume child inputs respond to parental inputs as well as to school inputs. Following Pop-Eleches and Urquiola [2013] the total policy effect then is:

$$\frac{dA_2}{d\Sigma_1} = \frac{dA_2}{d\left(\Sigma_1 - \bar{\Sigma}_1\right)} = \frac{\partial g_1}{\partial \Sigma_1} + \frac{\partial g_1}{\partial F_1^p} \frac{\partial F_1^p}{\partial \left(\Sigma_1 - \bar{\Sigma}_1\right)} + \frac{\partial g_1}{\partial F_1^c} \left[\frac{\partial F_1^c}{\partial \Sigma_1} + \frac{\partial F_1^c}{\partial F_1^p}\right], \tag{18}$$

where $\frac{\partial g_1}{\partial \Sigma_1}$ is the direct effect of school inputs on achievement, $\frac{\partial g_1}{\partial F_1^p} \frac{\partial F_1^p}{\partial \left(\Sigma_1 - \bar{\Sigma}_1\right)}$ is the indirect effect of parental inputs, and the final term $\frac{\partial g_1}{\partial F_1^c} \left[\frac{\partial F_1^c}{\partial \Sigma_1} + \frac{\partial F_1^c}{\partial F_1^p}\right]$ is the response of child-inputs to school and parental inputs. Setting $\frac{\partial g_1}{\partial \Sigma_1} = 0$ (so there are no short school-based responses to Ofsted as shown in Table A8), then finding an impact on post-treatment test scores ($\frac{dA_2}{d\Sigma_1} \gtreqless 0$) implies the relative total products of family to child inputs in generating test scores, can be assessed:

$$\frac{\partial g_1}{\partial F_1^p} \frac{\partial F_1^p}{\partial \left(\Sigma_1 - \bar{\Sigma}_1\right)} \gtreqless -\frac{\partial g_1}{\partial F_1^c} \left[\frac{\partial F_1^c}{\partial \Sigma_1} + \frac{\partial F_1^c}{\partial F_1^p}\right]. \tag{19}$$

In our context this is especially relevant because as documented above, parental and child time investments move in *opposite* directions in response to news about school quality.

6.1.1 Results

To study test score impacts of the school inspection regime we link the schools data with administrative data on children's test scores from the National Pupil Database (NPD), that records high stakes nationwide exam scores, taken at ages 11 and 16. We focus on students enrolled in schools inspected in the 2011, 2012 or 2013 academic years, and who were taking high stakes GCSE exams at age 16 at the end of these academic years. We thus estimate the within-academic year impact

on test scores following information received from Ofsted inspection.[30]

We estimate a value added model for test scores:

$$y_{i\sigma t} = \rho y_{it-1} + \beta_0 T_{\sigma t} + \beta_1 [T_{\sigma t}.good_{\sigma t}] + \beta_2 [T_{\sigma t}.bad_{\sigma t}] + \delta_1 good_{\sigma t} + \delta_2 bad_{\sigma t} + \mu_t + \gamma_0 X_{i\sigma t} + \gamma_1 Z_{\sigma t} + \varepsilon_{\sigma t}, \tag{20}$$

where $y_{i\sigma t}$ is student i's standardized average point score on the age-16 GCSE exams at the end of the academic year, y_{it-1} is her lagged test score at age-11. Treatment assignment is now determined at the school level: $T_{\sigma st}$ equals one if the school inspection took place early in the academic year (September through December), and is zero if the inspection takes place later in the year (January through April). $good_{\sigma t}$ and $bad_{\sigma t}$ are as previously defined. μ_t is a academic year fixed effect, $X_{i\sigma t}$ and $Z_{\sigma t}$ are student- and school-level controls derived from the NPD. We account for the generated regressors from the forecast model by deriving standard errors using the bootstrap method with $1,000$ iterations, allowing them to be clustered by local authority.[31]

The results are in Table 6, where we show the DD coefficients of interest (β_1, β_2) by academic year of inspection, and 90% confidence intervals on each estimate. As with our earlier research design, these identify causal impacts of news if treatment assignment $(T_{\sigma t})$ is as good as random.

Column 1 shows for students in schools inspected in the academic year when they are taking exams, good news about school quality earlier in the academic year leads to significantly *lower* test scores in these high stakes exams $(\widehat{\theta}_1 = .091sd)$. The 90% confidence interval rules out any impact larger than $-.018sd$. Bad news has no significant impact on test scores.

Recall the earlier findings suggested that good news causes parents to reduce their time input and children to increase their time input. Using the framework above and condition (19), the negative net impact on test scores $(\frac{dA_2}{d\Sigma_1} < 0)$ suggests the total product of children's own time investment is less than the total product of parental time investment in producing test scores. These results add to a handful of recent papers that have been able to measure the relative productivity of parental and child investments [Del Boca *et al.* 2017, Caetano *et al.* 2019].[32]

Using school level outcomes in the standard deviation and interquartile range of test scores $(y_{\sigma t})$, Columns 2 and 3 highlight the provision of news over school quality does not impact within-school inequality in test scores.[33] Given the earlier results on the distributional impacts on parental

[30] We drop schools inspected from May onwards in any academic year as this coincides with when GCSE exams are in progress. We also drop students in schools that received a failing inspection rating because such schools are known to be targeted for improvement [Hussain 2015].

[31] The student controls $X_{i\sigma t}$ comprise: eligibility for free lunch, ethnic minority status, special education needs status and gender. School-level controls $Z_{\sigma t}$ comprise: the type of school (e.g. community, academy, voluntary aided), the school's religious status, admission policy, single-sex entry, percentage of students eligible for free school meals, percentage of students speaking English as an additional language and total enrolment.

[32] Del Boca *et al.* [2017] and Caetano *et al.* [2019] examine child and parental inputs into test scores, using data on actual hours of investment to establish the relative marginal products of each. Del Boca *et al.* [2017] find that child time investments are more productive than maternal time investments; Caetano *et al.* [2019] find them to be equally productive (with grandparents active time investment being the most productive input).

[33] This is in slight contrast to results found for US accountability regimes, that have been documented to impact more positively test scores of low-achieving or marginal children [Feng *et al.* 2010, Neal and Schanzenbach 2010].

inputs of the inspections regime, this suggests that among those students whose schools are inspected early in the academic year of their high stakes exams, the inspection regime decreases educational attainment overall and decreases inequality in test scores between high and low quality schools (but not within a school).

7 Conclusion

While there is a voluminous literature studying family and school inputs into children's achievement, far less is known about interactions between school and family inputs. This is surprising because: (i) there is long-standing literature in public economics on public-private crowd in/out, but this issue has been less studied in educational contexts; (ii) input interactions are at the heart of the rapidly growing literature on early (pre-school) childhood development [Cunha *et al.* 2010].

We have studied input interactions by identifying the causal impact of information about school quality on parental time investment into children. Using administrative data linking children, households and schools, we find that when parents receive good news they significantly decrease time investment into their children. This implies that for the average household, beliefs over school quality and parental inputs are substitutes. Hence the ability of public investment into school quality to raise test scores is mitigated because of offsetting responses of parents.

Much of the current literature focuses on 'extensive margin' of school choice or house price responses to information on school quality or accountability. Indeed, the wider literature on information disclosure in public goods markets has also typically focused on the extensive margin [Dranove and Jin 2010]. In sharp contrast we examine one 'intensive margin' of parental responses to school quality ratings for children already in school. This margin is understudied, but affects a far larger cohort of parents than those facing the initial school choice problem.

That inputs endogenously respond to each other is the fundamental difficulty in structurally estimating underlying production functions in education [Becker and Tomes 1976, Todd and Wolpin 2003]. These input interactions: (i) drive a wedge between policy effects (estimated from experimental or quasi-experimental variation) and production function parameters; (ii) in turn, this makes interpreting the causal impact of any given input, especially school-based inputs, difficult without accounting for endogenous responses of family-based inputs; (iii) reinforce/mitigate inequalities across families and schools; (iv) shape the political economy of how the education system is organized and financed [Albornoz *et al.* 2019]. If behavioral responses of families to the same change in school inputs vary across contexts, this limits the external validity of any given study,

Longer term responses might be more muted or of opposite sign as they encompass further responses to the inspections regime by households and schools. For example, Pop-Eleches and Urquiola [2013] show using an RDD in Romania, how being assigned to a higher quality school causes reductions in parental help with homework in the short term, but then these reductions dissipate over academic years for such marginal children. Teacher turnover across academic years has been shown to be impacted by school accountability systems [Feng *et al.* 2010, Figlio and Loeb 2011, Dizon-Ross 2018].

and leads to conflicting results in a given literature.

Indeed, as Pop-Eleches and Urquiola [2013] and Albornoz *et al.* [2019] review, there is extensive literature examining the impact of school quality on test scores literature, but this has produced mixed findings.[34] The key insights arising from recognizing interactions between parental beliefs about school quality and inputs into children are that underlying differences in behavioral responses of households to school quality across contexts is a way to reconcile results across studies.

Given the global roll out of school accountability regimes [Figlio and Loeb 2011], all these issues will be relevant as middle and lower income countries either scale-up current interventions that provide information to parents about schools [Andrabi *et al.* 2017] or start to build school inspection regimes. Global survey data on parents suggests the kinds of issue we document in the English context will be even more relevant in these new settings. Figure 8 shows evidence from a global survey of parents conducted in 2017/18. Across countries at various stages of economic development, parents in lower income countries spend more time providing time inputs into their children. This suggests that household responses to information about school quality may be even larger in low-income settings.[35]

Our results open up a broad agenda of understanding the framing, targeting and specifics of information provision about schools, to increase efficiency in education markets.

A Appendix

A.1 Proofs

Proof of Result 1: Taking the first order condition from (1) and totally differentiating we obtain,

$$\begin{aligned}\frac{de}{dp_H} &= -\frac{u_e(s_H) - u_e(s_L)}{p_H\{u_{ee}(s_H) + u_{ll}(s_H)\} + (1-p_H)\{u_{ee}(s_L) + u_{ll}(s_L)\}} \\ &= -\frac{u_e(s_H) - u_e(s_L)}{\Delta},\end{aligned} \tag{21}$$

where $\Delta = p_H\{u_{ee}(s_H) + u_{ll}(s_H)\} + (1-p_H)\{u_{ee}(s_L) + u_{ll}(s_L)\} < 0$ given than $u(.)$ is concave in e and l. Hence the sign of $\frac{de}{dp_H}$depends on the numerator and it then follows directly that if $u_{es} < 0$ (> 0), then $\frac{de}{dp_H} < 0$ (> 0).■

Proof of Result 2A: From Result 1 we have that, $de = -\frac{1}{\Delta}[u_e(s_H) - u_e(s_L)]dp_H = Kdp_H$ where $K > 0$ in complements case and dp_H is the difference in posterior and prior belief that

[34] Early studies of school quality include Dale and Krueger [2002], Cullen *et al.* [2006] and Hastings *et al.* [2009]. A later wave of studies based on RDDs include Hoekstra [2009] and Jackson [2010]. These find marginal students just gaining admission to high achievement educational institutions have better academic and labor market outcomes. Other papers however find weaker evidence that school quality matters, including Cullen *et al.* [2006], Clark and Del Bono [2016], Duflo *et al.* [2011], Dobbie and Fryer [2011], and Abdulkadiroglu *et al.* [2014].

[35] The survey was conducted on behalf of the Varkey Foundation by Ipsos MORI. They interviewed 27,000 parents in 29 countries using an online survey, in December 2017 and January 2018.

$s = s_H$, p_H is the prior $prob(s = s_H)$, and the posterior beliefs are given in (2). Hence, in response to a positive signal ($j = H$),

$$\begin{aligned} de|_{j=H} &= K[prob(s = s_H|j = H) - p_H] \\ &= Kp_H[\frac{(1-p_H)(2q-1)}{(1-p_H)(1-q)+p_Hq}] \\ &> 0. \end{aligned} \tag{22}$$

Hence parental investment rises if the signal is good and e and s are complements ($K > 0$).■

Proof of Result 2B: Using the same approach as in Result 2A we have that,

$$de|_{j=L} = K\left[prob(s = s_H|j = L) - p_H\right] \tag{23}$$

$$\begin{aligned} &= Kp_H[\frac{(1-p_H)(1-2q)}{(1-p_H)q+p_H(1-q)}] \\ &< 0, \end{aligned} \tag{24}$$

because $(1 - 2q) < 0$. Hence parental investment falls if the signal is bad and e and s are complements ($K > 0$).■

Proof of Result 3: Note first that the difference in the absolute strength of response to a positive versus a negative signal is determined by the denominators in equations (23) and (22). These can be expressed as follows: $(1 - p_H)q + p_H(1 - q) = q - 2p_Hq + p_H \equiv x$ and $(1 - p_H)(1 - q) + p_Hq = 1 - (q - 2p_Hq + p_H) \equiv 1 - x$. Now, if $(x)^2 < (1 - x)^2$ then it follows that:

$$\begin{aligned} 1 - 2x &> 0 \\ q - 2p_Hq + p_H &< \frac{1}{2} \\ p_H &> \frac{1}{2}. \end{aligned}$$

Thus if $p_H > \frac{1}{2}$ then $\left|de|_{j=H}\right| < \left|de|_{j=L}\right|$. Similarly, if $p_H < \frac{1}{2}$ then $\left|de|_{j=H}\right| > \left|de|_{j=L}\right|$ and if $p_H = \frac{1}{2}$ then $\left|de|_{j=H}\right| = \left|de|_{j=L}\right|$.■

A.2 Further Evidence on the Identifying Assumptions

A.2.1 Time Trends

We provide additional evidence underpinning assumption (iii) of no time trends in $\Delta Y_{i\sigma t}$. First, we control for month of household interview in (8). Column 1 of Table A7 shows the results are robust to the inclusion of month of interview dummies, and these dummies are not jointly significant. Second, we construct a placebo check using across-school variation in inspection dates. More precisely, we take schools to be inspected in year $t + 1$ (so a year after survey waves 3 and 5) and

assign next year's inspection date in the current year. This sample is based on 5,242 inspections in 3,269 schools, where we assign all children ($N = 685$) the type of news shock experienced in year $t+1$. The result in Column 2 shows that these future inspection ratings have no relationship with changes in parental investment the year before.

A.2.2 Within-Year School Responses

We now provide evidence in relation to assumption (iv), of no within-year school responses to Ofsted ratings. We first reiterate that Hussain [2015] shows that schools labelled as failing change practices in the short run: they lengthen time devoted to instruction and change their instructional policies. However recall that in our sample only 7% of schools are ranked as failing, and Column 3 in Table A7 shows our core results are robust to dropping them.

All other schools still might be able to adjust on various other margins in the short run. No data on fine-grained adjustments in secondary schools exists for England. To thus shed light on the issue we use the Millennium Cohort Study (MCS), a panel of children tracked since birth in 2000/1, that can be linked to a detailed survey of their teachers. We link the MCS-4 teacher surveys (when the MCS children are age 7) and schools administrative data using school identifiers, to examine short run responses to good and bad news in schools attended by 7 year olds.

This linkage covers MCS schools with an Ofsted inspection in academic years 2007/8 or 2008/9. Our working sample comprises 735 schools and 1,304 teacher surveys (so there can be more than one per school). Schools in our final sample have an average enrolment of 86, as primary schools are smaller than secondary schools from the UKHLS data. 19% of schools have an outstanding rating, 49% are good, 30% are satisfactory and 2% are labelled as failing. This matches closely the evidence on the UKHLS schools in Panel C of Figure 2. In the MCS-4 school sample, 27% of schools have improved ratings over Ofsted cycles, 52% have no change and 21% worsen. This closely matches the distribution of ratings changes in Panel D of Figure 2.

Using information on exact inspection dates and the month of teacher survey, we create a treatment variable equal to one if the teacher interview takes place after the school inspection. We have 471 control observations and 833 treated observations. The samples are balanced on most measures including school size, school type and multiple margins of pupil achievement.

We build a rating forecast model for MCS-4 schools using the procedure described in the main text. We take the universe of inspections in academic years 2007/8 and 2008/9 and run forecasting models analogous to (6) that estimate a school's rating as a function of its past rating, school characteristics and past performance. We construct $news_{st}$ as in (5). The forecasting model displays similar properties as for the schools used in our core analysis. Across specifications we find: (i) a persistence in ratings across inspection cycles: $\widehat{\beta}_0 \in [.316, .415]$; (ii) we always reject a random walk forecasting model; (iii) past exam performance is statistically significant and economically meaningful in predicting inspection outcomes. These features hold when: (i) we restrict the sample

to the same schools as in our analysis of school inputs; (ii) we estimate (6) using an ordered probit model. Across specifications, the forecasting model never has an R-squared higher than .31.

Finally, we estimate a specification analogous to (8) where outcomes are various teaching practices as a function of treatment, $news_{\sigma t}$, and their interaction, conditional on school and teacher controls. We calculate bootstrapped standard errors. Table A8 reports results for various margins of school practice. We see that there is very little change in short run practices across this wide range of dimensions, including the quantity of homework set, the use of teaching assistants or supply teachers, time spent on numeracy and literacy, and the use of streaming, within class ability groups, or subject groups.

A.3 Robustness Checks

We present a battery of robustness checks on the core result. To begin with, Column 4 in Table A7 addresses the concern that the rating prediction model may be subject to greater misclassification error for schools at the tails of the rating distribution. Our results continue to hold when we drop both outstanding and failing schools.

The next check examines the possible strategic delay of bad news by schools. To do so we allow for a longer lag between inspection date and information release date and so address the concerns over non-compliance with treatment for schools with bad news. The core result is unchanged if we omit treated households that are interviewed two, three or four weeks post-inspection (Columns 5 to 7). Interestingly, the point estimates on the DDD in response to good and bad news are all slightly larger than in our baseline specification, suggesting some schools might be engaging in such strategic information delay.

In Column 8 we control for a wider set of school characteristics ($Z_{\sigma t}$); in Column 9 we additionally control for the baseline Ofsted ranking ($ranking_{\sigma t-1}$) in (8); in Column 10 we drop children aged 12 or younger (that are hardly ever in the same school in waves $t-2$ and t). The core findings are robust to all three modifications.

Table A9 then probes the robustness of the core result to using an alternative econometric approach. More precisely, we use a linear probability model for two outcomes: (i) whether the frequency of parental help with homework increases between $t-2$ and t (Panel A); (ii) whether the frequency of parental help with homework decreases between $t-2$ and t (Panel B). Using this alternative set-up delivers a very similar conclusion: there is strong evidence of substitution between parental beliefs about school quality and time investments into their children.

Finally, we examine the sensitivity of the results to alternative forecasting models to (6), thus allowing the assumed underlying information set parents use to vary. Table A10 presents these results, and Figure A4 plots the corresponding sets of marginal effects from each model. In Column 1 we assume parents use an AR(1) model that only conditions on past rating. Column 2 adds school characteristics, Column 3 adds school performance measures (our baseline specification), Column

4 uses an ordered probit first stage to construct the news content of the forecast (that then omits local education authority fixed effects λ_l), and Column 5 presents the naïve model where parents do not use a forecast model but update in response to the change in ratings over inspection cycles (so $news_{\sigma t} = \text{rating}_{\sigma t} - \text{rating}_{\sigma t-1}$). We find the core result to be robust to these alternatives, although the magnitude of responses varies depending on the assumed sophistication of parents. Columns 1 to 3 show that as we add more covariates to the forecasting model, there is a monotonic increase in the (absolute) response of treated households that receive good news. Reassuringly, this all suggests our core result is robust to any small misspecification in the forecasting model.

References

[1] ABDULKADIROGLU.A, J.ANGRIST AND P.A.PATHAK (2017) "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," *Econometrica* 82: 137-96.

[2] ABDULKADIROGLU.A, P.A.PATHAK, J.SCHELLENBERG AND C.R.WALTERS (2017) Do Parents Value School Effectiveness? NBER WP23912.

[3] AGARWAL.N AND P.SOMAINI (2018) "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism," *Econometrica* 86: 391-444.

[4] ALBORNOZ.F, S.BERLINSKI AND A.CABRALES (2019) "Motivation, Resources, and the Organization of the School system," *Journal of the European Economic Association* 16: 199-231.

[5] ANDRABI.T, J.DAS AND A.I.KHWAJA (2017) "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets," *American Economic Review* 107: 1535-63.

[6] AUTOR.D, D.FIGLIO, K.KARBOWNIK, J.ROTH AND M.WASSERMAN (2016) "School Quality and the Gender Gap in Educational Achievement," *American Economic Review* 106: 289-95.

[7] BECKER.G.S AND N.TOMES (1976) "Child Endowments and the Quantity and Quality of Children," *Journal of Political Economy* 84: S143-62.

[8] BLACK.S.E (1999) "Do Better Schools Matter? Parental Valuation of Elementary Education," *Quarterly Journal of Economics* 114: 577-99.

[9] BEUERMANN.D.W, C.K.JACKSON, L.NAVARRO-OLA AND F.PARDO (2018) What is a Good School, and Can Parents Tell? NBER WP25342.

[10] BURGESS.S, D.WILSON AND J.WORTH (2013) "A Natural Experiment in School Accountability: The Impact of School Performance Information on Pupil Progress," *Journal of Public Economics* 106: 57-67.

[11] BURGESS.S, E.GREAVES, A.VIGNOLES AND D.WILSON (2015) "What Parents Want: School Preferences and School Choice," *Economic Journal* 125: 1262-89.

[12] CAETANO.G, J.KINSLER AND H.TENG (2019) "Towards Causal Estimates of Children's Time Allocation on Skill Development," *Journal of Applied Econometrics*, forthcoming.

[13] CARNEIRO.P, K.V.LØKEN AND K.G.SALVANES (2015) "A Flying Start? Maternity Leave Benefits and Long-run Outcomes of Children," *Journal of Political Economy* 123: 365-412.

[14] CLARK.D AND E.DEL BONO (2016) "The Long-run Effects of Attending an Elite School: Evidence from the United Kingdom," *American Economic Journal: Applied Econ.* 8: 150-76.

[15] COLEMAN.J.S, E.Q.CAMPBELL, C.J.HOBSON, J.MCPAERTLAND, A.M.MOOD, F.D.WEINFELD AND R.L.YORK (1966) *Equality of Educational Opportunity*, Washington, DC: U.S. Government Printing Office.

[16] CULLEN.J.B, B.A.JACOB AND S.LEVITT (2006) "The Effect of School Choice on Participants: Evidence from Randomized Lotteries," *Econometrica* 74: 1191-230.

[17] CUNHA.F, J.J.HECKMAN AND S.M.SCHENNACH (2010) "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica* 78: 883-931.

[18] DALE.S.B AND A.B.KRUEGER (2002) "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables," *Quarterly Journal of Economics* 117: 1491-527.

[19] DAS.J, S.DERCON, J.HABYARIMANA, P.KRISHNAN, K.MURALIDHARAN AND V.SUNDARARAMAN (2013) "School Inputs, Household Substitution, and Test Scores," *American Economic Journal: Applied Economics* 5: 29-57.

[20] DATAR.A AND B.MASON (2008) "Do Reductions in Class Size "Crowd Out" Parental Investment in Education?," *Economics of Education Review* 27: 712-23.

[21] DEL BOCA.D, C.FLINN AND M.WISWALL (2014) "Household Choices and Child Development," *Review of Economic Studies* 81: 137-85.

[22] DEL BOCA.D, C.MONFARDINI AND C.NICOLETTI (2017) "Parental and Child Time Investments and the Cognitive Development of Adolescents," *Journal of Labor Econ.* 35: 565-608.

[23] DEL BONO.E, M.FRANCESCONI, Y.KELLY AND A.SACKER (2016) "Early Maternal Time Investment and Early Child Outcomes," *Economic Journal* 126: F96-135.

[24] DE FRAJA.G, T.D'OLIVEIRA AND L.ZANCHI (2010) "Must Try Harder: Evaluating the Role of Effort on Examination Results," *Review of Economics and Statistics* 92: 577-97.

[25] DEMING.D.J, S.COHODES, J.JENNINGS AND C.JENCKS (2016) "School Accountability, Post-secondary Attainment, and Earnings," *Review of Economics and Statistics* 98: 848-62.

[26] DIZON-ROSS.R (2018) "How Does School Accountability Affect Teachers? Evidence from New York City," forthcoming *Journal of Human Resources*.

[27] DIZON-ROSS.R (2019) "Parents' Beliefs and Children's Education: Experimental Evidence from Malawi," forthcoming *American Economic Review*.

[28] DOBBIE.W AND R.G.FRYER JR (2011) "Are High-quality Schools Enough to Increase Achievement Among the Poor? Evidence from the Harlem Children's Zone," *American Economic Journal: Applied Economics* 3: 158-87.

[29] DRANOVE.D AND G.Z.JIN (2010) "Quality Disclosure and Certification: Theory and Practice," *Journal of Economic Literature* 48: 935-63.

[30] DUFLO.E, P.DUPAS AND M.KREMER (2011) "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya," *American Economic Review* 101: 1739-74.

[31] FENG.L, D.N.FIGLIO AND T.SASS (2010) School Accountability and Teacher Mobility, NBER WP16070.

[32] FIGLIO.D.N AND L.W.KENNY (2009) "Public Sector Performance Measurement and Stakeholder Support," *Journal of Public Economics* 93: 1069-77.

[33] FIGLIO.D.N AND S.LOEB (2011) "School Accountability," in *Handbook of the Economics of Education* Vol. 3, pp383-421, Elsevier.

[34] FIGLIO.D.N AND M.E.LUCAS (2004) "What's in a Grade? School Report Cards and the Housing Market," *American Economic Review* 94: 591-604.

[35] FIORINI.M AND M.P.KEANE (2014) "How the Allocation of Children's Time Affects Cognitive and Noncognitive Development," *Journal of Labor Economics* 32: 787-836.

[36] FREDRIKSSON.P, B.OCKERT AND H.OOSTERBEEK (2016) "Parental Responses to Public Investments in Children: Evidence from a Maximum Class Size Rule," *Journal of Human Resources* 51: 832-68.

[37] HANUSHEK.E.A (2010) "Education Production Functions: Developed Country Evidence," in P.Peterson, E.Baker, B.McGaw (eds.) *International Encyclopedia of Education* 2: 407-11.

[38] HANUSHEK.E.A AND M.E.RAYMOND (2005) "Does School Accountability Lead to Improved Student Performance?" *Journal of Policy Analysis and Management* 24: 297-327.

[39] HASTINGS.J, T.J.KANE AND D.O.STAIGER (2009) Heterogeneous Preferences and the Efficacy of Public School Choice, NBER WP2145.

[40] HASTINGS.J AND J.M.WEINSTEIN (2008) "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," *Quarterly Journal of Economics*, 123: 1373-414.

[41] HOEKSTRA.M (2009) "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-based Approach," *Review of Economics and Statistics* 91: 717-24.

[42] HOUTENVILLE.A.J AND K.S.CONWAY (2008) "Parental Effort, School Resources, and Student Achievement," *Journal of Human Resources* 43: 437-53.

[43] HOXBY.C.M AND C.AVERY (2012) The Missing "one-offs": The Hidden Supply of High-achieving, Low Income Students, NBER WP 18586.

[44] HUSSAIN.I (2015) "Subjective Performance Evaluation in the Public Sector. Evidence from School Inspections," *Journal of Human Resources* 50: 189-221.

[45] HUSSAIN.I (2017) Do Consumers Respond to Short-Term Innovations in School Productivity? Evidence From the Housing Market and Parents' School Choices, University of Sussex mimeo.

[46] JACKSON.C.K (2010) "Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago," *Economic Journal* 120: 1399-429.

[47] JACOB.B.A (2005) "Accountability, Incentives and Behavior: The Impact of High-stakes Testing in the Chicago Public Schools," *Journal of Public Economics* 89: 761-796.

[48] JACOB.B.A AND L.LEFGREN (2007) "What Do Parents Value in Education? An Empirical Investigation of Parents' Revealed Preferences for Teachers," *Quarterly Journal of Economics* 122: 1603-37.

[49] JACOB.B.A AND S.D.LEVITT (2003) "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," *Quarterly Journal of Economics* 118: 843-77.

[50] JENSEN.R (2010) "The (Perceived) Returns to Education and the Demand for Schooling," *Quarterly Journal of Economics* 125: 515-48.

[51] JOHNSON.P (2004) "Education Policy in England," *Oxford Review of Econ. Pol.* 20: 173-97.

[52] KANE.T.J AND D.O.STAIGER (2002) "The Promise and Pitfalls of Using Imprecise School Accountability Measures," *Journal of Economic Perspectives* 16: 91-114.

[53] LIU.H, T.A.MROZ AND W.VAN DER KLAAUW (2010) "Maternal Employment, Migration, and Child Development," *Journal of Econometrics* 156: 212-28.

[54] MURPHY.K.M AND R.H.TOPEL (1985) "Estimation and Inference in Two-step Econometric Models," *Journal of Business and Economic Statistics* 3: 370-79.

[55] D.NEAL AND D.WHITMORE SCHANZENBACH "Left Behind by Design: Proficiency Counts and Test-Based Accountability," *Review of Economics and Statistics* 92: 263-83.

[56] POP-ELECHES.C AND M.URQUIOLA (2013) "Going to a Better School: Effects and Behavioral Responses," *American Economic Review* 103: 1289-324.

[57] ROUSE.C.E, J.HANNAWAY, D.GOLDHABER AND D.FIGLIO (2013), "Feeling the Florida Heat? How Low-performing Schools Respond to Voucher and Accountability Pressure," *American Economic Journal: Economic Policy* 5: 251-81.

[58] TODD.P.E AND K.I.WOLPIN (2003) "On the Specification and Estimation of the Production Function for Cognitive Achievement," *Economic Journal* 113: F3-33.

[59] UKHLS (2018) University of Essex Institute for Social and Economic Research, NatCen Social Research, Kantar Public. Understanding Society: Waves 1-8, 2009-2017 and Harmonised BHPS: Waves 1-18, 1991-2009, 11th Edition; UK Data Service, SN: 6614, http://doi.org/10.5255/UKDA-SN-6614-12.

Table 1: Balance

Means, Standard Deviation in Parentheses, p-values in Brackets

	(1) Treated: Interviewed After Ofsted Inspection	(2) Control: Interviewed Before Ofsted Inspection	(3) Normalized Difference	(4) Test of Equality [p-value]
Number of Children	402	288		
A. School Characteristics				
School Size: Number of Pupils	1128	1098	.058	[.308]
	(372)	(364)		
% Pupils Free School Meals	17.5	17.1	.020	[.754]
	(12.9)	(14.4)		
Academy School	.216	.288	-.117	[.035]
Boys School	.052	.035	.061	[.251]
% Pupils 5 or More A*-C grades	.772	.780	-.032	[.580]
	(.177)	(.152)		
Total Average GCSE Point Score	331	335	-.053	[.377]
	(49.4)	(40.1)		
B. Household Characteristics				
Household Size	4.13	4.20	-.037	[.512]
	(1.31)	(1.29)		
Home Owner	.632	.649	-.026	[.654]
C. Child Characteristics				
Female	.520	.476	.062	[.294]
Age	13.5	13.4	.034	[.518]
	(1.09)	(1.14)		
D. Mother Characteristics				
Married/cohabiting	.714	.722	-.014	[.803]
White Ethnicity	.745	.738	.012	[.826]
Education GCSE or Below	.432	.452	-.029	[.572]
E. Father Characteristics				
Married/cohabiting	.969	.938	.103	[.209]
White Ethnicity	.749	.793	-.075	[.359]
Education GCSE or Below	.466	.393	.104	[.194]

Notes: The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. Columns 1 and 2 show means and standard deviations in parentheses for treated and control households respectively. Column 3 shows normalized differences between treatment and control groups, namely the difference in sample means divided by the square root of the sum of the variances. The p-values shown in Column 4 are derived by regressing the characteristic on a treatment dummy and clustering standard errors by local authority.

Table 2: Ofsted Ratings and the Timing of Inspection

Dependent Variable: Ofsted Grade

Linear Regression, Standard Errors Clustered by Local Authority

	(1) Timing of Inspection	(2) Plus Previous Grade	(3) Plus Timing of Previous Inspection	(4) Plus School Characteristics	(5) Ordered Probit Model
Month of Ofsted Inspection					
September	-.269	.048	.045	.044	-.293
	(.216)	(.184)	(.196)	(.185)	(.327)
October	-.169	.155	.135	.085	.159
	(.192)	(.175)	(.179)	(.180)	(.327)
November	-.194	.054	.063	-.013	-.018
	(.229)	(.185)	(.196)	(.178)	(.331)
December	-.416	.037	.082	-.022	.135
	(.266)	(.24)	(.233)	(.225)	(.367)
January	-.365*	-.042	.040	-.107	.324
	(.213)	(.203)	(.212)	(.224)	(.345)
February	-.293	-.050	-.051	-.165	.542*
	(.215)	(.217)	(.222)	(.206)	(.326)
March	-.359*	.039	.093	-.101	.211
	(.2)	(.188)	(.198)	(.199)	(.289)
April	.130	.250	.310	.201	.011
	(.238)	(.207)	(.213)	(.170)	(.382)
May	-.222	.040	.0688	-.062	.161
	(.198)	(.169)	(.19)	(.183)	(.312)
June	-.169	-.052	.003	-.087	.237
	(.204)	(.176)	(.183)	(.178)	(.308)
Previous Ofsted Grade		.409***	.418***	.230***	-.502***
		(.055)	(.055)	(.054)	(.080)
F test: month dummies [p-value]	1.414 **[.180]**	.458 **[.914]**	.479 **[.901]**	.866 **[.567]**	Chi2 15.67**[.110]**
F test: Sept-Dec dummies [p-value]	.729 **[.574]**	.230 **[.921]**	.184 **[.946]**	.186 **[.945]**	Chi2 3.647**[.456]**
Timing of Previous Grade Controls	No	No	Yes	Yes	Yes
F test: timing of previous inspection dummies [p-value]			1.775 **[.071]**	1.527 **[.136]**	Chi2 15.08**[.129]**
School Characteristics	No	No	No	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	No
Number of Observations	690	690	690	690	690
Number of Schools	548	548	548	548	548

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The sample comprises schools that are inspected during the academic years in which the working sample of UKHLS households are interviewed. Hence, the unit of observation is a school. The outcome variable is the Ofsted inspection grade, where this can take the following values: 4 (Outstanding), 3 (Good), 2 (Requires Improvement) and 1 (Inadequate/failing). Columns 1 to 4 present OLS regressions of Ofsted grades on month of inspection and school level control variables. Column 1 controls for academic year, Column 2 additionally controls for the last Ofsted grade, Column 3 additionally controls for month of previous inspection and a dummy coding previous inspection month missing, Column 4 further adds controls for school composition, type and performance (14 controls). Column 5 presents the same specification as in Column 4 but using an ordered probit model. The lower panel shows F-tests (Chi-2 test in Column 5) and corresponding p-values in brackets for the joint significance of all month of inspection dummies, the joint significance of the September to December month of inspection dummies, and for the joint significance of the dummies indicating the timing of the previous inspection.

Table 3: Parental Response to Information on School Quality

Ordered Probit Regression Estimates

Bootstrapped Standard Errors in Parentheses, Clustered by Local Authority

	(1) Forecast, Unconditional	(2) Plus Child Characteristics	(3) Plus Parent Characteristics	(4) Plus School Characteristics
T-C | good news ($\beta_0+\beta_1$)	-.498**	-.488**	-.495**	-.518**
	(.196)	(.196)	(.207)	(.207)
T-C | bad news ($\beta_0+\beta_2$)	.041	.069	.078	.109
	(.228)	(.234)	(.218)	(.219)
Diff-in-Diff-in-Diff ($\beta_1-\beta_2$)	-.539*	-.558*	-.573*	-.627**
	(.310)	(.314)	(.303)	(.305)
Forecast Ofsted Rating	Yes	Yes	Yes	Yes
Child Characteristics	No	Yes	Yes	Yes
Parent Characteristics	No	No	Yes	Yes
School Characteristics	No	No	No	Yes
Observations	690	690	690	690

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. The outcome variable is the change in parental help with homework between survey wave 3 and 1, or between survey wave 5 and 3: this is coded as -1 if there is a decrease in parental help, 0 if there is no change in parental help and +1 is there an increase in parental help with homework. Ordered probit regression estimates are shown. In all Columns, the specification uses the predicted news shock. In Column 1 we control for a treatment dummy, interactions between the treatment dummy and dummies for whether a positive or negative news shock is observed, and the dummies for a positive or negative news shock. Column 2 additionally controls for child and household characteristics (gender and age dummies, household size, number of children in the household and dummies for housing tenure), Column 3 additionally controls for parental characteristics (ethnicity, highest educational degree and marital status), and Column 4 additionally controls for school characteristics (school size and proportion of children eligible for free school meals). Standard errors are derived using the bootstrap method with 1,000 iterations, clustered at the local authority level and shown in parentheses.

Table 4: Parental Response to Information on School Quality, Marginal Effects

Ordered Probit Marginal Effect Estimates

Bootstrapped Standard Errors in Parentheses, Clustered by Local Authority

	(1) Probability of Increasing Parental Time Investment	(2) Probability Parental Time Investment Unchanged	(3) Probability of Decreasing Parental Time Investment
T-C | good news ($\beta_0+\beta_1$)	-.112**	-.082***	.195***
	(.048)	(.030)	(.073)
T-C | bad news ($\beta_0+\beta_2$)	.028	.012	-.040
	(.054)	(.025)	(.078)
Diff-in-Diff-in-Diff ($\beta_1-\beta_2$)	-.141*	-.094**	.235**
	(.074)	(.040)	(.108)
Forecast Ofsted Rating	Yes	Yes	Yes
Child Characteristics	Yes	Yes	Yes
Parent Characteristics	Yes	Yes	Yes
School Characteristics	Yes	Yes	Yes
Number of Observations	690	690	690

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. The outcome variable is the change in parental help with homework between survey wave 3 and 1, or between survey wave 5 and 3: this is coded as -1 if there is a decrease in parental help, 0 if there is no change in parental help and +1 is there an increase in parental help with homework. The marginal effects of the coefficients from an ordered probit regression for increasing/not changing/decreasing help with homework at home are shown. The specification used controls for a treatment dummy, interactions between the treatment dummy and dummies for whether a positive or negative news shock is observed, dummies for a positive or negative news shock, child and household characteristics (gender and age dummies, household size, number of children in the household and dummies for housing tenure), parental characteristics (ethnicity, highest educational degree and marital status), and school characteristics (school size and proportion of children eligible for free school meals). Standard errors are derived using the bootstrap method with 1,000 iterations, clustered at the local authority level and shown in parentheses.

Table 5: The Distribution of News and Parental Time Investments

	News				Parental Time Investment
	(1) prob(good | s = s_k) *{Regression Adjusted}*	(2) prob(no | s = s_k) *{Regression Adjusted}*	(3) prob(bad | s = s_k) *{Regression Adjusted}*	Share of Schools	(4) Unconditional *{Regression Adjusted}*
Ofsted Rating from Earlier Inspection (s = s_k)					
Outstanding	.330	.510	.160	.182	.849
	{.513}	*{.183}*	*{.305}*		*{.145}*
Good	.222	.587	.191	.411	.825
	{.430}	*{.243}*	*{.326}*		*{.132}*
Requires Improvement	.254	.552	.194	.367	.803
	{.474}	*{.198}*	*{.328}*		*{.122}*
Inadequate	.136	.818	.045	.040	.725
	{.363}	*{.495}*	*{.142}*		*{.042}*
Share of Schools	.250	.569	.181		-
Observations	548	548	548		2,955

Notes: The sample in Columns 1 to 3 is based on those schools used in the main analysis. The news shock descriptives show the distribution of news shock (good news/no news/zero news) in Columns, by the schools Ofsted grade in the previous inspection cycle, in each row. The regression adjusted descriptives, in italics and parentheses, are residuals of linear regressions of a positive (no, negative) shock on child, household, parental and school characteristics and a local authority fixed effect, plus the constant of this regression. The sample in Columns 4 is based on the pooled sample of households with a non-missing outcome , non-missing school codes and covariates, and omitting schools inspected during the year t-2. Parental investment is a binary variable capturing high investment defined as frequency of help with homework being almost every day, or at least once a week (so the top two points of the five-point Likert scale). Column 4 shows raw and regression adjusted means are derived in the same way as for the news shock descriptives. The displayed regression adjusted value is the sum of the constant and the mean residual for each Ofsted score.

Table 6: Test Score Impacts of the School Inspection Regime

OLS Regression Estimates

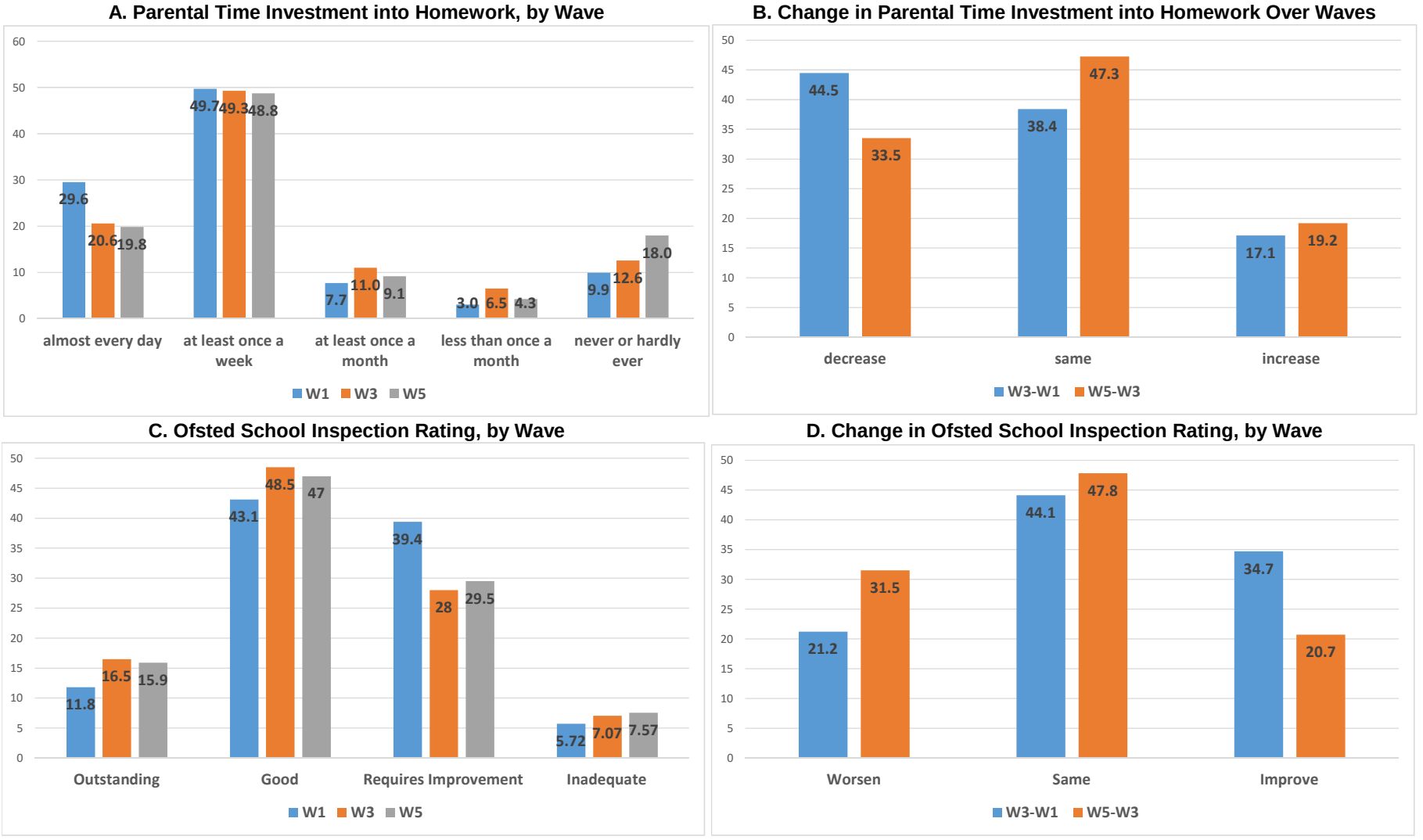
Bootstrapped Standard Errors in Parentheses, Clustered by Local Authority

90% Confidence Intervals in Brackets

	Student-level	School-level	
	(1) Standardized GCSE Average Point Score	(2) Standard Deviation	(3) Interquartile Range
T-C | good news ($\beta_0+\beta_1$)	-0.091**	0.002	-0.011
	(0.044)	(0.012)	(0.025)
	[-.164, -.0179]	[-.018, .022]	[-.051, .029]
T-C | bad news ($\beta_0+\beta_2$)	-0.048	0.002	0.005
	(0.045)	(0.019)	(0.026)
	[-.122, .026]	[-.028, .032]	[-.037, .047]
Forecast Ofsted Rating	Yes	Yes	Yes
School Characteristics	Yes	Yes	Yes
Pupil Characteristics	Yes	Yes	Yes
Number of Schools	1,143	1,143	1,143
Number of Pupils	203,500	-	-

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The sample comprises schools inspected by Ofsted in one of 2011, 2012 or 2013 academic years. If a school experiences more than one inspection in this three-year period, the first inspection event is selected. All regressions include year dummies. In Column 1, the dependent variable is the student-level standardized average point score on the age-16 GCSE exam. In Columns 2 and 3 the dependent variables are the school-level standard deviation in test scores and the interquartile range, respectively. For Column 1, the regression includes lagged test scores (student's age-11 Key Stage 2 performance). For Columns 2 and 3, regressions include lagged school-level standard deviation and lagged interquartile range, respectively. Treated (control) schools are defined as those where the OFSTED inspection took place in the early (late) part of the academic year. Early is defined as September through December; late is defined as January through April. GCSE exams take place in May and June. All regressions include a treatment dummy (inspected early) as well as dummies for positive and negative shocks. All regressions also include dummies for type of school (community, academy, voluntary aided, etc.), school's religious status, school's admission policy, single-sex entry, percent students eligible for free school meals, percent students speaking English as an additional language; total enrolment. Column 1 also includes student's eligibility for free lunch, ethnic minority dummy, special education needs status and gender. Schools failed in any of the years 2010 to 2013 are dropped (failed schools may be subject to local authority intervention). Standard errors are clustered at the Local Authority level. To account for generated regressors in the forecast model, standard errors are derived using the bootstrap method with 1,000 iterations. Standard errors are shown in parentheses, with 90% confidence intervals in brackets.

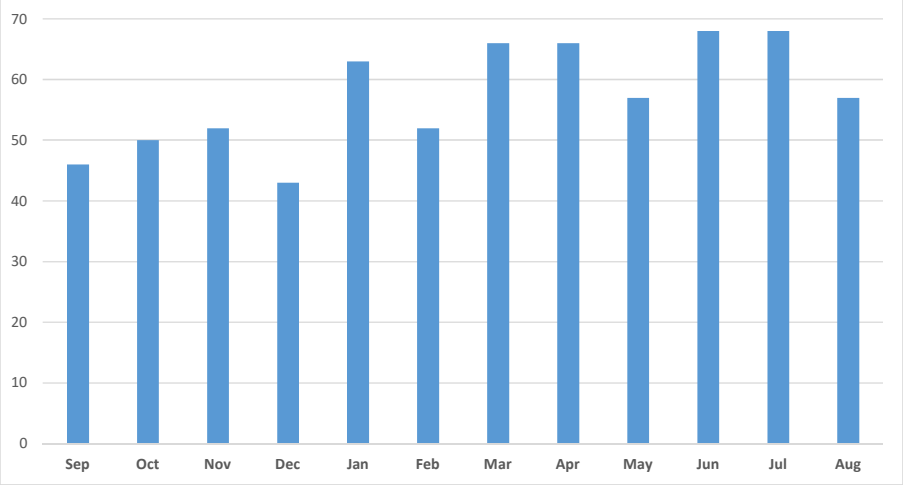
Figure 1: Parental Investment and Ofsted Ratings



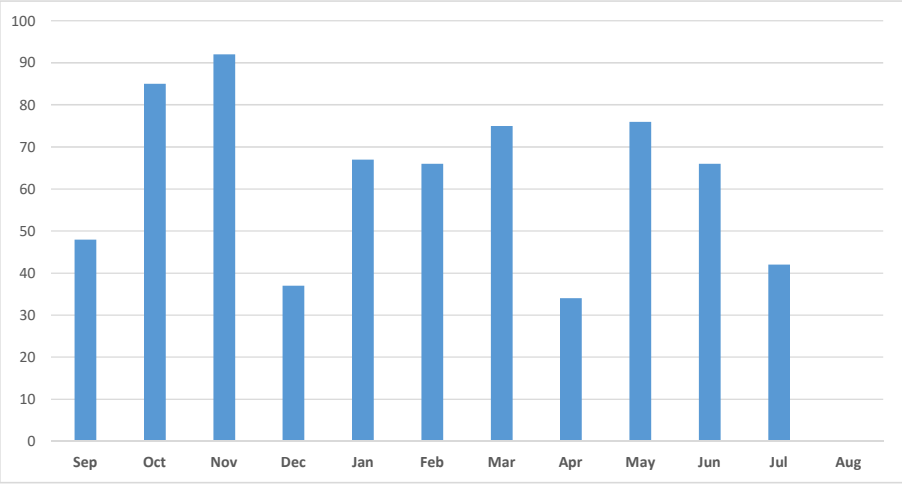
Notes: The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Panel A shows the distribution of parental time investment by wave, Panel B shows changes in parental time investment, separately for changes between survey waves 3 and 1 and survey waves 5 and 3. Panel C shows the distribution of children by the Ofsted inspection rating of their school and wave. Panel D shows the proportion of children with a worse, same or improved Ofsted rating, compared to the last rating of their school.

Figure 2 Timing of Household Surveys and OFSTED Inspections

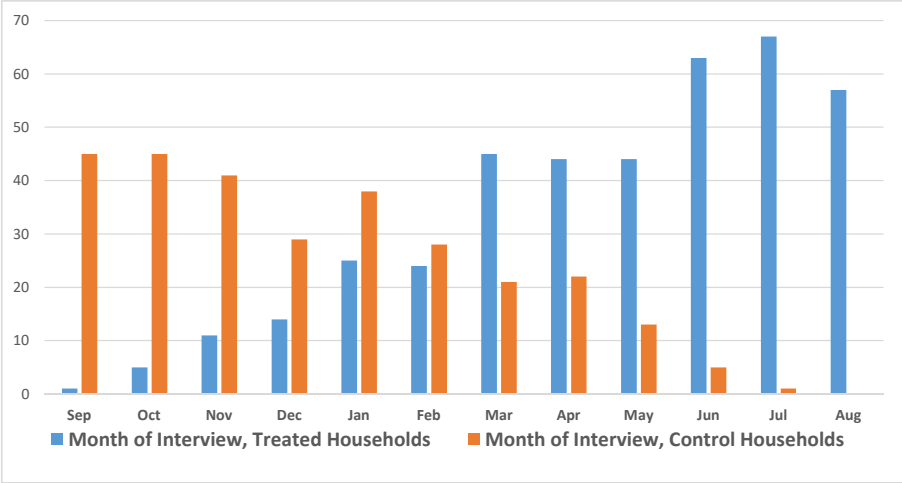
A. Household Survey Interviews



B. OFSTED Inspections



C. Household Survey Interviews: Treated and Control Groups



Notes: In Panels A and C, the sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. In Panel B the sample comprises the schools the children in these households attend. Treated (control) schools are defined as those whose date of inspection occurs after (before) the dates of UKHLS interviews.

Figure 3 Research Design

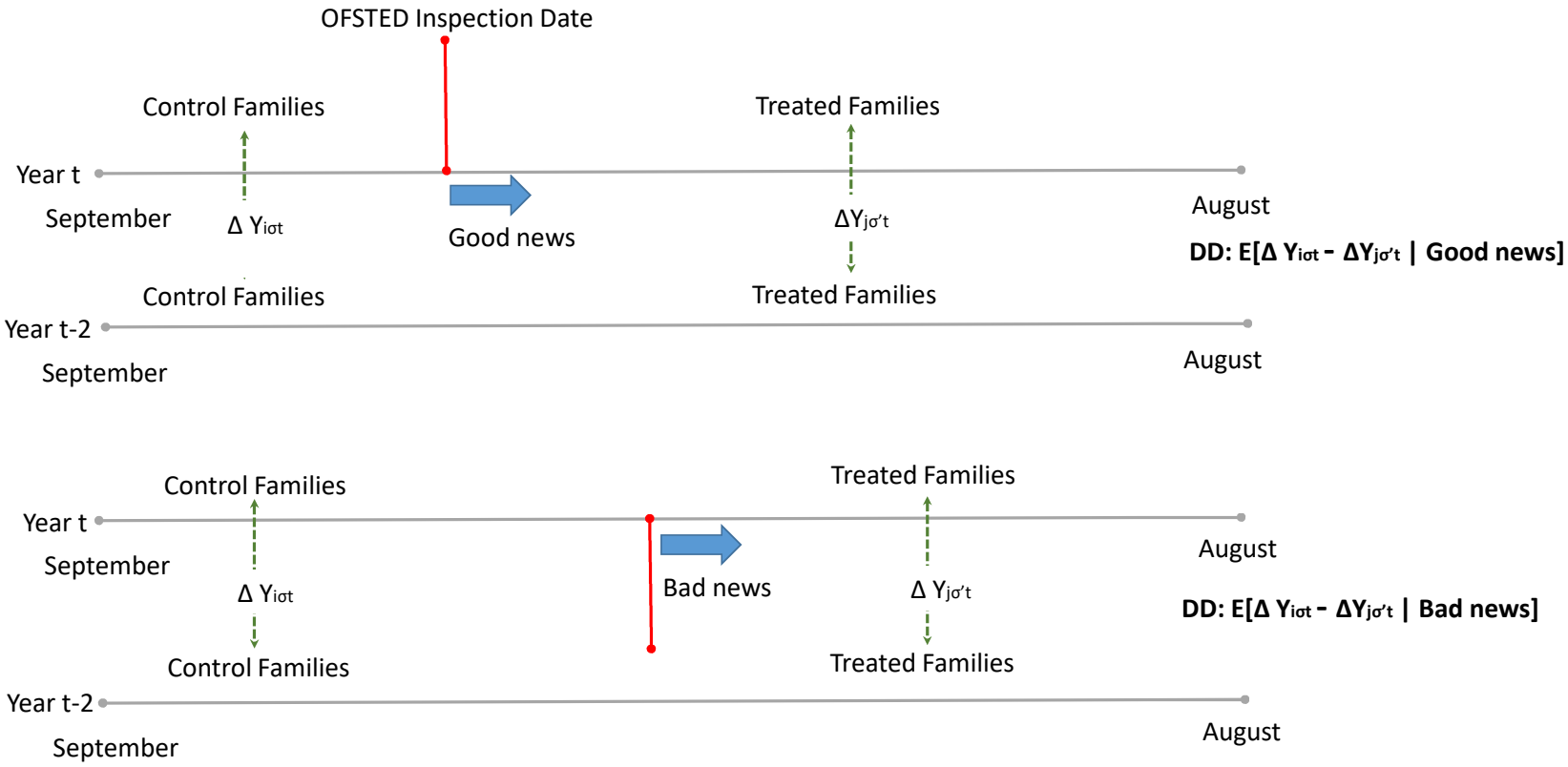
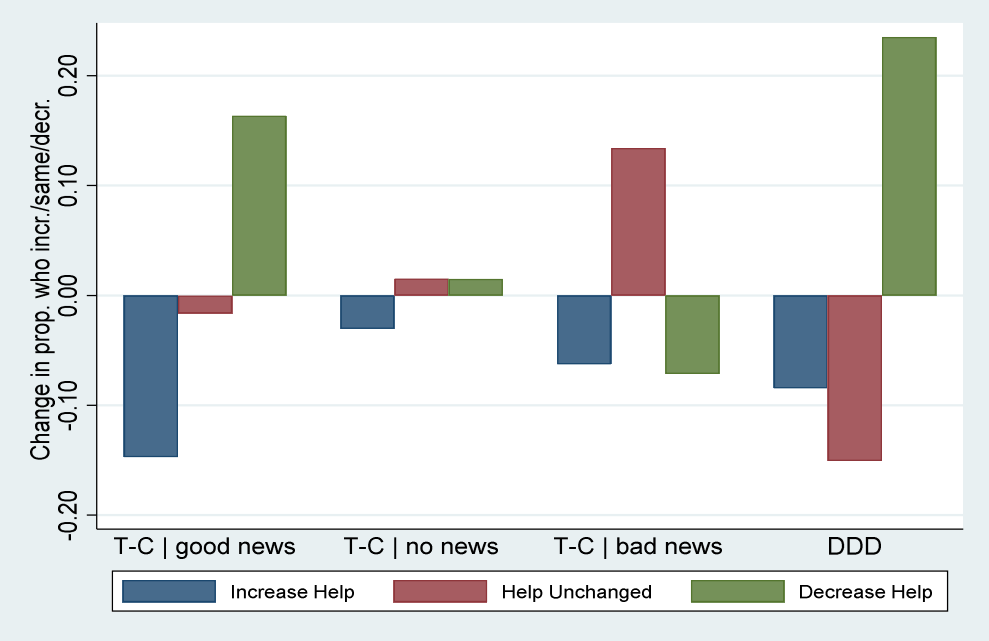
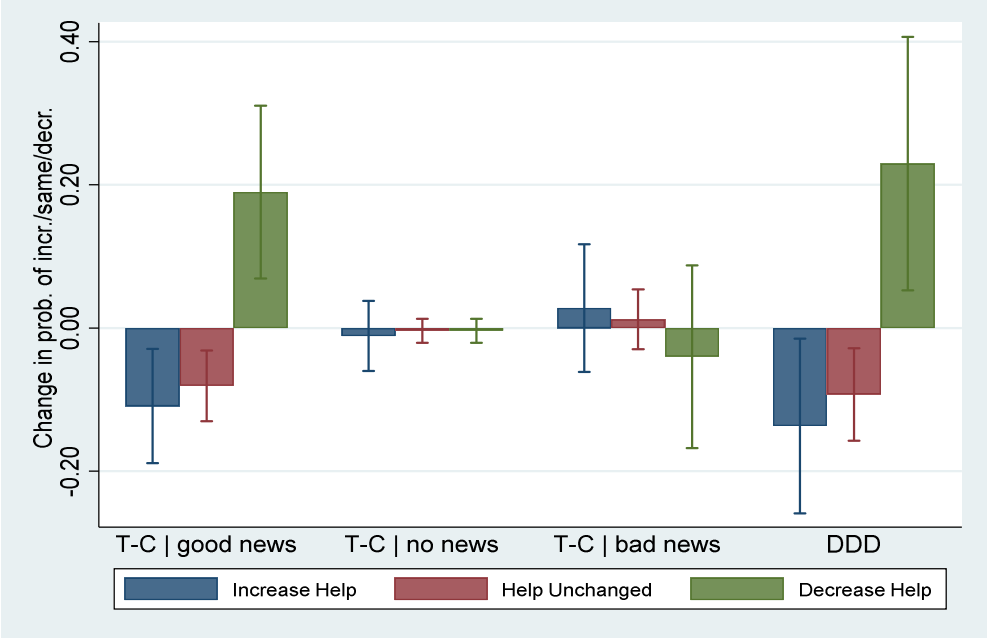


Figure 4: ΔParental Investment by News



Notes: The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. The Figure shows the difference in the proportion of parents who increase/do not change/decrease their help with homework when they receive a positive/no/negative shock about the quality of their child's school, compared to receiving the same shock in the future (i.e. the difference between treatment and control groups). The last set of bars show the triple difference, i.e. the difference in the proportion of parents between treatment and control group who increase/do not change/decrease their help with homework when they receive a positive rather than a negative shock. An increase (decrease) in parental help is defined as parents helping more (less) at wave 3 than at wave 1 or at wave 5 than at wave 3.

Figure 5: Marginal Impacts of Information on School Quality on Parental Investment



Notes: The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. The outcome variable is the change in parental help with homework between survey wave 3 and 1, or between survey wave 5 and 3: this is coded as -1 if there is a decrease in parental help, 0 if there is no change in parental help and +1 is there an increase in parental help with homework. The marginal effects of the coefficients from an ordered probit regression for increasing/not changing/decreasing help with homework at home are shown, along with 90% confidence intervals. The standard errors used to derive 90% confidence intervals are clustered at the local authority level and derived using the bootstrap method with 1,000 iterations.

Figure 6: Distributional Impacts of the Schools Inspection Regime on Parental Investment

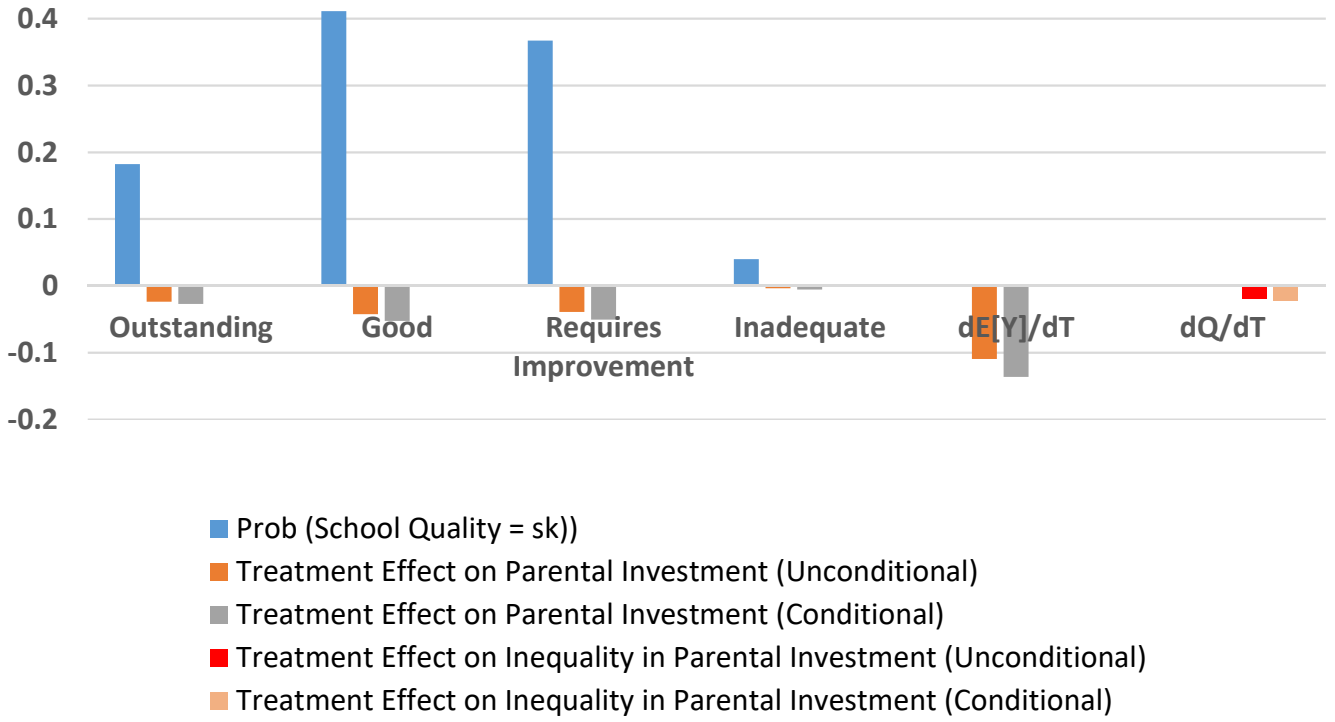
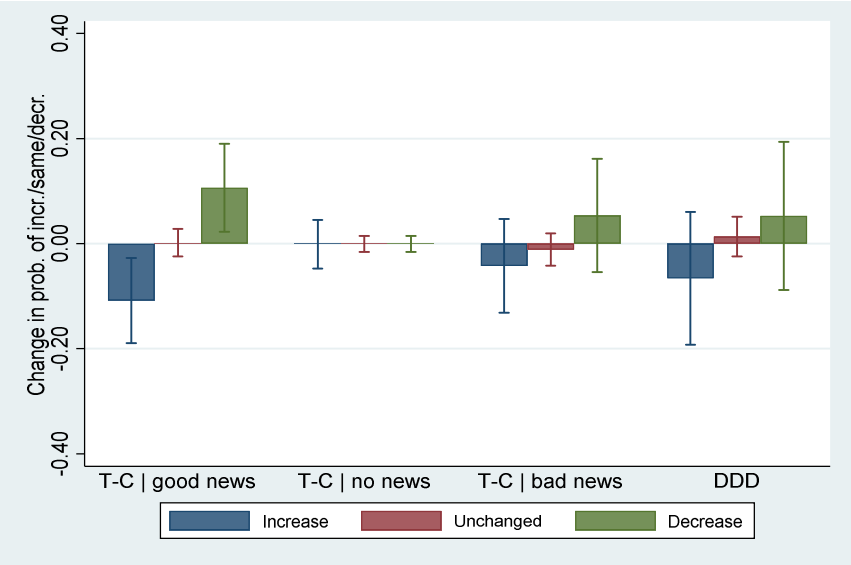
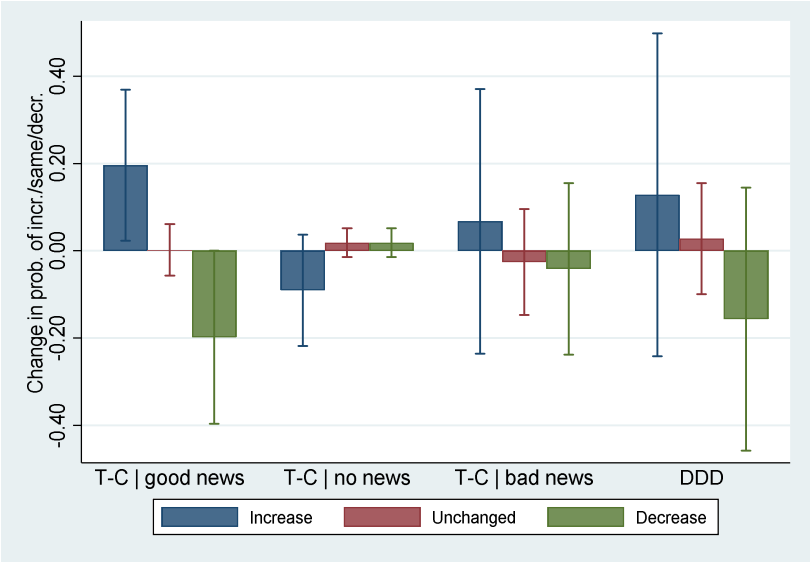


Figure 7: Academic Responses to School Quality Information

A. Child Talks with Parent About Things That Matter



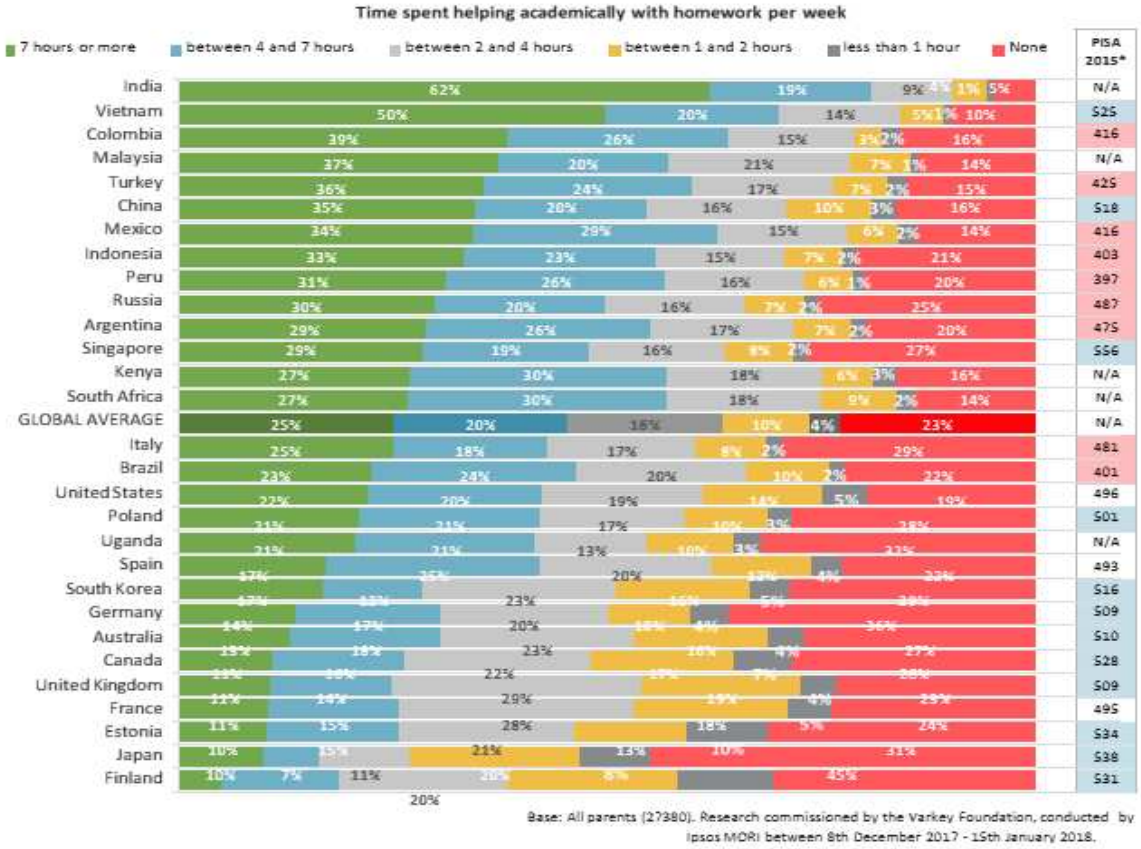
B. Child (Own) Time Investment: Hours Spent on Homework



Notes: Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. In Panel A, the sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing responses to the outcome collected in waves 1, 3 and 5. Change in how often child talks about things that matter is constructed using the question, "How often does your child/your children talk to you about things that matter?" (N=710). In Panel B the sample is based on UKHLS households with an Ofsted school inspection in the same academic year as interview and with non-missing outcome variable collected from the young person at waves 2 and 4. The change in hours the child spends doing homework is derived from the response to following question across waves, "When you do homework on a week-day evening during term time, how many hours do you usually spend doing your homework?" (N=244). The change in tutoring in school subjects is derived from responses to a list of after school activities (yes/no) (N=229). The marginal effects of the coefficients from an ordered probit regression for increasing/not changing/decreasing help with homework at home are shown, along with 90% confidence intervals. The standard errors used to derive 90% confidence intervals are clustered at the local authority level and derived using the bootstrap method with 1,000 iterations.

Figure 8: Global Survey Of Parents

On Average, How Much Time, if any, Do you Personally Spend Helping your Child Academically with their Education per Week?



Source: Varkey Foundation (2018), Global Parents Survey. The survey was conducted by Ipsos MORI. They interviewed 27,000 parents in 29 countries using an online survey, in December 2017 and January 2018. All countries had 1,000 interviews except Estonia (500), Kenya (501) and Uganda (371). The data presented in Figure 8 is weighted by age, gender and region of child and corrected for gender of parent. The survey is representative of parents of children aged 4-18 in education, based on these characteristics, with equal views from mothers and fathers. For countries with low internet penetration (India, Uganda, Kenya, Peru and Indonesia), the data is representative of the urban online population.

Table A1: Key Ofsted Judgements and Grade Descriptors

Part A: Criteria for Judgements

Judgement 1: Achievement of Pupils at the School

When evaluating the achievement of pupils, inspectors must consider:

(i) the standards attained by pupils by the time they leave the school, including their standards in reading, writing and mathematics and, in primary schools, pupils' attainment in reading by the end of Key Stage 1 and by the time they leave the school

(ii) how well pupils learn, the quality of their work in a range of subjects and the progress they have made since joining the school

(iii) how well pupils develop a range of skills, including reading, writing, communication and mathematical skills, and how well they apply these across the curriculum

(iv) how well disabled pupils and those who have special educational needs have achieved since joining the school

(v) how well gaps are narrowing between the performance of different groups of pupils in the school and compared to all pupils nationally

(vi) how well pupils make progress relative to their starting points.

Judgement 2: Quality of Teaching in the School

When evaluating the quality of teaching in the school, inspectors must consider:

(i) the extent to which teachers' expectations, reflected in their teaching and planning, including curriculum planning, are sufficiently high to extend the previous knowledge, skills and understanding of all pupils in a range of lessons and activities over time

(ii) how well teaching enables pupils to develop skills in reading, writing, communication and mathematics

(iii) the extent to which well judged teaching strategies, including setting challenging tasks matched to pupils' learning needs, successfully engage all pupils in their learning

(iv) how well pupils understand how to improve their learning as a result of frequent, detailed and accurate feedback from teachers following assessment of their learning

(v) how well pupils understand how to improve their learning as a result of frequent, detailed and accurate feedback from teachers following assessment of their learning

(vi) the extent to which teachers' questioning and use of discussion promote learning

(vii) the extent to which the pace and depth of learning are maximised as a result of teachers' monitoring of learning during lessons and any consequent actions in response to pupils' feedback

(viii) the extent to which teachers enthuse, engage and motivate pupils to learn and foster their curiosity and enthusiasm for learning

(ix) how well teachers use their expertise, including their subject knowledge, to develop pupils' knowledge, skills and understanding across a range of subjects and areas of learning

(x) the extent to which teachers enable pupils to develop the skills to learn for themselves, where appropriate, including setting appropriate homework to develop their understanding

(xi) the quality of teaching and other support provided for pupils with a range of aptitudes and needs, including disabled pupils and those who have special educational needs, so that their learning improves.

Source: OFSTED, The Evaluation Schedule for the Inspection of Maintained School and Academies, April 2012, OFSTED document reference number 090098. Available via the UCL Institute of Education Digital Education Resource Archive: http://dera.ioe.ac.uk/14076/1/The_evaluation_schedule_for_school_inspections_from_January_2012%5B1%5D.pdf

,

Table A1: Key Ofsted Judgements and Grade Descriptors (cont.)

Part A: Criteria for Judgements

Judgement 3: Behaviour and Safety of Pupils at the School

When evaluating the behaviour and safety of pupils at the school, inspectors must consider:

(i) pupils' attitudes to learning and conduct in lessons and around the school

(ii) pupils' behaviour towards, and respect for, other young people and adults, including, for example, freedom from bullying and harassment that may include cyber-bullying and prejudice-based bullying related to special educational need, sexual orientation, sex, race, religion and belief, gender reassignment or disability

(iii) how well teachers manage the behaviour and expectations of pupils to ensure that all pupils have an equal and fair chance to thrive and learn in an atmosphere of respect and dignity

(iv) pupils' ability to assess and manage risk appropriately and keep themselves safe

(v) pupils' attendance and punctuality at school and in lessons

(vi) how well the school ensures the systematic and consistent management of behaviour.

Judgement 4: Quality of Leadership in and Management of the School

When evaluating the quality of leadership and management inspectors must consider whether the school's leadership:

(i) demonstrates an ambitious vision for the school and high expectations for what every pupil and teacher can achieve, and sets high standards for quality and performance

(ii) improves teaching and learning, including the management of pupils' behaviour

(iii) provides a broad and balanced curriculum that: meets the needs of all pupils; enables all pupils to achieve their full educational potential and make progress in their learning; and promotes their good behaviour and safety and their spiritual, moral, social and cultural development

(iv) evaluates the school's strengths and weaknesses and uses their findings to promote improvement

(v) improves the school and develops its capacity for sustaining improvement by developing leadership capacity and high professional standards among all staff

(vi) engages with parents and carers in supporting pupils' achievement, behaviour and safety and their spiritual, moral, social and cultural development

(vii) ensures that all pupils are safe.

Source: OFSTED, The Evaluation Schedule for the Inspection of Maintained School and Academies, April 2012, OFSTED document reference number 090098. Available via the UCL Institute of Education Digital Education Resource Archive: http://dera.ioe.ac.uk/14076/1/The_evaluation_schedule_for_school_inspections_from_January_2012%5B1%5D.pdf

Table A1: Key Ofsted Judgements and Grade Descriptors (cont.)

Part B: Rating Descriptors

Rating	Achievement of Pupils at the School	Quality of Teaching in the School
Outstanding (coded 4)	Almost all pupils, including where applicable disabled pupils and those with special educational needs, are making rapid and sustained progress in most subjects over time given their starting points. They learn exceptionally well and as a result acquire knowledge quickly and in depth and are developing their understanding rapidly in a wide range of different subjects across the curriculum, including those in the sixth form and areas of learning in the Early Years Foundation Stage. They develop and apply a wide range of skills to great effect, including reading, writing, communication and mathematical skills across the curriculum that will ensure they are exceptionally well prepared for the next stage in their education, training or employment. The standards of attainment of almost all groups of pupils are likely to be at least in line with national averages for all pupils with many above average. In exceptional circumstances where standards of attainment, including attainment in reading in primary schools, ofived from responses to a list of after school activities (yes/no) (N=229). The marginal effects of the coefficients from an ordered probit regression for increasing/no	Much of the teaching in all key stages and most subjects is outstanding and never less than consistently good. As a result, almost all pupils are making rapid and sustained progress. All teachers have consistently high expectations of all pupils. Drawing on excellent subject knowledge, teachers plan astutely and set challenging tasks based on systematic, accurate assessment of pupils' prior skills, knowledge and understanding. They use well judged and often imaginative teaching strategies that, together with sharply focused and timely support and intervention, match individual needs accurately. Consequently, pupils learn exceptionally well across the curriculum. The teaching of reading, writing, communication and mathematics is highly effective. Teachers and other adults generate high levels of enthusiasm for, participation in and commitment to learning. Teaching promotes pupils' high levels of resilience, confidence and independence when they tackle challenging activities. Teachers systematically and effectiived from responses to a list of after school activities (yes/no) (N=229). The marginal effects of the coefficients from an ordered probit regression for increasing/not changing/decreasing help with homework at home are shown, along with 90% confidence intervals. The standard errors used to derive 90% confidence intervals are clustered at the local authority level and derived using the bootstrap method with
Good (coded 3)	Pupils are making better progress than all pupils nationally given their starting points. Groups of pupils, including disabled pupils and those with special educational needs, are also making better progress than similar groups of pupils nationally. Performance will exceed floor standards. Pupils acquire knowledge quickly and are secure in their understanding in different subjects. They develop and apply a range of skills well, including reading, writing, communication and mathematical skills, across the curriculum that will ensure they are well prepared for the next stage in their education, training or employment. The standards of attainment of the large majority of groups of pupils are likely to be at least in line with national averages for all pupils. Where standards of any group of pupils are below those of all pupils nationally, the gaps are closing. In exceptional circumstances, where attainment, including attainment in reading in primary schools, is low overall, it is improving at a faster rate than ived from responses to a list of af	As a result of teaching that is mainly good, with examples of outstanding teaching, most pupils and groups of pupils, including disabled pupils and those who have special educational needs, are achieving well over time. Teachers have high expectations of all pupils. Teachers in most subjects and key stages use their well developed subject knowledge and their accurate assessment of pupils' prior skills, knowledge and understanding to plan effectively and set challenging tasks. They use effective teaching strategies that, together with appropriately targeted support and intervention, match most pupils' individual needs so that pupils learn well across the curriculum. The teaching of reading, writing, communication and mathematics is very efficient. Teachers and other adults enthuse and motivate most pupils to participate. Teaching generally promotes pupils' resilience, confidence and independence when tackling challenging activities. Teachers regularly listen astutely to, carefully observe and skillfully questiived from responses to a list of after school activities (yes/no) (N=229). The marginal effects of the coefficients from an ordered probit regression for increasing/not changing/decreasing help with homework at home are shown, along with 90% confidence intervals. The standard errors used to derive 90% confidence intervals are clustered at the local authority level and derived using the bootstrap method with 1,000 iterations.total enrolment. Column 1 also includes student's eligibility for free lunch, ethnic minority dummy, special education needs st
Requires Improvement (coded 2)	Pupils are progressing at least as well as all pupils nationally given their starting points. Groups of pupils, including disabled pupils and those who have special educational needs, are also making progress in line with similar groups of pupils nationally. Performance is usually at least in line with floor standards. Pupils generally learn well in most subjects, with no major weaknesses. As a result, they are acquiring the knowledge, understanding and skills, including those in reading, writing, communication and mathematics, to ensure that they are prepared adequately for the next stage in their education, training or employment. The standards of attainment of the majority of groups of pupils are likely to be in line with national averages for all pupils. Where standards of groups of pupils are below those of all pupils nationally, the gaps are closing overall. In exceptional circumstances, where attainment, including attainment in reading in primary schools, is low overall, it is improving over a sustaineived from	Teaching results in most pupils, and groups of pupils, currently in the school making progress that is broadly in line with that made by pupils nationally with similar starting points. There is likely to be some good teaching and there are no endemic inadequacies in particular subjects, across year groups or for particular groups of pupils. Teachers' expectations enable most pupils to work hard and achieve satisfactorily and encourage them to make progress. Due attention is often given to the careful assessment of pupils' learning but this is not always conducted rigorously enough and may result in some unnecessary repetition of work for pupils and tasks being planned and set that do not fully challenge. Teachers monitor pupils' work during lessons, picking up any general misconceptions and adjust their plans accordingly to support learning. These adaptations are usually successful but occasionally are not timely or relevant and this slows learning for some pupils. Teaching strategies ensure that the individual needs of pupils are usually met. Teachers carefully deploy any available additional support and set appropriate homework, and these contribute reasonably well to the quality of learning for pupils, including disabled pupils and those who have special educational needs. Pupils are informed about the progress they are making and how to improve further through marking and dialogue with adults that is usually timely and encouraging. This approach ensures that most pupils want to work hard and improve. Communication skills, including reading and writing, and mathematics may be taught inconsistently across the curriculum.
Inadequate (coded 1)	Achievement is likely to be inadequate if any of the following apply: Pupils' learning and progress overall, or the learning and progress of particular groups, is consistently below those of all pupils nationally given their starting point; Learning and progress in any key subject10 or key stage, including the sixth form, lead to underachievement; The learning, quality of work and progress of disabled pupils and those who have special educational needs show that this group is underachieving; Pupils' communication skills, including in reading and writing and proficiency in mathematics overall, or those of particular groups, are not sufficient for the next stage of education or training; Attainment is consistently low, showing little, fragile or inconsistent improvement, or is in decline; There are wide gaps in attainment and in learning and progress between different groups of pupils and of all pupils nationally that are showing little sign of closing or are widening; There are wide gaps in attainment and in learning and progress between different groups of pupils that are barely closing or are widening.	Teaching is likely to be inadequate where any of the following apply: As a result of weak teaching over time, pupils or groups of pupils currently in the school are making inadequate progress; Teachers do not have sufficiently high expectations and teaching over time fails to excite, enthuse, engage or motivate particular groups of pupils, including disabled pupils and those who have special educational needs; Pupils can not communicate, read, write or use mathematics as well as they should; Learning activities are not sufficiently well matched to the needs of pupils so that they make inadequate progress.

Source: OFSTED, The Evaluation Schedule for the Inspection of Maintained School and Academies, April 2012, OFSTED document reference number 090098. Available via the UCL Institute of Education Digital Education Resource Archive: http://dera.ioe.ac.uk/14076/1/The_evaluation_schedule_for_school_inspections_from_January_2012%5B1%5D.pdf

Table A1: Key Ofsted Judgements and Grade Descriptors (cont.)

Part B: Rating Descriptors

Rating	Behaviour and Safety of Pupils at the School	Quality of Leadership in and Management of the School
Outstanding (coded 4)	Parents, carers, staff and pupils are highly positive about behaviour and safety. Pupils make an exceptional contribution to a safe, positive learning environment. They make every effort to ensure that others learn and thrive in an atmosphere of respect and dignity. Pupils show very high levels of engagement, courtesy, collaboration and cooperation in and out of lessons. They have excellent, enthusiastic attitudes to learning, enabling lessons to proceed without interruption. Pupils are consistently punctual in arriving at school and lessons. They are highly adept at managing their own behaviour in the classroom and in social situations, supported by systematic, consistently applied approaches to behaviour management. They are very calm, orderly and considerate when moving around the school. There are excellent improvements in behaviour over time for any individuals or groups with particular behavioural difficulties. Instances of bullying, including for example, cyber-bullying and prejudice-based bullying related to special educational need, sexual orientation, sex, race, religion and belief, gender reassignment or disability, are extremely rare. Pupils are acutely aware of different forms of bullying and actively try to prevent it from occurring. The school has an active and highly effective approach to identifying and tackling bullying. All groups of pupils feel safe at school at all times. They understand very clearly what constitutes unsafe situations and are highly aware of how to keep themselves and others safe. It is likely that attendance will be above average for all groups of pupils or will show sustained and convincing improvement over time.	The pursuit of excellence in all of the school's activities is demonstrated by an uncompromising and highly successful drive to strongly improve achievement, or maintain the highest levels of achievement, for all pupils, including disabled pupils and those who have special educational needs, over a sustained period of time. All leaders and managers, including the governing body, are highly ambitious for the school and lead by example. They base their actions on a deep and accurate understanding of the school's performance and of staff and pupils' skills and attributes. Key leaders focus relentlessly on improving teaching and learning, resulting in teaching that is likely to be outstanding and at least consistently good. The school's curriculum: provides highly positive, memorable experiences and rich opportunities for high quality learning; has a very positive impact on all pupils' behaviour and safety; and contributes very well to pupils' achievement and to their spiritual, moral, social and cultural development. The school has highly successful strategies for engaging with parents and carers to the very obvious benefit of pupils, including those who might traditionally find working with the school difficult. The school's arrangements for safeguarding pupils meet statutory requirements and give no cause for concern.
Good (coded 3)	There are few well founded concerns expressed by parents, carers, staff and pupils about behaviour and safety. Pupils are typically considerate, respectful and courteous to staff and each other and consistently meet the school's expectations. This makes a very positive contribution to a well ordered, safe school. The very large majority of pupils are consistently punctual to school and to lessons. In lessons, pupils demonstrate positive attitudes towards the teacher, their learning and each other. Their good levels of engagement allow lessons to flow smoothly throughout so that disruption is unusual. Pupils, including those with identified behavioural difficulties, respond very well to the school's strategies for managing and improving behaviour, which are applied consistently. Disruptive incidents seldom occur. There are marked improvements in behaviour over time for individuals or groups with particular needs. Instances of bullying, including for example, cyber-bullying and prejudice-based bullying related to special educational need, sexual orientation, sex, race, religion and belief, gender reassignment or disability, are rare. Pupils have a good awareness of different forms of bullying and take active steps to prevent it from occurring. The school swiftly and successfully addresses any incidents of bullying that do occur, thus gaining the full confidence of pupils, parents and carers. Pupils feel safe at school. They understand clearly what constitutes unsafe situations and how to keep themselves safe. Where pupils are able to influence their own attendance, it is likely that attendance will be above average for all sizeable groups of pupils, or showing sustained and convincing improvement over time.	Key leaders and managers, including the governing body, consistently communicate high expectations and ambition. They model good practice and demonstrably work to monitor, improve and support teaching, encouraging the enthusiasm of staff and channeling their efforts and skills to good effect. As a result, teaching is improving and is at least satisfactory, with much that is good. Planned actions based on accurate self-evaluation to overcome weaknesses have been concerted and effective. As a result, achievement has improved or consolidated previous good performance. The school's curriculum provides well organised, imaginative and effective opportunities for learning for all groups of pupils including disabled pupils and those with special educational needs, promotes positive behaviour and safety and provides a broad range of experiences that contribute well to the pupils' achievement and to their spiritual, moral, social and cultural development. The school usually works well with parents and carers, including those who might traditionally find working with the school difficult, to achieve positive benefits for pupils. The school's arrangements for safeguarding pupils meet statutory requirements and give no cause for concern.
Requires Improvement (coded 2)	Parents, carers, pupils and staff are generally positive about behaviour, although some concerns may be raised. Pupils' behaviour and engagement, including their punctuality to school and lessons contributes to a safe and orderly school environment. In lessons, pupils respond promptly to teachers' direction and work cooperatively with each other. Major disruption to learning is uncommon. The school's behaviour management procedures are clear and usually applied but some inconsistencies exist and low-level disruption may occur occasionally. However, it is not endemic in any subject, class or group, or key stage. Pupils, including those with identified behavioural difficulties, are well aware of the school's strategies for managing and improving behaviour; they try hard to respond and improvements over time are evident for individuals and groups, including for those with particular needs. Instances of bullying, including for example, cyber-bullying and prejudice-based bullying related to special educational need, sexual orientation, sex, race, religion and belief, gender reassignment or disability, are infrequent and pupils are aware of different forms of bullying and the importance of preventing them. The school generally deals with any incidents of bullying promptly and effectively, thus gaining the confidence of pupils, parents and carers. Pupils feel safe at school. They know about the main risks they might face and understand how these risks may threaten their own and others' safety. Attendance will usually be at least average but if it is below average, for all pupils or particular groups, it will be improving over time.	The head teacher and most other key leaders, including the governing body, provide a concerted approach to school improvement. Planned actions by leaders and managers have improved the quality of teaching so that very little is inadequate. Most, but not all, staff and pupils are fully committed to the drive and ambition demonstrated by key leaders. Capacity to improve is demonstrated by a trend of sustained improvement in achievement, behaviour and safety, although a few significant weaknesses remain. Essential systems are embedded sufficiently to enable the school to continue improving and do not depend solely on only one or two senior leaders. The curriculum is generally matched to pupils' needs, interests and aspirations and provides adequate preparation for the next stage of their lives, whatever their starting points. The school usually works well with parents and carers, although may be less successful in engaging those who might traditionally find working with the school difficult. The school's arrangements for safeguarding pupils meet statutory requirements and give no cause for concern.
Inadequate (coded 1)	Behaviour and safety are likely to be inadequate when any of the following apply: Parents, carers, pupils or staff raise major and/or well founded concerns about behaviour that are not being addressed; Pupils' lack of engagement and persistent low-level disruption contribute more than occasionally to reduced learning and/or a disorderly classroom environment; A significant minority of pupils show a lack of respect and intolerance for each other or staff and a lack of self-discipline, resulting in poor behaviour around the school; Incidents of bullying overall or specific types of bullying, including for example, cyber-bullying and prejudice-based bullying related to special educational need, sexual orientation, sex, race, religion and belief, gender reassignment or disability, are frequent or pupils have little confidence in the school's ability to address bullying successfully; Pupils or specific groups of pupils do not feel safe; Attendance is consistently low for all pupils or groups of pupils and shows little or no sign of improvement.	Leadership and management are likely to be inadequate if any of the following apply: Capacity for further improvement is limited because current leaders and managers have been ineffective in securing essential improvements since the last inspection; Leaders and managers are not taking effective steps to secure satisfactory and better teaching for all groups of pupils, including disabled pupils and those who have special educational needs; The curriculum fails to meet the needs of pupils or particular groups of pupils; Despite remedying a few small areas of weakness, perhaps recently, improvements are fragile, too slow or depend on external support; The school's strategies for engaging with parents and carers are weak so that parents and carers are not involved sufficiently in supporting their children's learning and development; The school's arrangements for safeguarding pupils do not meet statutory requirements and give serious cause for concern.

Source: OFSTED, The Evaluation Schedule for the Inspection of Maintained School and Academies, April 2012, OFSTED document reference number 090098. Available via the UCL Institute of Education Digital Education Resource Archive: http://dera.ioe.ac.uk/14076/1/The_evaluation_schedule_for_school_inspections_from_January_2012%5B1%5D.pdf

Table A1: Key Ofsted Judgements and Grade Descriptors (cont.)

Part B: Rating Descriptors

For Overall Effectiveness, inspectors must consider the evidence gathered in support of their evaluations of the four key judgements.

Rating	Overall Effectiveness
Outstanding (coded 4)	The school's practice consistently reflects the highest aspirations for pupils and expectations of staff. It ensures that best practice is spread effectively in a drive for continuous improvement. Teaching is likely to be outstanding and together with a rich curriculum, which is highly relevant to pupils' needs, it contributes to outstanding learning and achievement or, in exceptional circumstances, achievement that is good and rapidly improving. Other principal aspects of the school's work are good or outstanding. The school's thoughtful and wide ranging promotion of the pupils' spiritual, moral, social and cultural development enables them to thrive in a supportive, highly cohesive learning community. Consequently, pupils and groups of pupils have excellent experiences at school, ensuring that they are very well equipped for the next stage of their education, training or employment.
Good (coded 3)	The school takes effective action to enable most pupils to reach their potential. Pupils benefit from teaching that is at least good. This promotes very positive attitudes to learning and ensures that achievement is at least good. Leadership and management play a significant role in this and are good overall. Behaviour and safety are strong features. Deliberate and effective action is taken to create a cohesive learning community by promoting the pupils' spiritual, moral, social and cultural development. A positive climate for learning exists and pupils and groups of pupils have highly positive experiences at school so that they are well prepared for the next stage in their education, training or employment.
Requires Improvement (coded 2)	Achievement, behaviour and safety, the quality of teaching and learning, and leadership and management are all likely to be at least satisfactory with some significant good practice. In addition, the school takes reasonable steps to promote pupils' spiritual, moral, social and cultural development, enabling them to develop the skills and personal qualities needed to work together in a generally cohesive learning community. As a result, pupils and groups of pupils have a generally positive experience at school and are not disadvantaged as they move to the next stage of their education, training or employment.
Inadequate (coded 1)	Overall effectiveness is likely to be inadequate if any of the following apply: Achievement is inadequate; Quality of teaching is inadequate; Behaviour and safety are inadequate; Leadership and management are inadequate. There are important weaknesses in the school's promotion of pupils' spiritual, moral, social and cultural development, resulting in a poor climate for learning and an incohesive school community where pupils or groups of pupils are unable to thrive.

Source: OFSTED, The Evaluation Schedule for the Inspection of Maintained School and Academies, April 2012, OFSTED document reference number 090098. Available via the UCL Institute of Education Digital Education Resource Archive: http://dera.ioe.ac.uk/14076/1/The_evaluation_schedule_for_school_inspections_from_January_2012%5B1%5D.pdf

Table A2: Sample Selection of Households

Means, Standard Deviation in Parentheses
Pooling Across Survey Waves 1, 3 and 5

	(1) Children Aged 10-15 (England)	(2) Homework Variable Can be Constructed Across Waves	(3) Non-missing School Code	(4) School Inspected in Academic Year of Interview	(5) Final Sample
Sample Size (children):	14,092	4,661	2,899	747	690
Household Characteristics					
Household Size	4.51	4.46	4.21	4.19	4.16
	(1.47)	(1.40)	(1.32)	(1.32)	(1.30)
Home Owner	.611	.644	.633	.620	.633
Mother characteristics					
Married/cohabiting	.758	.753	.728	.716	.718
White Ethnicity	.702	.721	.724	.741	.743
Education GCSE or Below	.458	.432	.439	.455	.440
Father characteristics					
Married/cohabiting	.972	.969	.960	.953	.955
White Ethnicity	.703	.726	.741	.769	.770
Education GCSE or Below	.418	.379	.404	.431	.435

Notes: Column 1 is based on the initial sample of UKHLS households with children aged 10-15 observed at waves 1, 3 or 5. Columns 2 is restricted to those households in which the parental help with homework variable is observed at two consecutive times. Column 3 is further restricted to those that also have a non-missing school code. Column 4 is further restricted to those whose school was Ofsted inspected in the academic year of observation. Column 5 is further restricted by dropping those whose household interview was on the same day as the school inspection or with missing predicted inspection grades (mostly new Academy schools with missing past Ofsted grade). This is our final sample used for the main analysis.

Table A3: Sample Selection of Schools

Means, Standard Deviation in Parentheses

	School Inspected by Ofsted		School Not Inspected by Ofsted	
	(1) Wave 3	(2) Wave 5	(4) Wave 3	(5) Wave 5
	Jan 2011- Dec 2012	Jan 2013-Dec 2014	Reference Year: 2011/12	Reference Year: 2013/14
Number of Schools	2,102	2,437	1,686	1,438
School composition				
School Size: Number of Pupils	815.2	788.5	887.2	873.3
	(494.9)	(488.2)	(490.2)	(504.4)
% Pupils Free School Meals	21.85	22.43	18.23	18.44
	(15.68)	(15.18)	(15.19)	(15.02)
% Pupils English as an Additional Language	12.13	13.71	13.24	13.76
	(18.13)	(19.36)	(19.17)	(18.77)
School type				
Academy School	.244	.129	.375	.111
Community School	.284	.197	.233	.132
Voluntary Aided or Controlled School	.115	.090	.130	.092
Foundation School	.143	.093	.097	.054
Special School	.214	.190	.163	.131
School performance				
% Pupils 5 or More A*-C grades incl. English & Maths	.462	.438	.539	.542
	(.255)	(.246)	(.273)	(.284)
% Pupils 5 or More A*-C grades	.680	.513	.737	.617
	(.332)	(.275)	(.316)	(.296)
% Pupils 5 or More A*-G grades	.812	.781	.850	.831
	(.341)	(.349)	(.316)	(.323)
% Pupils with Entries in all English Baccalaureate Subjects	.106	.167	.177	.266
	(.134)	(.150)	(.206)	(.230)
Total Average Point Score	291.9	256.1	313.7	287.3
	(110.7)	(110.1)	(106.9)	(113.5)

Notes: Columns 1 and 2 show the number and characteristics of secondary schools that were inspected at some point during the two-year survey periods of Waves 3 and 5 of Understanding Society. Each two-year survey period covers all or part of three academic years, with academic years running from September of one calendar year to August of the next year. Columns 3 and 4 show the numbers and characteristics of non-inspected secondary schools during survey waves 3 and 5 respectively. For the non-inspected schools we define a reference year that falls within the survey period.

Table A4: Predicting Ofsted Inspection Ratings

Dependent Variable: Ofsted Grade

Linear Regression, Standard Errors Clustered by Local Authority

	(1) AR (1)	(2) School Characteristics	(3) School Performance	(4) Progress in English & Maths	(5) GCSE Performance	(6) All Performance Measures	(7) Sample Schools	(8) Full Model, Ordered Probit
Last Ofsted grade	.439***	.336***	.331***	.306***	.273***	.268***	.121***	.475***
	(.014)	(.016)	(.016)	(.016)	(.015)	(.015)	(.037)	(.026)
School Size: Number of Pupils		.000**	-.000	-.000***	.000**	-.000*	.000	-.000**
		(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
% Pupils FSM		-.013***	-.009***	-.002	.003*	.002	.003	.004**
		(.001)	(.001)	(.002)	(.002)	(.002)	(.004)	(.002)
% Pupils EAL		.004***	.003***	.001	.003***	.001	.003	.003**
		(.001)	(.001)	(.001)	(.001)	(.001)	(.002)	(.001)
Academy School		-.555***	-1.140***	-1.149***	-.793***	-.824***	-.635***	-1.38***
		(.058)	(.102)	(.096)	(.096)	(.099)	(.238)	(.178)
Community School		-.700***	-1.255***	-1.258***	-.897***	-.926***	-.565**	-1.54***
		(.055)	(.092)	(.088)	(.091)	(.0940)	(.242)	(.168)
Voluntary Aided or Controlled School		-.504***	-1.037***	-1.065***	-.790***	-.806***	-.576**	-1.375***
		(.069)	(.107)	(.105)	(.098)	(.102)	(.254)	(.182)
Foundation School		-.619***	-1.179***	-1.180***	-.836***	-.855***	-.632**	-1.415***
		(.063)	(.100)	(.096)	(.096)	(.099)	(.251)	(.179)
Has Sixth Form		-.001	-.039	-.044	-.016	-.020	.015	-.040
		(.036)	(.034)	(.034)	(.026)	(.027)	(.067)	(.039)
Christian Denomination		-.090*	-.129**	-.134**	-.074	-.080*	.045	-.112
		(.052)	(.054)	(.053)	(.048)	(.047)	(.091)	(.080)
Other Religious Denomination		-.671**	-.738**	-.805***	-.721***	-.666**	-1.417***	-1.169***
		(.306)	(.303)	(.296)	(.252)	(.269)	(.164)	(.447)
Mixed Gender School		-.299***	-.238***	-.100**	-.008	-.0552	-.014	-.132
		(.055)	(.052)	(.047)	(.048)	(.048)	(.086)	(.081)
Boys School		-.324***	-.318***	-.276***	-.131**	-.144**	-.116	-.244**
		(.0694)	(.067)	(.063)	(.063)	(.059)	(.126)	(.108)
% Pupils 5 or More A*-C Grades			1.087***		.375***	.389***	1.715***	.555**
			(.108)		(.140)	(.141)	(.244)	(.230)
% Pupils 5 or More A*-G Grades					-1.871***	-2.384***	-3.880***	-3.972***
					(.165)	(.188)	(.387)	(.315)
% Pupils 5 or More A*-C Grades incl. English & Maths					2.381***	1.135***	1.053**	1.730***
					(.198)	(.251)	(.482)	(.406)
Total Average Point Score					.002***	.002***	.002***	.004***
					(.000)	(.000)	(.000)	(.000)
% making expected progress in English				.883***		.869***	1.427***	1.546***
				(.119)		(.143)	(.264)	(.239)
% Making Expected Progress in Maths				.853***		1.101***	1.588***	1.907***
				(.114)		(.169)	(.392)	(.279)
% Pupils Achieving English Baccalaureate				.346***		-.359***	-.862***	-.330
				(.122)		(.108)	(.241)	(.202)
Null: β_0=1 F-statistic [p-value]	1537 [.000]	1813 [.000]	1834 [.000]	1812 [.000]	2281 [.000]	2270 [.000]	558.6 [.000]	Chi2: 418.4 [.000]
Adjusted R-squared	.177	.206	.234	.275	.346	.360	.423	-
Number of LEAs	-	151	151	151	151	151	138	151
LEA Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	Yes	No
Number of Observations	4419	4419	4419	4419	4419	4419	826	4419
Number of Schools	3113	3113	3113	3113	3113	3113	548	3113

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The sample in Columns 1 to 6 comprises schools that are inspected during the survey period covering waves 1, 3 and 5 of UKHLS. In Column 7 we restrict the sample to those schools in which our final sample of UKHLS households are interviewed. Hence, the unit of observation is a school. The outcome variable is the OFSTED inspection grade in period t (the last time the school was inspected), where this can take the following values: 4 (Outstanding), 3 (Good), 2 (Requires Improvement) and 1 (Inadequate/failing). Column 1 only controls for the past Ofsted inspection grade. Column 2 additionally controls for schools characteristics, Column 3 controls for school characteristics plus key school performance, Column 4 controls for school characteristics plus measures of progress in English and Maths, Column 5 controls for school performance plus detailed GCSE performance measures, Columns 6 to 8 control for school characteristics plus all performance measures. Column 1 is based on a linear regression with standard errors clustered at the local authority level, Columns 2 to 7 are based on linear regression with local authority fixed effects and standard errors clustered by local authority, and Column 8 is based on an ordered probit regression with standard errors clustered at the local authority level. The omitted category for school type dummies is 'Special and other schools'. All controls refer to the previous academic year as that of inspection, except for last Ofsted grade which refers to the year of the last inspection. At the foot of each Column we report the F-statistic and p-value on the null hypothesis that the coefficient on the last Ofsted grade is equal to one.

Table A5: Balance, by News Shock

Means, Standard Deviation in Parentheses, p-values in Brackets

	Good News			No News			Bad News		
	(1) Treated: Interviewed After Ofsted Inspection	(2) Control: Interviewed Before Inspection	(3) Test of Equality [p-value]	(4) Treated: Interviewed After Ofsted Inspection	(5) Control: Interviewed Before Inspection	(6) Test of Equality [p-value]	(7) Treated: Interviewed After Ofsted Inspection	(8) Control: Interviewed Before Inspection	(9) Test of Equality [p-value]
Number of children	107	68		218	175		77	45	
A. School Characteristics									
School Size: Number of Pupils	1125.3	1114.6	[.853]	1142.7	1089.7	[.259]	1089.1	1102.4	[.864]
	(366.8)	(349.5)		(369.4)	(367.6)		(387.1)	(375.9)	
% Pupils Free School Meals	16.59	19.27	[.472]	18.24	16.52	[.170]	16.84	16.34	[.816]
	(11.12)	(17.67)		(14.06)	(13.68)		(11.69)	(10.99)	
Academy School	.159	.412	[.002]	.261	.234	[.499]	.169	.311	[.081]
Boys School	.037	.000	[.105]	.055	.040	[.474]	.065	.067	[.975]
% Pupils 5 or More A*-C grades	.817	.822	[.826]	.775	.777	[.929]	.701	.727	[.411]
	(.139)	(.128)		(.165)	(.157)		(.229)	(.152)	
Total Average GCSE Point Score	339.5	342.5	[.588]	331.7	333.6	[.710]	318.2	326.4	[.361]
	(39.62)	(25.35)		(46.63)	(44.52)		(64.83)	(38.75)	
B. Household Characteristics									
Household Size	3.97	4.26	[.110]	4.24	4.14	[.472]	4.06	4.33	[.338]
	(1.068)	(1.192)		(1.400)	(1.325)		(1.33)	(1.28)	
Home Owner	.626	.559	[.428]	.633	.686	[.300]	0.64	0.64	[.926]
C. Child characteristics									
Female	.551	.456	[.192]	.528	.491	[.509]	.455	.444	[.913]
Age	13.56	13.54	[.918]	13.39	13.41	[.915]	13.51	13.20	[.172]
	(1.183)	(1.152)		(1.061)	(1.145)		(1.034)	(1.079)	
D. Mother characteristics									
Married/cohabiting	.709	.651	[.428]	1.06	1.15	[.711]	.662	.721	[.511]
White Ethnicity	.777	.730	[.559]	.715	.777	[.178]	.784	.605	[.062]
Education GCSE or Below	.417	.444	[.693]	.435	.427	[.859]	.446	.558	[.295]
E. Father characteristics									
Married/cohabiting	.925	1.000	[.047]	.981	.900	[.022]	1.000	1.000	.
White Ethnicity	.736	.788	[.598]	.788	.822	[.575]	.647	.682	[.774]
Education GCSE or Below	.585	.364	[.069]	.423	.400	[.719]	.412	.409	[.986]

Notes: The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. For schools that receive a positive news shock, Columns 1 and 2 show means and standard deviations in parentheses for treated and control households respectively. Column 3 shows the p-values on the test of equality of the mean, derived by regressing the characteristic on a treatment dummy and clustering standard errors by local authority. The remaining Columns show the same information among those schools that receive no news and those that receive a negative news shock.

Table A6: Parental Response to Information on School Quality

Ordered Probit Regression Estimates
Bootstrapped Standard Errors in Parentheses, Clustered by Local Authority

	(1) Forecast, Unconditional	(2) Plus Child Characteristics	(3) Plus Parent Characteristics	(4) Plus School Characteristics
Treated (β_0)	-.034	-.047	-.037	-.036
	(.121)	(.124)	(.114)	(.112)
Treated x Good news (β_1)	-.464*	-.442*	-.458*	-.482**
	(.239)	(.243)	(.241)	(.243)
Treated x Bad news (β_2)	.075	.116	.114	.145
	(.254)	(.258)	(.243)	(.244)
Good news (δ_1)	.012	-.014	.013	.018
	(.184)	(.187)	(.184)	(.185)
Bad news (δ_2)	-.073	-.109	-.095	-.128
	(.215)	(.222)	(.215)	(.215)
T-C | good news ($\beta_0+\beta_1$)	-.498**	-.488**	-.495**	-.518**
	(.196)	(.196)	(.207)	(.207)
T-C | bad news ($\beta_0+\beta_2$)	.041	.069	.078	.109
	(.228)	(.234)	(.218)	(.219)
Diff-in-Diff-in-Diff ($\beta_1-\beta_2$)	-.539*	-.558*	-.573*	-.627**
	(.310)	(.314)	(.303)	(.305)
Forecast Ofsted Rating	Yes	Yes	Yes	Yes
Child Characteristics	No	Yes	Yes	Yes
Parent Characteristics	No	No	Yes	Yes
School Characteristics	No	No	No	Yes
Observations	690	690	690	690

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. The outcome variable is the change in parental help with homework between survey wave 3 and 1, or between survey wave 5 and 3: this is coded as -1 if there is a decrease in parental help, 0 if there is no change in parental help and +1 is there an increase in parental help with homework. Ordered probit regression estimates are shown. In all Columns, the specification uses the predicted news shock. In Column 1 we control for a treatment dummy, interactions between the treatment dummy and dummies for whether a positive or negative news shock is observed, and the dummies for a positive or negative news shock. Column 2 additionally controls for child and household characteristics (gender and age dummies, household size, number of children in the household and dummies for housing tenure), Column 3 additionally controls for parental characteristics (ethnicity, highest educational degree and marital status), and Column 4 additionally controls for school characteristics (school size and proportion of children eligible for free school meals). Standard errors are derived using the bootstrap method with 1,000 iterations, clustered at the local authority level and shown in parentheses.

Table A7: Further Support for Identifying Assumptions and Robustness Checks

Ordered Probit Regression Estimates

Bootstrapped Standard Errors in Parentheses, Clustered by Local Authority

	(1) Controlling for Month of Interview	(2) Placebo: Next Year's Inspections	(3) Drop Failing Schools	(4) Drop Failing & Outstanding Schools	(5) Omitting Obs 2 Weeks Post-Inspection	(6) Omitting Obs 3 Weeks Post-Inspection	(7) Omitting Obs 4 Weeks Post-Inspection	(8) Controlling for Full Set of school Characteristics	(9) Controlling for Previous Ofsted Rating	(10) Dropping 12 Year Olds
T-C | good news ($\beta_0+\beta_1$)	-.596***	.015	-.524**	-.552*	-.578***	-.583***	-.590***	-.509**	-.517**	-.468*
	(.213)	(.159)	(.209)	(.291)	(.201)	(.195)	(.195)	(.212)	(.208)	(.256)
T-C | bad news ($\beta_0+\beta_2$)	.023	.111	.076	.059	.155	.141	.116	.122	.110	.187
	(.258)	(.206)	(.276)	(.282)	(.221)	(.231)	(.232)	(.227)	(.219)	(.269)
Diff-in-Diff-in-Diff ($\beta_1-\beta_2$)	-.619*	-.096	-.600*	-.610	-.733**	-.724**	-.705**	-.631**	-.628**	-.655*
	(.329)	(.258)	(.340)	(.372)	(.307)	(.307)	(.313)	(.314)	(.305)	(.377)
F-test: month dummies [p-value]	3.91 [.951]									
Forecast Ofsted Rating	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Child Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parent Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	690	685	639	525	655	653	648	690	690	503

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. The outcome variable is the change in parental help with homework between survey wave 3 and 1, or between survey wave 5 and 3: this is coded as -1 if there is a decrease in parental help, 0 if there is no change in parental help and +1 is there an increase in parental help with homework. Ordered probit regression estimates are shown. In all Columns, the specification uses a predicted news shock. All specifications are as in the baseline with some modification. Column 1 additionally controls for month of household interview. The foot of this Column reports the F-statistic (and p-value) on the null that these month of interview dummies are jointly insignificant. The sample in Column 2 is based on UKHLS households with an Ofsted school inspection in the following academic year of interview and a non-missing change in help with homework. Schools are then assigned next year's Ofsted rating to generate the placebo news shock. Column 3 drops failing schools (those ranked as 1=failing by Ofsted in year t). Columns 4 to 6 omit households interviewed within 2, 3 and 4 weeks of an Ofsted inspection. Column 7 additionally controls for a full set of school characteristics. Column 8 additionally controls for the previous Ofsted rating. Column 9 drops 12-year-olds (so drops children that are most likely to have changed schools across survey waves). Standard errors are derived using the bootstrap method with 1,000 iterations, clustered at the local authority level and shown in parentheses.

Table A8: Short Run School Responses to Information on School Quality

Linear Regression Estimates

Bootstrapped Standard Errors in Parentheses, Clustered by Local Authority

	(1) Hours Children Expected to do Homework	(2) Number of Class Support Teachers	(3) Days Used Supply Teachers	(4) Time Teaching Numeracy and Literacy	(5) % Time in Numeracy and Literacy Groups	(6) % Time Numeracy and Literacy with Individuals
T-C | good news ($\beta_0+\beta_1$)	-6.771	-0.151	-1.924	-0.377	1.516	-0.199
	(8.838)	(.193)	(1.517)	(.496)	(1.850)	(3.598)
T-C | bad news ($\beta_0+\beta_2$)	-11.901	0.223	-1.62	0.357	0.9	1.089
	(7.848)	(.210)	(1.538)	(.484)	(2.133)	(4.098)
Diff-in-Diff-in-Diff ($\beta_1-\beta_2$)	5.13	-0.375	-0.304	-0.734	0.615	-1.288
	(9.861)	(.253)	(2.032)	(.637)	(2.578)	(4.906)
Forecast Ofsted Rating	Yes	Yes	Yes	Yes	Yes	Yes
School Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,198	1,132	1,138	893	1,081	1,076

	(7) Use Streaming	(8) Use Sets: Literacy	(9) Use Sets: Numeracy	(10) Use Ability Grouping	(11) Use Subject Groups: Literacy	(12) Use Subject Groups: Numeracy
T-C | good news ($\beta_0+\beta_1$)	0.088	0.182	0.152	-0.161*	-0.061	-0.098
	(.100)	(.121)	(.128)	(.096)	(.073)	(.077)
T-C | bad news ($\beta_0+\beta_2$)	0.08	0.097	0.085	-0.004	0.007	-0.04
	(.106)	(.126)	(.135)	(.101)	(.088)	(.089)
Diff-in-Diff-in-Diff ($\beta_1-\beta_2$)	0.008	0.085	0.067	-0.157	-0.068	-0.058
	(.131)	(.161)	(.165)	(.128)	(.107)	(.104)
Forecast Ofsted Rating	Yes	Yes	Yes	Yes	Yes	Yes
School Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,258	1,229	1,218	1,277	1,277	1,275

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The sample comprises teachers of MCS children whose school had an Ofsted inspection during the academic year of the survey (academic years 2007/08 and 2008/09). Treated (control) teachers are defined as those whose teacher survey was filled in after (before) the date of OFSTED inspection. Outcome variables are teaching practices used in schools of MCS children. Linear regression estimates are shown. Column 4 is teaching time in hours/week, Columns 5 and 6 are the proportion of teaching time in numeracy and literacy devoted to the whole class, groups and individuals. Streaming in Column 7 is the practice of dividing a class into hierarchical ability groups according to overall ability. Setting (Columns 8 and 9) is ability grouping by subject. School-level controls include school size, the proportion of pupils on free lunches, school type and religious denomination. Robust standard errors are derived using the bootstrap method with 1,000 iterations and shown in parentheses.

Table A9: LPM Estimates of the Parental Response to Information on School Quality

Linear Regressions, Standard Errors Clustered by Local Authority

	Panel A: Frequency of Help Increases (0 / 1)			Panel B: Frequency of Help Decreases (0 / 1)		
	(1) Child Characteristics	(2) Plus Parent Characteristics	(3) Plus School Characteristics	(4) Child Characteristics	(5) Plus Parent Characteristics	(6) Plus School Characteristics
T-C | good news ($\beta_0+\beta_1$)	-.146**	-.146**	-.148**	.146**	.147**	.155**
	(.059)	(.064)	(.064)	(.075)	(.076)	(.075)
T-C | bad news ($\beta_0+\beta_2$)	-.057	-.055	-.049	-.094	-.098	-.111
	(.078)	(.077)	(.077)	(.091)	(.085)	(.085)
Diff-in-Diff-in-Diff ($\beta_1-\beta_2$)	-.089	-.091	-.100	.240*	.246**	.266**
	(.100)	(.102)	(.104)	(.125)	(.119)	(.120)
Forecast Ofsted Rating	Yes	Yes	Yes	Yes	Yes	Yes
Child Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Parent Characteristics	No	Yes	Yes	No	Yes	Yes
School Characteristics	No	No	Yes	No	No	Yes
Adjusted R-squared	.012	.012	.012	.031	.034	.041
Observations	690	690	690	690	690	690

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. The outcome variable is the change in parental help with homework between survey wave 3 and 1, or between survey wave 5 and 3: this is coded as -1 if there is a decrease in parental help, 0 if there is no change in parental help and +1 is there an increase in parental help with homework. The outcome variables are binary indicators for increasing (Panel A) and decreasing (Panel B) parental help with homework between survey wave 3 compared to 1 and 5 compared to 3 respectively. Linear probability estimates are shown. In all Columns, the specification uses the predicted news shock. In Column (1) we control for a treatment dummy, interactions between the treatment dummy and dummies for whether a positive or negative news shock is observed, and the dummies for a positive or negative news shock. Columns 1 and 4 additionally control for child and household characteristics (gender and age dummies, household size, number of children in the household and dummies for housing tenure). Columns 2 and 5 additionally control for parental characteristics (ethnicity, highest educational degree and marital status). Columns 3 and 6 additionally control for school characteristics (school size and proportion of children eligible for free school meals). Standard errors are derived using the bootstrap method with 1,000 iterations, clustered at the local authority level and shown in parentheses.

Table A10: Alternative Forecasting Models

Ordered Probit Regression Estimates

Bootstrapped Standard Errors in Parentheses, Clustered by Local Authority

	(1) AR(1)	(2) School Characteristics	(3) Full Model	(4) Full Model: Ordered Probit in First Stage	(5) Naïve Model
T-C | good news ($\beta_0+\beta_1$)	-.407**	-.477**	-.518**	-.520**	-.407**
	(.180)	(.188)	(.207)	(.245)	(.187)
T-C | bad news ($\beta_0+\beta_2$)	.117	.067	.109	.076	.268
	(.227)	(.227)	(.219)	(.210)	(.181)
Diff-in-Diff-in-Diff ($\beta_1-\beta_2$)	-.525*	-.544*	-.627**	-.596*	-.675**
	(.306)	(.311)	(.305)	(.326)	(.276)
Forecast Ofsted Rating	Yes	Yes	Yes	Yes	No
Child Characteristics	Yes	Yes	Yes	Yes	Yes
Parent Characteristics	Yes	Yes	Yes	Yes	Yes
School Characteristics	Yes	Yes	Yes	Yes	Yes
Observations	690	690	690	690	690

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. The outcome variable is the change in parental help with homework between survey wave 3 and 1, or between survey wave 5 and 3: this is coded as -1 if there is a decrease in parental help, 0 if there is no change in parental help and +1 is there an increase in parental help with homework. Ordered probit regression estimates are shown. The Columns vary in the underlying forecasting model used. Column 1 uses a model to forecast the Ofsted rating that controls for the past Ofsted inspection grade. Column 2 additionally controls for schools characteristics, Columns 3 and 4 additionally control for school characteristics plus all performance measures. Column 5 uses the past Ofsted grade as the forecast grade to derive the news shock variable. The forecasting models used in Columns 1 to 3 are based on linear regressions, while the forecasting model used in Column 4 is an ordered probit regression. In all Columns we control for a treatment dummy, interactions between the treatment dummy and dummies for whether a positive or negative news shock is observed, and the dummies for a positive or negative news shock, controls for child and household characteristics (gender and age dummies, household size, number of children in the household and dummies for housing tenure), controls for parental characteristics (ethnicity, highest educational degree and marital status), and controls for school characteristics (school size and proportion of children eligible for free school meals). Standard errors are derived using the bootstrap method with 1,000 iterations, clustered at the local authority level and shown in parentheses.

Figure A1: Example Ofsted Letter to Parents

This letter is provided for the school, parents and carers to share with their children. It describes Ofsted's main findings from the inspection of their school.

Ofsted
raising standards
improving lives

21 November 2011

Dear Students

Inspection of [redacted], [redacted]

We very much enjoyed the two days we spent in your school recently and wish we had had time to talk to more of you. [redacted] was nearly outstanding when it was inspected four years ago, and now it is! GCSE results have risen dramatically and A-level results are also better than they were. You clearly work hard and respond well to the good teaching, and many of you spoke appreciatively about the extra time teachers give you to help you do well. You receive outstanding care and support.

You admitted (some of you slightly unenthusiastically) that the recent focus on behaviour, attendance and punctuality has been welcome. Your behaviour is very good and your attendance is now well above the national average. If you continue to adhere to the high standards expected, then there is no reason why you should not achieve even better results.

The school is already focused on increasing the number of top GCSE and A-level grades; we support this by making it one of the issues for improvement. The other one relates to the limited amount of physical education in Year 10 upwards and the fact that too few of you, especially in the sixth form, join in extra-curricular sports.

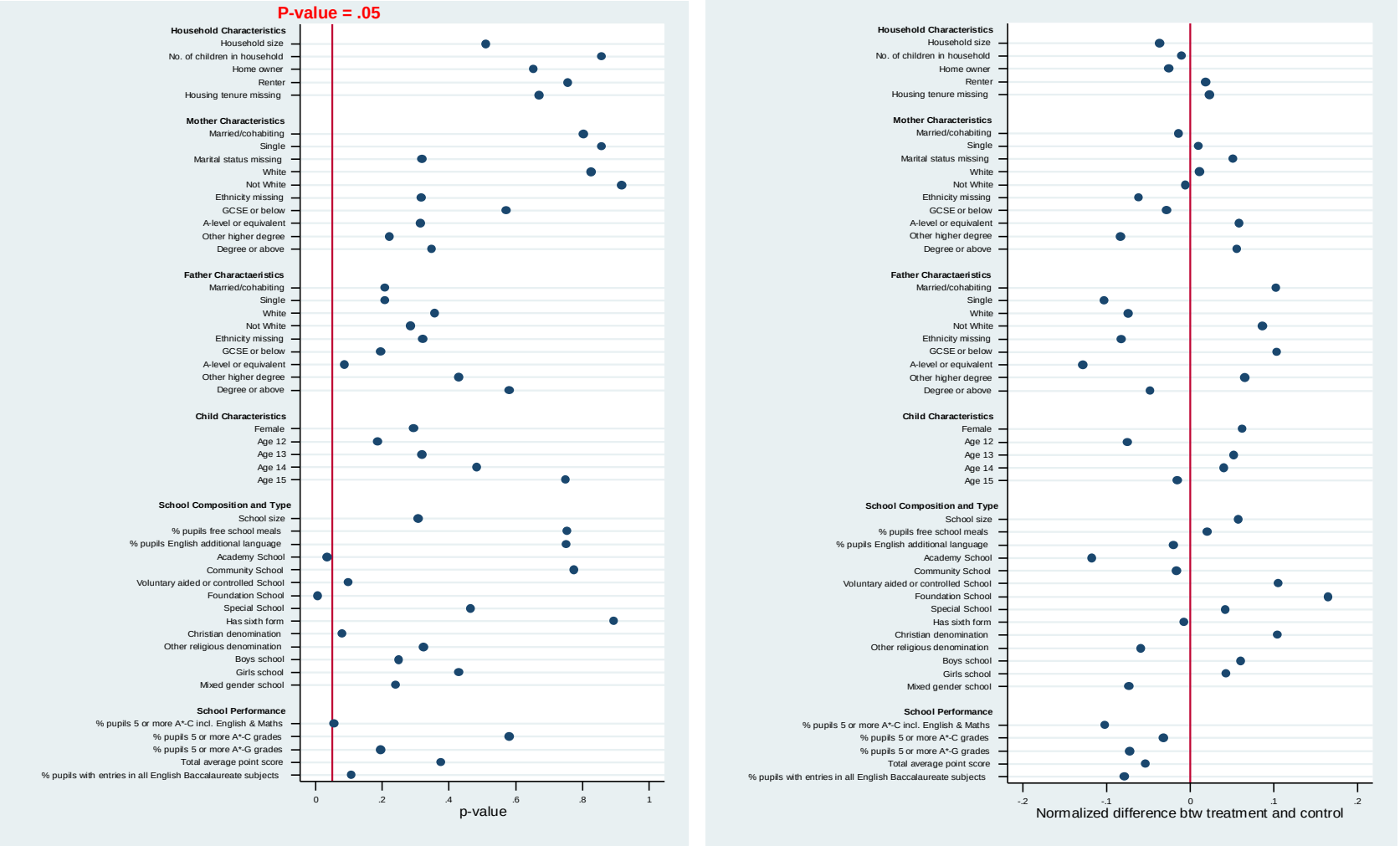
Many of you will read the full report and notice other small things we mention that could be even better. We spent some time investigating the comments that your parents and carers sent us and realise that home/school relationships are not yet perfect!

Enjoy the festive season when it arrives and we hope you all have successful and enjoyable futures.

Yours sincerely

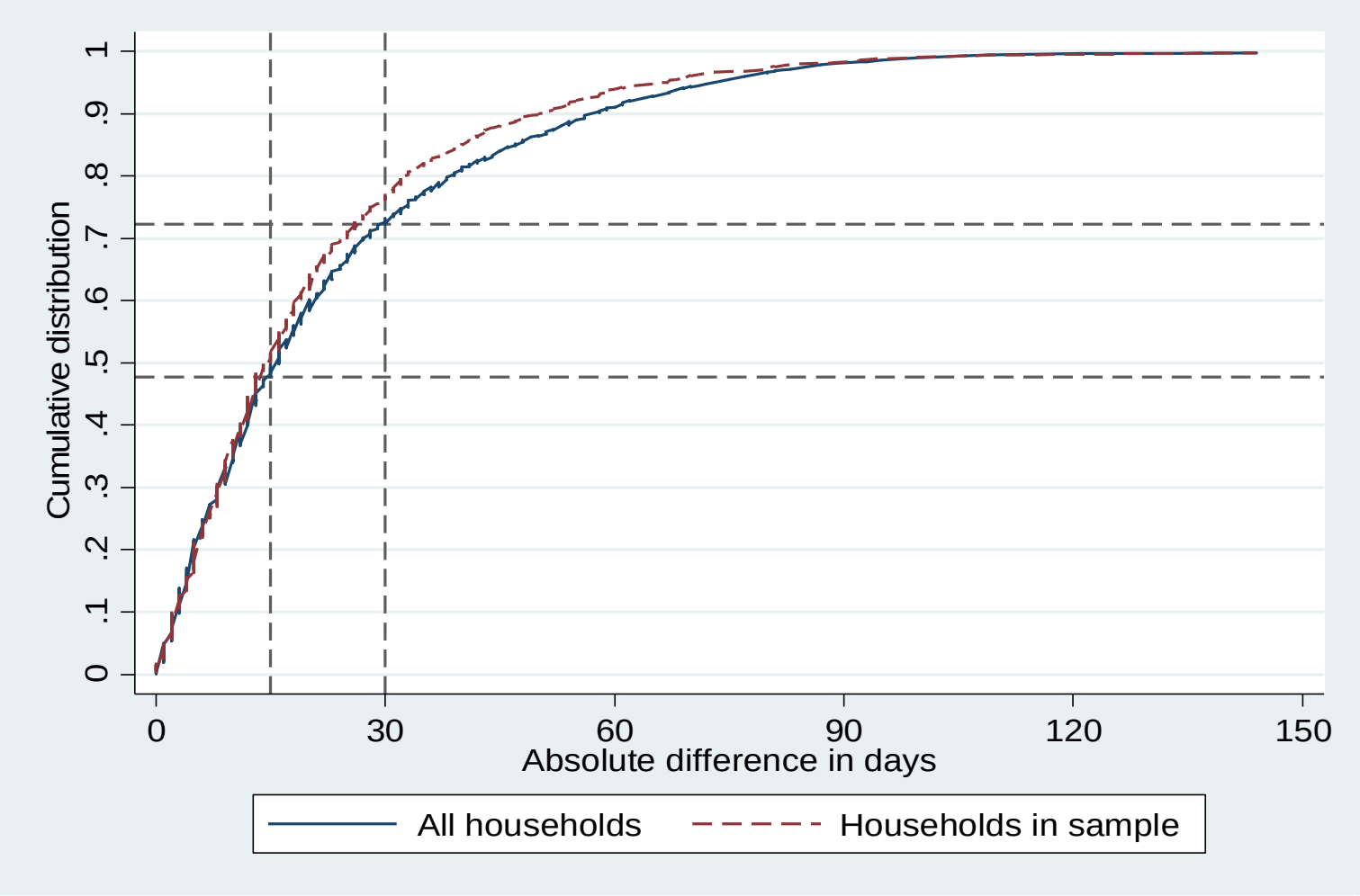
[redacted]
Lead inspector

Figure A2: Balance and Normalized Differences



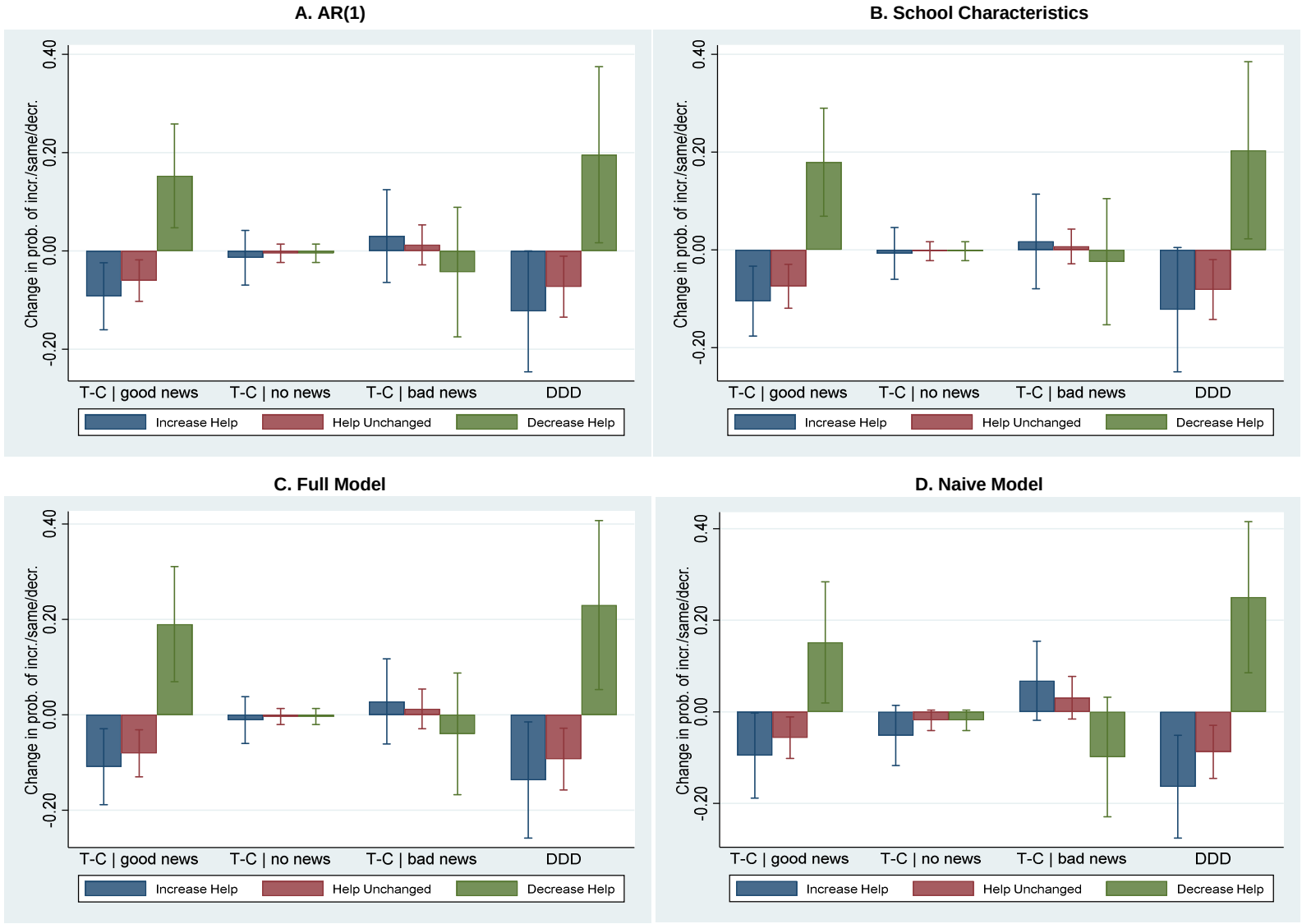
Notes: The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. The left hand panel shows p-values for test on the equality of household, child, mother, father and school characteristics across treated and control households. These are derived by regressing characteristics on treatment dummy and clustering standard errors by local authority. The vertical line indicates a p-value of 0.05. The right hand panel rows displays normalized difference of the means of household, child, mother, father and school characteristics between treatment and control groups, derived by dividing the raw mean difference by the square root of the sum of the variances.

Figure A3: Absolute Difference in Interview Dates Between Survey Waves



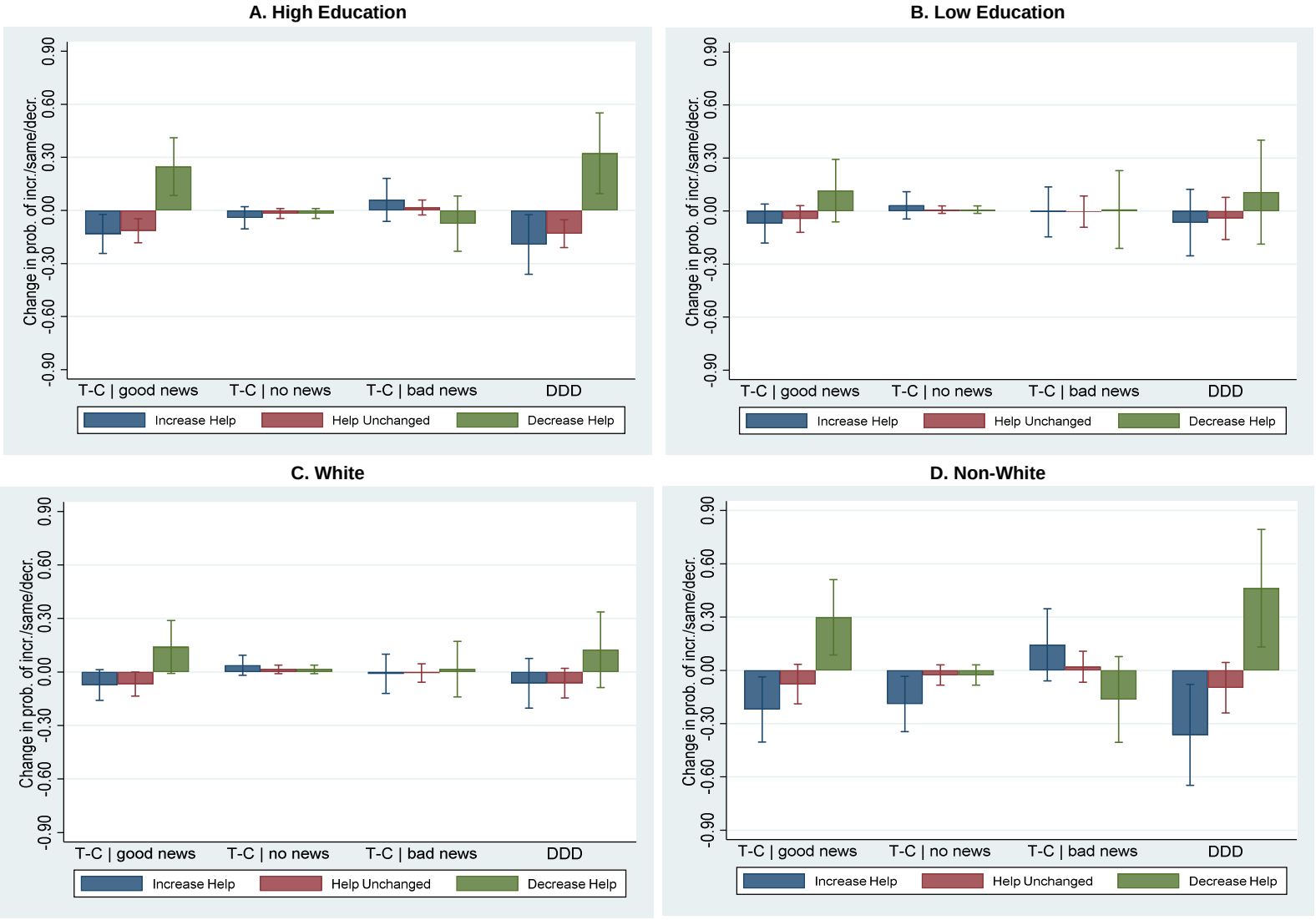
Notes: Two household samples are used. The first are all those observed in the UKHLS across consecutive waves in waves 1, 3 and 5 (N=4661). The second sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question (N=690). The Figure shows the cumulative distribution in the absolute difference in date of interview at wave t and t-2. Vertical lines are marked at 15 and 30 day differences, and horizontal lines mark the cumulative distribution at the median and at 30 days.

Figure A4: Marginal Impacts of Information on School Quality on Parental Investment, by Forecasting Model



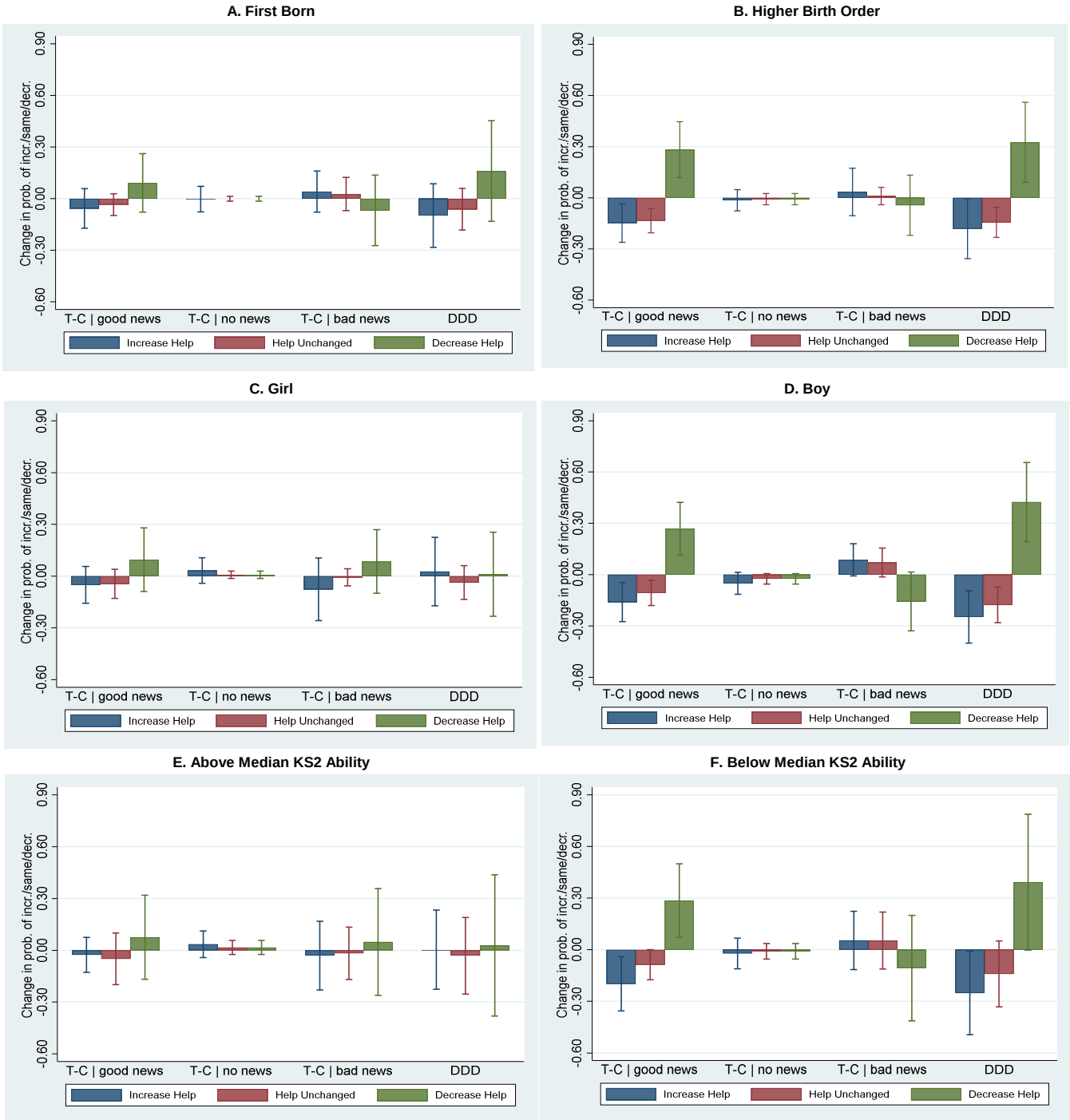
Notes: The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. The outcome variable is the change in parental help with homework between survey wave 3 and 1, or between survey wave 5 and 3: this is coded as -1 if there is a decrease in parental help, 0 if there is no change in parental help and +1 is there an increase in parental help with homework. The Panels vary in the underlying forecasting model used. Panel A uses a model to forecast the Ofsted rating that controls for the past Ofsted inspection grade. Panel B additionally controls for schools characteristics, Columns Panel C additionally controls for school characteristics plus all performance measures. Panel D uses the past Ofsted grade as the forecast grade to derive the news shock variable. The marginal effects of the coefficients from an ordered probit regression for increasing/not changing/decreasing help with homework at home are shown, along with 90% confidence intervals. The standard errors used to derive 90% confidence intervals are clustered at the local authority level and derived using the bootstrap method with 1,000 iterations.

Figure A5A: Marginal Impacts of Information on School Quality on Parental Investment, by Family Characteristic



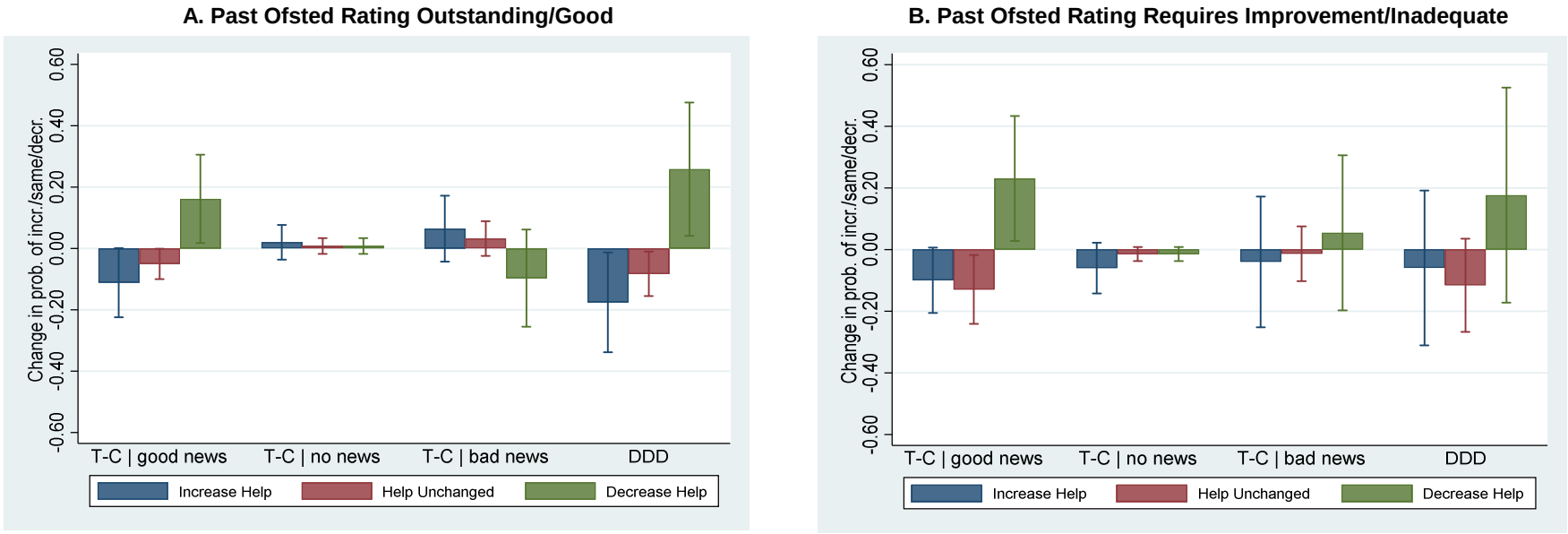
Notes: The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. The outcome variable is the change in parental help with homework between survey wave 3 and 1, or between survey wave 5 and 3: this is coded as -1 if there is a decrease in parental help, 0 if there is no change in parental help and +1 is there an increase in parental help with homework. Panels A and B split the sample to households where parents are highly educated defined as having an A-level or higher education (panel A, N=385) or low educated, defined as having GCSEs or no qualification (panel B, N=305). Panels C and D split households by ethnicity (White, N=511, Non-White, N=179). The marginal effects of the coefficients from an ordered probit regression for increasing/not changing/decreasing help with homework at home are shown, along with 90% confidence intervals. The standard errors used to derive 90% confidence intervals are clustered at the local authority level and derived using the bootstrap method with 1,000 iterations.

Figure A5A: Marginal Impacts of Information on School Quality on Parental Investment, by Child Characteristic



Notes: The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. The outcome variable is the change in parental help with homework between survey wave 3 and 1, or between survey wave 5 and 3: this is coded as -1 if there is a decrease in parental help, 0 if there is no change in parental help and +1 is there an increase in parental help with homework. Panels A and B split the sample by whether the child is a firstborn (N=316) or a higher order child (N=374). Panels C and D split the sample by whether the child is a girl (N=346) or a boy (N=344). Panels E and F split the sample by whether the child had above median KS2 ability, measured as the average of math and English fine points in national curriculum tests (N=193) or below median KS2 ability (N=192). Ability measures are taken from linked National Pupil Database data, available for children with valid linkage consents and successful links. The marginal effects of the coefficients from an ordered probit regression for increasing/not changing/decreasing help with homework at home are shown, along with 90% confidence intervals. The standard errors used to derive 90% confidence intervals are clustered at the local authority level and derived using the bootstrap method with 1,000 iterations.

Figure A6: Marginal Impacts of Information on School Quality on Parental Investment, by Past OFSTED Rating



Notes: The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of Ofsted inspection. The outcome variable is the change in parental help with homework between survey wave 3 and 1, or between survey wave 5 and 3: this is coded as -1 if there is a decrease in parental help, 0 if there is no change in parental help and +1 is there an increase in parental help with homework. The samples are split between schools that were rated Outstanding or Good in their last Ofsted inspection (Panel A, N=403) or were last rated as Requires Improvement or Fail (Panel B, N=287). The marginal effects of the coefficients from an ordered probit regression for increasing/not changing/decreasing help with homework at home are shown, along with 90% confidence intervals. The standard errors used to derive 90% confidence intervals are clustered at the local authority level and derived using the bootstrap method with 1,000 iterations.