THE ECONOMIC JOURNAL

The Economic Journal, **133** (*August*), 2334–2402 https://doi.org/10.1093/ej/uead015
Advance Access Publication Date: 2 February 2023

PARENTAL RESPONSES TO INFORMATION ABOUT SCHOOL QUALITY: EVIDENCE FROM LINKED SURVEY AND ADMINISTRATIVE DATA*

Ellen Greaves, Iftikhar Hussain, Birgitta Rabe and Imran Rasul

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. Inspection ratings provide news that shifts parental beliefs about school quality, and hence investment into children. We study this using household panel data from England, linked to administrative records on school inspection ratings. We find that parents receiving good news over school quality significantly decrease time investment into their children. We provide insights on the distributional and test score impacts of the nationwide inspections regime, through multiple margins of endogenous response of parents and children.

Family- and school-based inputs determine children's human development and academic achievement. It has long been recognised that family and school inputs can be substitutes or complements (Becker and Tomes, 1976; Todd and Wolpin, 2003). We extend this literature to study interactions between family inputs and parental beliefs over school quality. We do so by identifying the causal impact of exogenously released new information on school quality, on parental time investments into their children, on children's own time investments and the ultimate impact these multiple household responses have on the high-stake test scores of children.

Our study context is England, where a source of credible information on school quality is an established nationwide school inspection regime. Most of the existing literature on school accountability makes the strong assumption that parents are perfectly informed on school quality. Relaxing this assumption lies at the heart of our analysis. We examine how parental time investments shift in response to news, or exogenous changes in the beliefs they hold over school quality. We thus help reconcile work on school accountability regimes with the wider literature on parental

* Corresponding author: Imran Rasul, Department of Economics, University College London, Drayton House, 30 Gordon Street, London WC1H 0AX, UK. Email: i.rasul@ucl.ac.uk

This paper was received on 18 December 2019 and accepted on 17 January 2023. The Editor was Frederic Vermeulen.

The authors were granted an exemption to publish their data because access to the data is restricted. However, the authors provided the Journal with temporary access to the data, which allowed the Journal to run their codes. The codes are available on the Journal repository. The data and codes were checked for their ability to reproduce the results presented in the paper. The replication package for this paper is available at the following address: https://doi.org/10.5281/zenodo.7450645.

The article was submitted to The Economic Journal prior to the implementation of The Economic Journal Ethics Policy on 6 October 2021, by which the current Chair of the Publications Committee is not allowed to submit to the journal during their term to mitigate conflicts of interest. Imran Rasul served as Chair of the Royal Economic Society Publications Committee from 01 October 2018 to 10 October 2021 and has been a member of the RES Trustee Board since 21 November 2018.

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), Centre for the Microeconomic Analysis of Public Policy (RES-544-28-0001) and Research Centre on Micro-Social Change: ES/S012486/1 and ES/L009153/1. We thank the editor (Frederic Vermeulen) and three anonymous referees for valuable feedback. 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.

educational choices, which suggests that households lack critical information in relation to schools and aspects of education systems more broadly, that can lead to sub-optimal choices (Pathak and Sönmez, 2008; Lucas and Mbiti, 2012; Abdulkadiroğlu *et al*., 2014; Ajayi *et al*., 2017).

English school inspections are conducted by the Office for Standards in Education (Ofsted). Schools are typically inspected every few 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 a number of 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 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, are based on soft information not necessarily known to parents ex ante, and children might be unable or unwilling to provide such information to their parents (Weinberg, 2001; Jacob and Lefgren, 2007; Burgess *et al.*, 2015; Beuermann *et al.,* 2018; Bergman, 2021). There can thus be informational content in school 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 these 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, and exploiting only the ordinal information in inspection ratings. We use the model to define whether the inspection rating reveals good, bad or no news to parents about school quality.[2]

To study the impact this news has on parental behaviour, we exploit household panel data from the UK Household Longitudinal Survey (UKHLS) that records parental time investment into their children's academic studies, as well as children's own time investments and other parent-child interactions. 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 that this treatment assignment is as good as random.

Our research design can be summarised as follows. Consider schools inspected in a given year t, and *hold constant* whether parents will receive good, bad or no news. The control group comprises households interviewed in survey year t, but prior to the inspection actually taking place. Treated households are *also* in schools inspected in the same year, but happen to be

[1] In the United States, 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 no child left behind (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 school ratings: two notable exceptions are Rouse *et al.* (2013) and Feng *et al.* (2018), who built 'accountability shocks' in the context of NCLB.

interviewed after the inspection takes place (and so know whatever news is released about 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 have updated their 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 that (*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 stylised framework to make clear parental preferences, the production function for child human capital (or child quality) and parents' optimisation problem. This makes precise how parental investments respond to informational shocks parents receive about school quality, and the conditions under which parental beliefs about school quality and parental time investments into children are complements or substitutes in the production function for child human capital. Following Todd and Wolpin (2003) and Pop-Eleches and Urquiola (2013), we then extend the framework to derive the overall impacts of multiple household responses on children's academic achievement, and so make precise what can be inferred about the relative total products of these various input margins in producing test scores. These elements are 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 more likely to reduce 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 the production function for child human capital.

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. We then use our data to help calibrate a simple model of parental investments. This calibration exercise shows that, given the distribution of news across schools, the impacts of the information released by the inspection regime are to (*i*) reduce the expected level of parental inputs marginally; (*ii*) reduce across-school inequality in parental inputs by 18%. 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 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 behavioural 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 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 around 200,000 children in nationwide high-stake exams taken at age 16.

We find that 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] What can explain the overall fall in test scores as a result of parents and children receiving good news over school quality? As discussed in more detail later, given imperfect information of parents and children, their combined responses to news on school quality can potentially lead them to make mistakes that reduce children's human capital development (at least in the short run as our design allows us to measure).

Our work provides novel insights for three important 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 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; Das *et al*., 2013).

Our contribution to this literature is to understand how households' beliefs about school quality (rather than measures of school quality constructed by researchers) affect parental investments into children, and so our work is closest to Ainsworth *et al.* (2020).

Furthermore, we provide novel evidence on parental and child responses to new information on school quality, studying within-household interactions in the production of children's human capital. In doing so, we complement a nascent literature on parents' educational investment response to child-level information interventions. While other papers have shown—some using field experiments—that providing information to parents can affect educational decisions, either in terms of school selection (Hastings and Weinstein, 2008; Hoxby and Avery, 2012; Ajayi *et al*., 2017; Andrabi *et al*., 2017; Neilson *et al.,* 2019; Ainsworth *et al.,* 2020) or improving student effort/behaviour (Jensen, 2010; Avvisati *et al.,* 2013), few have done so to examine parent and child time investments, or their interaction.[4]

The third literature we contribute to is 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; Hastings and Weinstein, 2008; Figlio and Loeb, 2011; Hussain, 2020). 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

[3] Hanushek and Raymond (2005), Jacob (2005), Figlio and Loeb (2011) and Burgess *et al.* (2013) studied 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).

[4] On parental responses to child-level information interventions, Dizon-Ross (2019) investigated the effect of revealing the child's ability on school enrolment and other education inputs. Bursztyn and Coffman (2012) and Bergman (2021) addressed information interventions in the form of monitoring technologies, designed to address the strategic interactions arising from diverging parent-child preferences and the inability of parents to perfectly monitor child actions.

children in any school grade) than those facing an initial school choice problem.[5] Such policies can also reinforce/mitigate inequalities within and across schools and families, as we document. Given the global roll out of school accountability regimes and widespread use of report cards (Figlio and Loeb, 2011), these are relevant issues for education systems around the world.

The paper is organised as follows. Section 1 develops a framework to understand how parental inputs vary with beliefs over school quality. Section 2 describes our linked household and school administrative data. Section 3 presents our research design and identifying assumptions. Section 4 contains our core findings on parental responses to news, and calibrates the distributional impacts of the inspections regime. Section 5 examines impacts on test scores through multiple endogenous responses of parents and children. Section 6 concludes. The appendices contain proofs and further robustness checks.

1. Conceptual Framework

1.1. *Set-Up*

We present a simple framework to understand how shocks to parental beliefs over school quality impact their time investment into their children.[6] Parents are assumed to invest in one child, and be uncertain over school quality. Parental utility is $U(C, H)$, where C denotes consumption and H denotes the child's human capital, and this is taken as the numeraire good. Utility $U(\cdot)$ is concave in each argument. The production function for child human capital is determined by school quality (S) and parental investments (I), $H = f(S, I)$, where f is concave in each argument. We consider time investments made by parents into their child. This matches what we empirically measure, and it is well recognised that time investments are an important input into children's human capital (Cunha *et al.,* 2010; Avvisati *et al.,* 2013; Del Boca *et al.,* 2013; Fiorini and Keane, 2014; Carneiro *et al.*, 2015; Bono *et al.*, 2016; Bergman, 2021; Maldonado *et al.*, 2021).

Parents are uncertain over school quality, and their prior belief is denoted $\overline{S}$. Although earlier work has shown that inspection ratings drive school choice, this does not mean that school quality is necessarily fully revealed to parents once their child is enrolled. Indeed, there is mixed evidence on how well informed parents are about school quality (Abdulkadiroğlu *et al.*, 2014; Beuermann *et al.*, 2018). Furthermore, children may not correctly report school quality to their parents, perhaps because they are inexperienced regarding public services in general, and hence a poor judge of the quality of education being delivered. Alternatively, they may strategically misreport quality because truthful reports may entail parental demands on margins such as child effort.[7]

New information generated from Ofsted inspection ratings leads parents to update their beliefs about school quality to $\overline{S} + \mu$, where the news shock μ may be positive or negative. Parents then

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

[6] Currie and Almond (2011) and Yi *et al.* (2015) presented related models investigating the impact of health shocks on parental investments. Greenwood (2019) provided similar examples from household production theory.

[7] This literature on parent-child interactions recognises parents' need to invest, motivate and monitor their children's academic progress by providing incentives (Weinberg, 2001; Hao *et al.*, 2008) or using certain parenting styles (Burton *et al.*, 2002; Doepke *et al.*, 2019). Todd and Wolpin (2003) also discussed the possibility that there may be a deviation between the school-level inputs chosen by the household at the time of the school entry and the level of school inputs actually received.

re-optimise their investment in response to this shock. Thus, the child's human capital is given by $H = f(\overline{S} + \mu, I(\mu))$.

Parental investment responses to the change in perceived school quality, $\partial I/\partial \mu$, depends on the crowding in or out of parental inputs as beliefs about school quality change. This relates to the broad literature on the interplay between public and private investments, and as we document later, this also has implications for how the school inspection regime impacts inequality of parental inputs within and across schools. We assume that schools face short-run adjustment costs and so do not immediately respond to the release of information on school quality (an assumption that matches our institutional setting and that is empirically validated below).[8]

Parents have a unit of time at their disposal and choose how to allocate this time between investments into their child and work, which earns w per unit of time. Therefore, parent's time budget constraint is given by $C = w(1 - I)$, and their optimisation problem is

$$\max_{C,I} U = U(C, H), \qquad H = f(\overline{S} + \mu, I), \qquad C = w(1 - I).$$

As parents choose their investment and consumption after information about school quality is revealed, the maximisation problem yields the first-order condition

$$\frac{\partial U}{\partial H}\frac{\partial H}{\partial I} = \frac{\partial U}{\partial C}w. \tag{1}$$

This simply states that parents invest in children up to the point at which the marginal benefit of time investment is equal to its marginal cost in terms of foregone consumption.

1.2. *Parental Time Investment*

We place a little more structure onto the problem to proceed further. Following Currie and Almond (2011), we assume Cobb–Douglas parental preferences and a constant elasticity of substitution production function for the child's human capital:

$$U = \theta \ln(C) + (1 - \theta)\ln(H), \qquad H = a[\gamma(\overline{S} + \mu)^{\rho} + (1 - \gamma)I^{\rho}]^{1/\rho}, \tag{2}$$

where $0 \leq \theta \leq 1$ and $0 \leq \gamma \leq 1$. The degree of complementarity between school quality and parental investment is determined by ρ, where $\rho \leq 1$. Here $\rho \ll 0$ implies strong complementarity between parental investment and school quality; ρ close to 1 implies that parental investment and school quality are readily substitutable; $\rho = 0$ implies a Cobb–Douglas production function for the child's human capital. As shown in Appendix A, substituting these functional forms into the first-order condition (1) yields

$$\frac{(1 - \theta)(1 - \gamma)I^{\rho - 1}}{a[\gamma(\overline{S} + \mu)^{\rho} + (1 - \gamma)I^{\rho}]} = \frac{\theta}{1 - I}, \tag{3}$$

where the left-hand side represents the marginal benefit of child investment and the right-hand side represents the marginal cost. Note that the marginal cost is rising in I and is independent of the school quality shock, μ. The left-hand side of (3) shows that the marginal benefit is falling in the amount of time investment I, both because of diminishing marginal utility of child human

[8] In England schools make staffing decisions towards the end of each academic year. Evidence in favour of such short-run adjustment costs has been documented for the United States and the UK (Rouse *et al.*, 2013; Hussain, 2015). Of course, in the longer term, school accountability systems might well impact teacher turnover (Figlio and Loeb, 2011; Feng *et al.*, 2018; Dizon-Ross, 2020).

capital and also because of diminishing marginal product of investment. This makes clear that two margins of substitution are relevant for the level of parental time investment: substitution between consumption and the child's human capital in parents' utility (θ), and the technical substitution between inputs in the production function for the child's human capital (ρ). The former affects both the marginal cost and benefit of investing, while the response of the marginal benefit curve to the school quality shock depends on the latter.

If $0 < \rho \leq 1$ then a positive news shock on school quality (a rise in μ) leads to a downward shift in the marginal benefit curve, leading to a fall in child investment. This follows because, as perceived school quality rises, the expected level of child human capital also rises, and parents can increase utility by raising their level of consumption goods and cutting back on time investment I. Any gain via higher marginal productivity of child investment is not sufficient to offset these forces. In this case public investments that raise μ do indeed crowd out private investments, as originally emphasised in the classic study by Becker and Tomes (1976).

The situation is reversed when $\rho < 0$: the marginal benefit curve shifts upwards in response to a positive news shock to μ and child investment rises. In this case the impact on the marginal product of child investment is large enough to offset any tendency to cut back on these investments arising from diminishing marginal utility of child human capital.

For the special case where $\rho = 0$, even though parents are uncertain over school quality, there is no response in parental child investment to the news shock μ.

This result can also be demonstrated more formally as follows. Using the implicit function theorem, differentiating the first order condition (FOC) (3) by μ yields

$$\frac{(1-\theta)(1-\gamma)}{\theta}[(\rho-1)I^{\rho-2}-\rho I^{\rho-1}]\frac{\partial I}{\partial \mu} = a\rho\gamma(\overline{S}+\mu)^{\rho-1} + a\rho(1-\gamma)I^{\rho-1}\frac{\partial I}{\partial \mu},$$

which, as shown in Appendix A, simplifies to

$$\frac{\partial I}{\partial \mu} = \frac{-\rho\gamma(\overline{S}+\mu)^{-1}I^{\rho-1}(1-I)}{\gamma+(1-\gamma)(\overline{S}+\mu)^{-\rho}I^{\rho}-\rho\gamma(1-I)}. \tag{4}$$

The denominator on the right-hand side is positive since $\rho \leq 1$. Thus, the sign of $\partial I/\partial \mu$ depends on the value of ρ: in line with the discussion above, $\partial I/\partial \mu > 0$ if $\rho < 0$; $\partial I/\partial \mu < 0$ if $\rho > 0$ and $\partial I/\partial \mu = 0$ if $\rho = 0$. Given the Cobb–Douglas functional form assumption, the substitution between consumption and the child's human capital in parents' utility (θ) does not matter for the marginal response of parental investment to news on school quality.[9]

Two further points are of note. First, given the global roll out of school accountability regimes and widespread use of report cards (Figlio and Loeb, 2011), the model makes precise how family input responses to information on school quality might differ across contexts. In particular, parental priors $\overline{S}$ will differ (hence, the response to new information will differ) if the market for information on school quality is better developed, or because mechanisms enabling households to sort into schools differ across contexts. Second, our modelling framework follows much of the existing literature in assuming that parents make one investment over the academic year. In reality, parents continually invest. How they respond to news about school quality will depend on these earlier investments, and how far along the school year information is revealed. We address these points when we later set out our empirical research design.

[9] Yi *et al.* (2015) discussed the more general point that functional form assumptions determine which parameters drive investment on the margin. Our formulation is as in Currie and Almond (2011). In contrast, Behrman *et al.* (1982) assumed a constant elasticity of substitution parental utility function and a Cobb–Douglas production function for child human capital. They showed that the optimal investment strategy is uniquely determined by parental preference parameters.

1.3. *Multiple Response Margins and Test Scores*

Our second set of empirical results consider a wider set of household responses to news on school quality, and their subsequent impact on children's test scores. In contrast to the literature on early childhood development, where it is reasonable to assume that parents fully control investments into their children's skills (Cunha and Heckman, 2007; Heckman and Mosso, 2014), for older children in the age range that we study, it is more accurate to recognise adolescents as economic agents with an ability to influence their own outcomes.[10]

To understand how equilibrium responses by parents and children combine to influence test scores, we follow Pop-Eleches and Urquiola (2013), and adapt the Todd and Wolpin (2003) framework to our context. We continue to assume that parents have prior beliefs about school quality, $\overline{S}$, and that an Ofsted inspection rating leads to a revision of beliefs on school quality, $\overline{S}+\mu$. We then allow for multiple family investments of parents and children, denoted I^P and I^C, respectively. We assume that parents act as Stackelberg leaders in the investment game, and so child investments respond to parental investments as well as to μ. The child's human capital, proxied by test scores or achievement, A, is given by

$$A = g(\overline{S}+\mu, I^P, I^C).$$

We can then write the total policy effect of a positive news shock on school quality (a rise in μ) as

$$\frac{dA}{d\mu} = \frac{\partial g}{\partial \mu} + \frac{\partial g}{\partial I^P}\frac{\partial I^P}{\partial \mu} + \frac{\partial g}{\partial I^C}\left[\frac{\partial I^C}{\partial \mu} + \frac{\partial I^C}{\partial I^P}\frac{\partial I^P}{\partial \mu}\right]. \tag{5}$$

The first term on the right-hand side of (5), $\partial g/\partial \mu$, is the direct effect of the school quality information shock on achievement via any school response to Ofsted inspection. The second term represents the indirect parental investment response: this is a product of the parental investment response to the information shock and the marginal impact of parental inputs. The final term represents the indirect response of children's own investments: these respond to the information shock ($\partial I^C/\partial \mu$), but may also be mediated via a response to the change in parental inputs ($(\partial I^C/\partial I^P)(\partial I^P/\partial \mu)$).

As with parental inputs, the responsiveness of child inputs can vary with the degree of complementarity or substitutability with school quality; indeed, the child and parental responses could conceivably be of opposite signs. For example, if there is relatively strong (weak) complementarity between child (parental) inputs and school quality, then it is possible that child investments increase whilst parental investments decrease. The final component in (5), which includes the term capturing child input responses to changes in parental input, $\partial I^C/\partial I^P$, reflects the possibility that there may be interaction between parental inputs and child inputs (De Fraja *et al.*, 2010; Bergman, 2021). For instance, parental help with homework may lead to higher levels of child own investment; on the other hand, in the presence of information frictions the child may slack and lower their own inputs when parental effort rises.

Finally, setting $\partial g/\partial \mu = 0$ (so that there are no short-term school-based responses to Ofsted as validated below), then finding an impact on post-treatment test scores ($dA/d\mu \gtreqless 0$) implies

[10] Indeed, researchers typically use non-cooperative game theory when modelling interactions between parents and their adolescent children (Weinberg, 2001; Burton *et al.*, 2002; Hao *et al.*, 2008). Our framework does not include such bargaining, but rather follows Todd and Wolpin (2003) in assuming that parents are Stackelberg leaders in making investments into the child's human capital, while maintaining the focus on how investments and children both respond to changes in beliefs over school quality.

that the relative total products of family to child inputs in generating test scores can be assessed:

$$\frac{\partial g}{\partial I^P}\frac{\partial I^P}{\partial \mu} \gtreqless -\frac{\partial g}{\partial I^C}\left[\frac{\partial I^C}{\partial \mu}+\frac{\partial I^C}{\partial I^P}\frac{\partial I^P}{\partial \mu}\right]. \tag{6}$$

To be clear, parent and child responses to news on school quality can move in opposite directions. Hence, the framework does not guarantee that multiple investment responses within the household to new information will leave test scores unchanged or higher. Given imperfect information of parents and children, their combined responses to news on school quality can potentially lead them to make mistakes that reduce children's human development (at least in the short run as our design allows us to measure). For example, parents might be imperfectly informed about the skills of their child, or the marginal productivity of the child's own time investment ($\partial g/\partial I^C$). A growing evidence base suggests that parental investments into children are related to their beliefs over child skill and the productivity of inputs into the production function for a child's human capital (De Fraja *et al.*, 2010; Boneva and Rauh, 2018; Dizon-Ross, 2019; Attanasio *et al.*, 2020a; 2020b). These typically find parents have upwards biased beliefs about their children's skills or academic performance (Dizon-Ross, 2019; Bergman, 2021; Kinsler and Pavan, 2021).

We return to this issue in more detail in Section 5 once we examine the impact of news on school quality on a wider range of household behaviours and parent-child interactions, and on high-stake test scores.

2. Context and Data

2.1. *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.[11]

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. Tables E1 and E2 in Appendix E detail 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, and children might be unable or unwilling to provide such information to their parents (Weinberg, 2001; Jacob and Lefgren, 2007; Beuermann *et al.*, 2018; Bergman, 2021). There can thus be informational

[11] 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.

content in school ratings, constituting news to parents that shifts time investments into their children.[12,13]

2.1.1. *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). 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 D1 in Appendix D provides an example of an Ofsted letter sent to parents. The letter is simple, concise and clearly states the headline inspections rating.

The salience of inspection ratings to parents is corroborated by evidence: an annual survey of parents undertaken by YouGov since 2015 shows that 90% of parents are aware of their child's school's inspection rating.[14]

Beyond the headline rating, four sub-components of school quality are also rated by inspectors: achievement of pupils, quality of teaching, behaviour and safety of pupils, and leadership and management. While these aspects might also be 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. Table E1 in Appendix E shows the sub-components rated (and the dimensions considered in each); Table E2 in Appendix E 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 is not easily available to parents ex ante.

2.2. *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 high-stake nationwide test score impacts of the inspections regime.

2.2.1. *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 measure of time investment, is collected). The exact interview date is recorded in each wave.[15]

[12] High-stake 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.

[13] 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).

[14] 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.

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

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 E3 in Appendix E details sample characteristics as we make each selection towards our working sample of 621 households.[16]

The 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).[17] To be clear, this measure of parental investment cannot capture the total time spent with children. However, to map the data to the model, we only need to measure the change in time spent with children in response to news shocks, and the measure can capture such changes.

The change in parental time investment, $\Delta I_t = I_t - I_{t-2}$, is measured between waves 1 and 3, and 3 and 5. These survey waves—two years apart—are those in which data on parental time investment are available. Given the inspection cycle, the majority of schools (around 80%) are last inspected further back in time than $t-2$. Here $\Delta I_t = -1$ 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 the same school across waves (although the majority of children are).

Figure 1(a) 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 11%–12% report never or hardly ever helping. Panel (b) then shows within-household changes over time. Averaging across these changes we see that (*i*) 19% of parents increase time investments; (*ii*) 43% keep constant their time investment; (*iii*) 38% of parents decrease their time investment.[18]

We later complement these with data on children's behaviours 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. This allows us to map directly to the second part of the conceptual framework, and shed light on how parent and child time investments into the child's

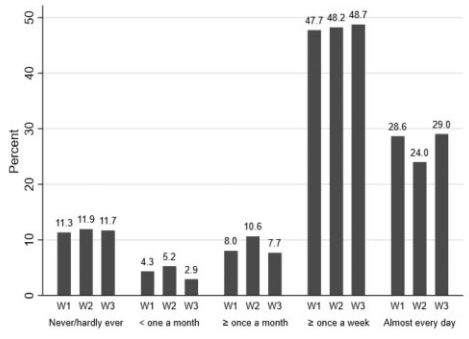
covers academic years 2010/11, 2011/12, 2012/13; survey wave 5 takes place between January 2013 and December 2014, and thus (partly) covers academic years 2012/13, 2013/14, 2014/15.

[16] Appendix Table E3 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 that 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.

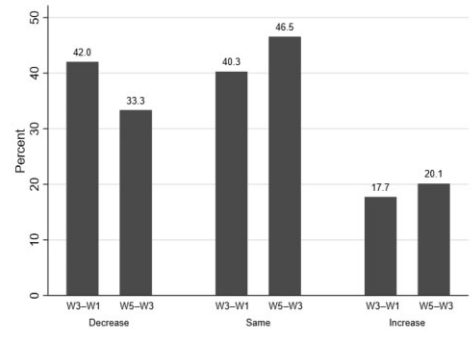
[17] This question is asked separately of both parents if they have one or more children aged 10−15. Where responses are available for both parents, we choose the dominant parental helper, defined as the one helping more. When both parents are present, 41% of the time the dominant parent is the mother. Where both parents help equally (41% of the time) we choose the mother's response.

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.* (2013) presented evidence from a dynamic structural model of child development that suggests that maternal and paternal investments are equally productive.

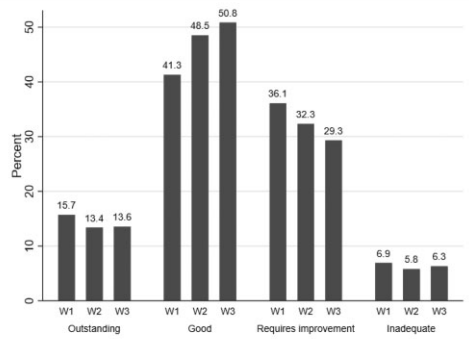
[18] To be clear, our analysis takes the selection of children into schools as given. However, as previous research has shown, this selection is likely driven by past Ofsted ratings, and this in turn might then impact the level of parental investments in our sample.



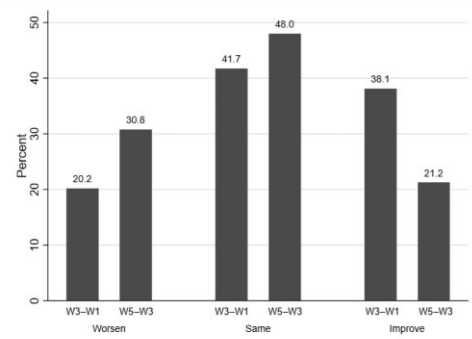
(a) Parental time investment into homework, by wave



(b) Change in parental time investment into homework, by wave



(c) Ofsted school inspection rating, by wave



(d) Change in Ofsted school inspection rating, by wave

Fig. 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: 'W1' is wave 1; 'W2' is wave 2; 'W3' is wave 3. Panel (b) shows changes in parental time investment, separately for changes between survey waves 3 and 1 ('W3–W1') and survey waves 5 and 3 ('W5–W3'). 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.

human capital respond to news on school quality, and the subsequent impact on high-stake test scores for the child.

Sample sizes do not permit us to examine in any detail parental investments beyond time—that of course could respond in opposite directions to news on school quality—although we reiterate that time investments are recognised as an important input into children's human capital.[19]

[19] We note that alternative measures of parental time investments have been used in the related literature, such as parent-child interactions around dinner times that have been argued to be important for educational outcomes and other dimensions of child well being (Houtenville and Conway, 2008; Cunha and Heckman, 2009).

2.2.2. *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—these provide characteristics of the student body and school type; (*iii*) Ofsted inspections data—these provide inspection outcomes and the exact date of inspection.[20]

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.[21]

Ofsted data cover all inspections from September 2005 until December 2014, covering 63,337 inspections 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 E4 in Appendix E. 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.

Figure 1(c) shows inspection ratings by survey wave. The distribution of ratings is relatively stable over time: around 13%–16% of schools receive an outstanding rating, 41%–51% receive a good rating, 29%–36% receive a rating of requires improvement and 6%–7% of schools are rated as failing/inadequate. Panel (d) shows within-school rating changes. The majority of schools change rating: 21%–38% of schools experience an improved rating and 20%–31% of schools have a worse rating.[22]

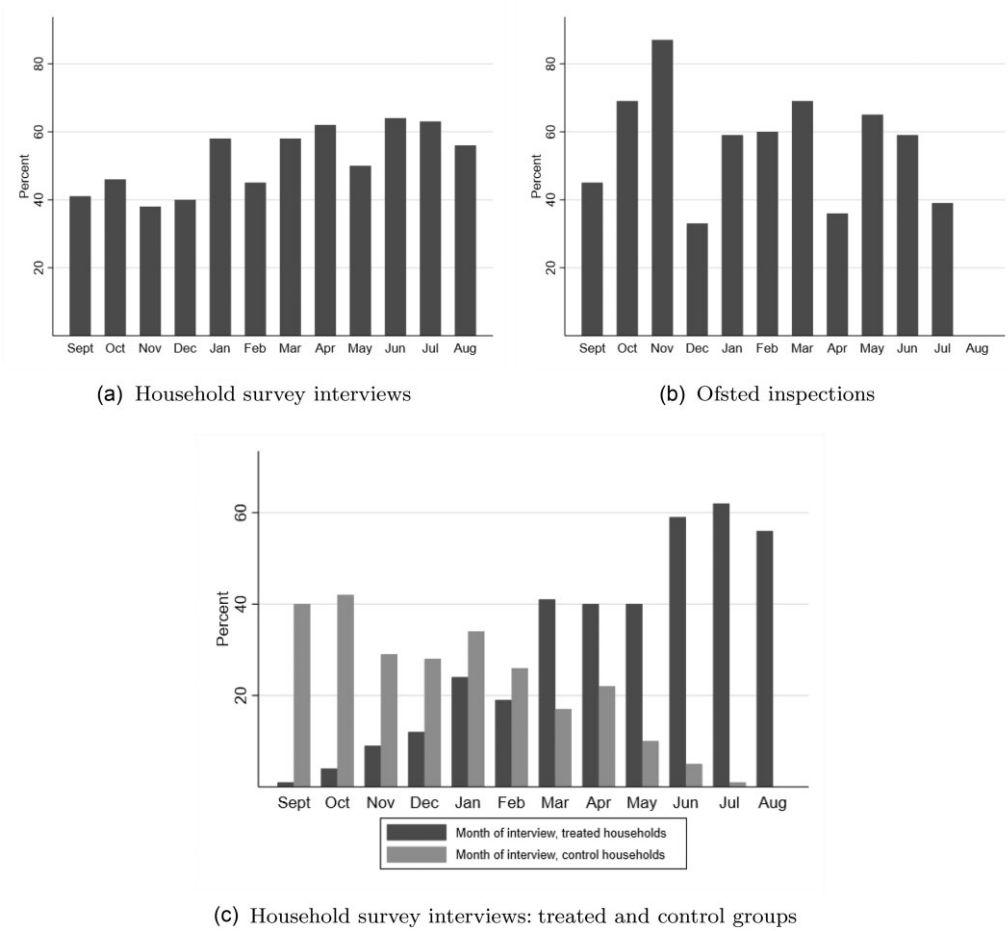
Finally, we link these school data sets to administrative data on individual child test scores from the National Pupil Database (NPD). We use this to examine the test score impacts of news generated by the nationwide inspections regime.[23]

[20] The school identifier is collected in waves 1, 3, 5. Households were also asked to provide consent to link their children's data to test score records in the 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 the 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.

[21] The following school test score indicators are available: the percentage of pupils with five or more A*–C grades, the percentage with five or more A*–C grades including English and maths, the percentage with five 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 contain information on school size (number of pupils), the percentage of pupils entitled to free school meals, the school type (academy, community, voluntary aided, controlled, foundation, special), whether it has a sixth form, any Christian or other religious denomination and whether it is a mixed gender school.

[22] We note that, while the level of Ofsted inspection ratings do correlate with academic attainment, changes in Ofsted ratings are not associated with immediate changes in academic performance. As such, Ofsted ratings provide more holistic information to parents than achievement or test score data alone. This is in line with the aims of the regime, which covers four sub-components of school quality as described above: achievement of pupils, quality of teaching, behaviour and safety of pupils, and leadership and management.

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



(a) Household survey interviews

(b) Ofsted inspections

(c) Household survey interviews: treated and control groups

Fig. 2. *Timing of Household Surveys and Ofsted Inspections.*
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 the 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.

3. Empirical Method

3.1. *Defining Treatment and Control Households*

School inspections take place 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).

Figure 2(a) 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 patterns of timing of household interviews and school inspections ameliorate concerns over UKHLS enumerators or inspectors front/back-loading their effort, which could otherwise have led to measurement error in parental behaviour 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 periods 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. Both treatment and control households are therefore assumed to have the same beliefs about the likelihood and timing of school inspection during the year, and hence are expected to undertake similar time paths of investments pre-inspection.

Panel (c) shows the 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. Of the sample households, 41% are controls and 59% are treated.

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

3.2. *Measuring News*

Parental beliefs on school quality 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

$$news_{\sigma t} = \text{rating}_{\sigma t} - \text{predicted rating}_{\sigma t}. \tag{7}$$

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, which we can check for (although in our context, children are aged 10–15 and are mostly in secondary schools with over 1,000 pupils).[24]

[24] The forecast is constructed at the school level: the UKHLS has no information at the household level of expected inspection ratings. Beuermann *et al.* (2018) and Abdulkadiroğlu *et al.* (2020) overviewed recent work examining whether parents can tell what constitutes a good school. We further note that in a theoretical work that allows for multidimensional school quality, it can be rational for households to prioritise school attributes other than value added or whatever is measured by inspection ratings (Beuermann *et al.*, 2018; MacLeod and Urquiola, 2019; Abdulkadiroğlu *et al.*, 2020).

To maximise 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 ($rating_{\sigma t-1}$).

To construct the predicted rating, we have to respect the fact that Ofsted ratings provide ordinal, not cardinal information on school quality. We thus work primarily with transition probabilities of inspection ratings from one inspection cycle to the next ($rating_{\sigma t-1}$, $rating_{\sigma t}$). The unconditional transition matrix is shown in Table 1. Panel A shows this for all schools in our sample of inspections. In line with the earlier evidence from Figure 1(d) on changes in inspection rating over time, we see that there are transitions to different ratings over time, especially moves of one rating up or down. Panel B shows the transition matrix for the schools that our working sample of children attend. Comparing the two samples, we see a very similar proportion of transitions in the majority of cells, with there being some small differences in a few—two of which are for the small number of schools rated as inadequate in their earlier inspection.

We then use an ordered probit model to estimate each transition probability:

$$prob(rating_{\sigma t} = b | rating_{\sigma t-1} = a). \tag{8}$$

Here $a, b = 1, 2, 3, 4$ correspond to the four-point scale used by Ofsted as headline ratings: 4 (outstanding), 3 (good), 2 (requires improvement) and 1 (inadequate/failing). We estimate (8) conditioning on school performance and school characteristics ($Z_{\sigma t}$). Table E5 in Appendix E shows the results. Columns (1) to (4) estimate (8) for each previous rating $rating_{\sigma t-1} = a$. At the foot of each column we report the mean predicted rating for each origin grade. For example, for the 297 schools with $rating_{\sigma t-1} = inadequate$, 0.5% are predicted an outstanding rating in the next inspection ($rating_{\sigma t}$), 17.0% are predicted a rating of good, 69.8% are predicted as requiring improvement and 12.7% are predicted as still being inadequate. The average predicted rating is 2.04 (where 2 is 'requires improvement').

To only exploit ordinal information contained in inspection ratings, the predicted rating is then the most likely rating in the current cycle:

$$\text{predicted rating}_{\sigma t} = \text{mode}(\widehat{prob}(rating_{\sigma t} = b | rating_{\sigma t-1} = a)).$$

Using the mode for the predicted rating is an assumption that we make to retain the ordinal nature of Ofsted ratings in our empirical application.[25] However, the modal Ofsted rating is predicted with relative certainty, which suggests that, given $rating_{\sigma t-1}$ and observable time-varying school covariates, parents would have reasonable certainty about the expected school inspection rating at period t. As the foot of Appendix Table E5 shows, for each Ofsted rating in $t-1$, the median mode predicted probability is above 0.54, with the 25th percentile also being high, with a minimum value across prior Ofsted ratings of 0.48.

We use the model prediction to define $news_{\sigma t}$ as follows:

$$news_{\sigma t} = \begin{cases} good_{\sigma t} & \text{if (actual rating} - \text{predicted rating}_{\sigma t} > 0), \\ none_{\sigma t} & \text{if (actual rating} - \text{predicted rating}_{\sigma t} = 0), \\ bad_{\sigma t} & \text{if (actual rating} - \text{predicted rating}_{\sigma t} < 0). \end{cases} \tag{9}$$

[25] Ignoring the ordinality requirement, it could be equally reasonable to assume that parents do not use as an expected quality the mode of their ex ante expectation, and might use more information from the entire probability distribution of school quality outcomes.

Table 1. *Transitions in School Inspection Ratings across Inspection Cycles.*

		Ofsted rating from earlier inspection (rating $\sigma_{t-1} = a$)				
		Outstanding	Good	Requires improvement	Inadequate	Total
Panel A: all schools						
Ofsted rating from new inspection (rating $\sigma_t = a$)	Outstanding	261 (41.7%)	327 (18.6%)	45 (2.6%)	1 (0.3%)	634
	Good	291 (46.5%)	873 (49.5%)	697 (40.6%)	55 (17.4%)	1,916
	Requires improvement	52 (8.3%)	433 (24.6%)	727 (42.4%)	219 (69.3%)	1,431
	Inadequate	22 (3.5%)	129 (7.3%)	246 (14.3%)	41 (13.0%)	438
	Total	626	1,762	1,715	316	4,419
Panel B: estimation sample						
Ofsted rating from new inspection (rating $\sigma_t = a$)	Outstanding	44 (38.3%)	48 (19.4%)	7 (3.0%)	0 (0.0%)	99
	Good	53 (46.1%)	139 (56.0%)	96 (40.9%)	4 (17.4%)	292
	Requires improvement	12 (10.4%)	53 (21.4%)	101 (43.0%)	18 (78.3%)	184
	Inadequate	6 (5.2%)	8 (3.2%)	31 (13.2%)	1 (4.3%)	46
	Total	115	248	235	23	621

Notes: The table shows transition numbers, with column percentages in parentheses. The sample in panel A comprises schools that are inspected during the survey period covering waves 1, 3 and 5 of UKHLS. The sample in panel B 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, our main estimation sample. The table shows, in each panel, the number and proportion of schools that transition from each Ofsted rating at an earlier inspection (comprising the ratings outstanding, good, requires improvement and inadequate), shown in columns, to a new inspection rating at time t, shown in rows.

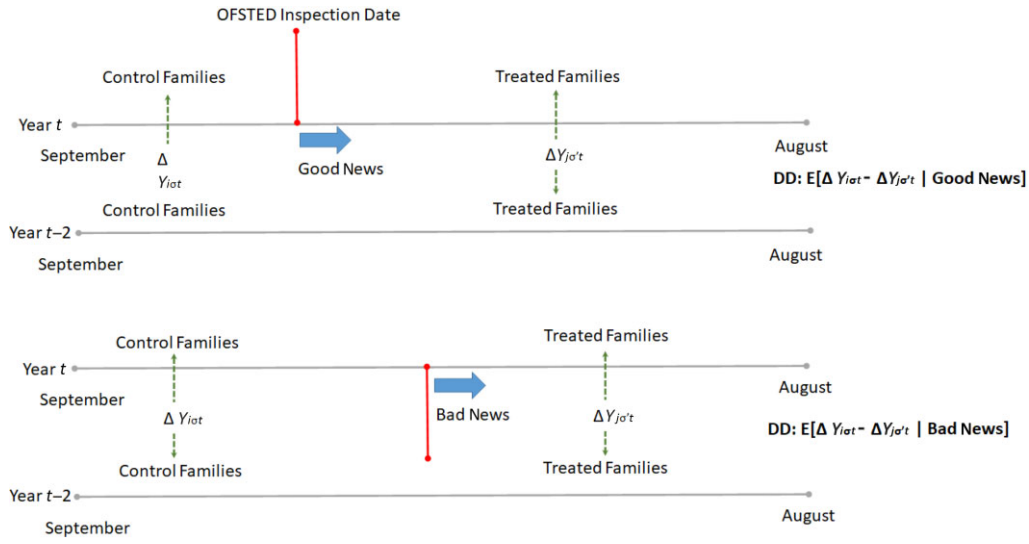


Fig. 3. *Research Design.*

News $news_{\sigma t}$ corresponds to μ, the change in belief in school quality. The (actual rating − predicted rating$_{\sigma t}$) ranges from −2 to +1, with 20% of parents receiving bad news, 19% receiving good news and 61% receiving no news (because (actual rating − predicted rating$_{\sigma t}$) = 0) (see Table E6 in Appendix E). We have variation in $news_{\sigma t}$ to identify parental responses to new information: 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. We later exploit this full variation in news across the schools to shed light on distributional impacts for parental inputs across schools.

3.3. *Research Design*

Figure 3 shows our research design, combining all the elements above. Parental time investment in household i whose children attend school σ in period t is denoted $Y_{i\sigma t}$. 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 $E[\Delta Y_{i\sigma t} - \Delta Y_{j\sigma' t} \mid good_{\sigma t}, good_{\sigma' t}]$, mapping to $\mu > 0$; (*ii*) bad news is received (lower panel), so the key DD is $E[\Delta Y_{i\sigma t} - \Delta Y_{j\sigma' t} \mid bad_{\sigma t}, bad_{\sigma' t}]$, mapping to $\mu < 0$ (there is a corresponding DD estimated for parents receiving no news that we do not show in Figure 3 for expositional clarity). 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:

$$\Delta Y_{i\sigma t} = \delta_0 + \beta_0 T_{i\sigma t} + \beta_1 [T_{i\sigma t}.good_{\sigma t}] + \beta_2 [T_{i\sigma t}.bad_{\sigma t}] + \delta_1 good_{\sigma t} + \delta_2 bad_{\sigma t} + \gamma_1 X_{i\sigma t} + \gamma_1 Z_{\sigma t} + \varepsilon_{i\sigma t}. \quad (10)$$

Here $\Delta Y_{i\sigma t}$ is the 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 t; the $X_{i\sigma t}$ are child- and family-level controls and the $Z_{\sigma t}$ are school-level controls.[26] As $\Delta Y_{i\sigma t} \in \{-1, 0, 1\}$, we estimate (10) using an ordered probit model.[27]

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.[28]

By examining the change in parental time investment, $\Delta Y_{i\sigma t}$, 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; instead, 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 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. Both treatment and control households are assumed to have the same beliefs about the likelihood and timing of school inspection during the year, and are hence expected to have similar time paths of investments pre-inspection.

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.

3.4. *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.

We examine assumption (*i*) using two strategies. First, panel A of Table 2 shows school characteristics by treatment and control (Figure D2 in Appendix D shows *p*-values on balance

[26] The $X_{i\sigma t}$ 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), the mother's ethnicity (White, non-White, missing), the mother's highest education (five binary indicators) and the mother's marital status (married/cohabiting, single, missing). Where the mother's information is missing, the 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.

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

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

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 that $\varepsilon_{i\sigma t} \sim N(0, 1)$ yields the ordered probit model where $prob(\Delta Y_{i\sigma t} = j) = \phi(\mu_j - \beta' \mathbf{X}_{i\sigma t}) - \phi(\mu_{j-1} - \beta' \mathbf{X}_{i\sigma t})$.

[28] There are analytic correction methods to correct for generated regressors, such as the two-step variance estimator derived in Murphy and Topel (2002), 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 1,000 replications.

Table 2. *Balance.*

	Treated: interviewed after Ofsted inspection (1)	Control: interviewed before Ofsted inspection (2)	Normalised difference (3)	Test of equality [*p*-value] (4)
Number of children	367	254		
Panel A: school characteristics				
School size: number of pupils	1,146	1,091	0.107	[0.077]
	(375)	(347)		
% Pupils free school meals	17.50	16.75	0.040	[0.553]
	(12.50)	(14.09)		
Academy school	0.21	0.26	−0.090	[0.094]
Boys' school	0.06	0.04	0.073	[0.177]
% Pupils 5 or more A*–C grades	0.77	0.77	−0.025	[0.681]
	(0.17)	(0.14)		
Total average GCSE point score	331.16	335.00	−0.067	[0.264]
	(46.20)	(33.90)		
Panel B: household characteristics				
Household size	4.11	4.13	−0.011	[0.856]
	(1.32)	(1.27)		
Home owner	0.64	0.69	−0.077	[0.197]
Panel C: child characteristics				
Female	0.52	0.48	0.047	[0.479]
Age	13.46	13.46	0.002	[0.979]
	(1.09)	(1.17)		
Panel D: mother characteristics				
Married/cohabiting	0.70	0.74	−0.049	[0.426]
White ethnicity	0.74	0.74	−0.001	[0.980]
Education GCSE or below	0.41	0.44	−0.040	[0.449]
Panel E: father characteristics				
Married/cohabiting	0.97	0.93	0.131	[0.134]
White ethnicity	0.75	0.83	−0.150	[0.065]
Education GCSE or below	0.46	0.40	0.074	[0.356]

Notes: The table shows means, with SDs in parentheses, and *p*-values in brackets. 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 SDs in parentheses for treated and control households, respectively. Column (3) shows normalised 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.

tests and normalised 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 the month of inspection. Table 3 shows the results: there is no statistically significant impact of the month of inspection on the rating once the 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 ($p = 0.417$ in our preferred specification in column (4)). Linking back to Figure 2(b) on inspection timing, we note that the September to December month dummies in Table 3 are also not jointly significant ($p = 0.749$). 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.

Table 3. *Ofsted Ratings and the Timing of Inspection.*

	Timing of inspection (1)	Plus prev. grade (2)	Plus timing of prev. inspection (3)	Plus school characteristics (4)	Ordered probit model (5)
Month of Ofsted inspection					
September	−0.422	−0.139	−0.129	−0.193	−0.164
	(0.276)	(0.236)	(0.242)	(0.249)	(0.363)
October	−0.148	0.130	0.106	−0.016	0.029
	(0.244)	(0.218)	(0.217)	(0.229)	(0.359)
November	−0.122	0.077	0.105	−0.059	−0.164
	(0.277)	(0.226)	(0.240)	(0.236)	(0.350)
December	−0.511	−0.137	−0.104	−0.255	0.125
	(0.326)	(0.294)	(0.281)	(0.273)	(0.383)
January	−0.311	−0.042	0.012	−0.136	0.074
	(0.249)	(0.231)	(0.244)	(0.275)	(0.361)
February	−0.312	−0.111	−0.098	−0.260	0.386
	(0.254)	(0.251)	(0.252)	(0.250)	(0.345)
March	−0.440*	−0.098	−0.042	−0.297	0.133
	(0.257)	(0.232)	(0.242)	(0.236)	(0.328)
April	0.0179	0.104	0.147	0.151	−0.197
	(0.283)	(0.243)	(0.243)	(0.229)	(0.404)
May	−0.246	0.021	0.057	−0.141	−0.099
	(0.242)	(0.204)	(0.219)	(0.226)	(0.317)
June	−0.195	−0.130	−0.059	−0.119	0.072
	(0.260)	(0.221)	(0.234)	(0.239)	(0.348)
Prev. Ofsted grade		0.385***	0.381***	0.223***	−0.520***
		(0.066)	(0.066)	(0.064)	(0.084)
F-test: month dummies	1.315	0.605	0.516	1.036	χ^2 11.82
[*p*-value]	[0.229]	[0.808]	[0.877]	[0.417]	[0.297]
F-test: Sept–Dec dummies	1.311	0.662	0.554	0.482	χ^2 1.690
[*p*-value]	[0.269]	[0.619]	[0.696]	[0.749]	[0.792]
Timing of prev. grade controls	No	No	Yes	Yes	Yes
F-test: timing of previous inspection dummies			1.031	0.836	χ^2 10.92
[*p*-value]			[0.421]	[0.595]	[0.364]
School char.	No	No	No	Yes	Yes
School FEs	Yes	Yes	Yes	Yes	No
Number of obs.	621	621	621	621	621
Number of schools	496	496	496	496	496

Notes: The dependent variable is the Ofsted grade. Columns (1) to (4) show the results of a linear regression model. All standard errors are clustered by local authority. Here *** denotes significance at 1%. 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 the 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 the month of the 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 (χ^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.

On assumption (*ii*), panels B to E of Table 2 show balance between treatment and controls on characteristics of the household, child, mother and father (Figure D2 in Appendix D shows *p*-values on balance tests and normalised differences over a wider set of outcomes). We find no imbalances. Given that heterogeneous treatment effects are central in our research design, we further show balance by type of news shock. Table E6 in Appendix E 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 D3 in Appendix D shows the cumulative distribution in the absolute difference in interview date across waves. Around 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 the month of interview (recall that Figure 2 showed variation 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, along with all the other checks of our identification assumptions, are presented in Appendix B.

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 exist for England. However, to shed light on the issue, we use the Millennium Cohort Study (MCS), a panel of children tracked since birth in 2000/01, 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 organisation, 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 presented in Appendix B, 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.[29]

4. 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 realisation 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 that beliefs about school quality and parental time investment are substitutes in the production function for child human capital. 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: $E[\Delta Y_{i\sigma t} - \Delta Y_{j\sigma' t} \mid good_{\sigma t}, good_{\sigma' t}] - E[\Delta Y_{i\sigma t} - \Delta Y_{j\sigma' t} \mid bad_{\sigma t}, bad_{\sigma' t}]$. Parents are far more likely to decrease time investment in response to good news.

Overall, the raw unconditional evidence suggests that parental time investment and beliefs over school quality are substitutes in the production function for child human capital H, so that, with the functional form assumptions in (2), $\rho > 1$ and $\partial I / \partial \mu < 0$ as shown in (4).

[29] Hussain (2015) provided 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 improve. Recall that in our sample only 7% of schools are ranked as failing (see Table 1), and our core results are robust to dropping them.

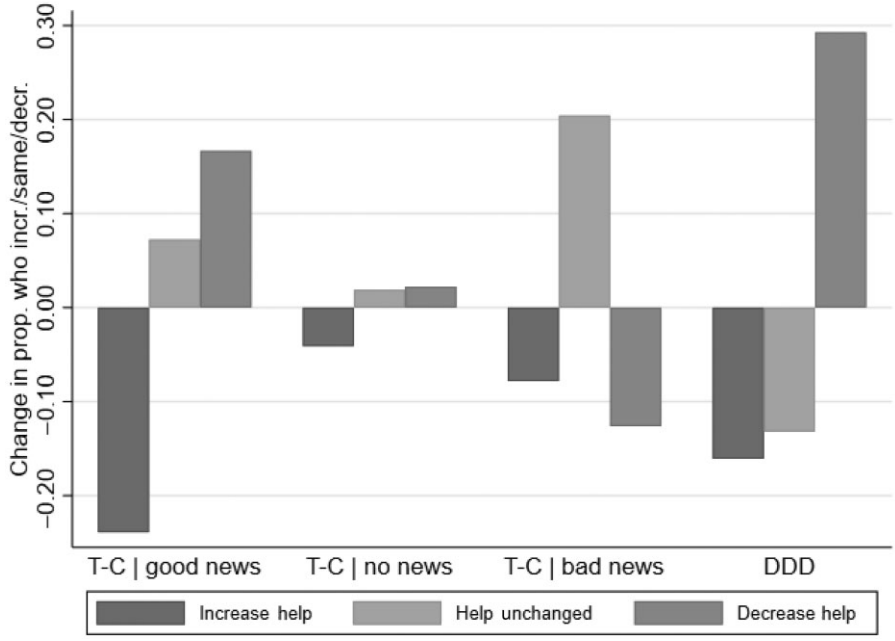


Fig. 4. *Marginal Impacts of Information on School Quality on Parental Investment.*
Notes: The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. The figure shows the 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 the treatment and control groups). The last set of bars shows the triple difference, i.e., the difference in the proportion of parents between the treatment and control groups 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.

4.1. *Regression Results*

Table 4 presents our core results that estimate (10). 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. Given that this identifies a causal impact under stronger assumptions than those discussed earlier, this is a secondary focus of our analysis.

Across the columns in Table 4 we sequentially add in covariates ($X_{i\sigma t}$, $Z_{\sigma t}$). The estimates are stable across specifications, suggesting that there is not a high correlation between these child, parent, household and school characteristics, and the influence of parental beliefs over school quality on time investments.

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 ($\widehat{\beta}_0 + \widehat{\beta}_2 = 0$).

Table 4. *Parental Response to Information on School Quality.*

	Forecast, unconditional (1)	Plus child characteristics (2)	Plus parent characteristics (3)	Plus school characteristics (4)
T-C | good news ($\beta_0 + \beta_1$)	−0.636**	−0.623**	−0.636**	−0.636**
	(0.251)	(0.251)	(0.289)	(0.289)
T-C | bad news ($\beta_0 + \beta_2$)	0.124	0.146	0.142	0.155
	(0.232)	(0.233)	(0.227)	(0.230)
Diff-in-diff-in-diff ($\beta_1 - \beta_2$)	−0.760**	−0.769**	−0.778**	−0.791**
	(0.356)	(0.355)	(0.369)	(0.370)
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	621	621	621	621

Notes: The table shows ordered probit regression estimates. Bootstrapped standard errors, clustered by local authority, are shown in parentheses. Here ** denotes significance at 5%. 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 the Ofsted inspection. The outcome variable is the change in parental help with homework between survey waves 3 and 1, or between survey waves 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 if there is 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.

This implies that, for the average family, there is substitutability between beliefs about school quality and parental time investment in the production function for child human capital.

Table E7 in Appendix E shows all coefficients from (10). We have $\widehat{\beta}_0 = 0$ 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 that $\widehat{\delta}_1 = \widehat{\delta}_2 = 0$, 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 quantify impacts, in Table 5 we report averaged marginal effect estimates from our preferred specification (column (4), Table 4). 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 that their time investment (*i*) increases falls by 15pp; (*ii*) remains unchanged falls by 9pp; (*iii*) decreases rises by 23pp. 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 differences between responses to good and bad news are significantly different, as shown by the third row ($\widehat{\beta}_1 - \widehat{\beta}_2 < 0$). If these convey similar amounts of information and are transmitted to similar kinds of households, then we can say that parents respond differentially to good and bad news. That would be consistent with parents holding unduly pessimistic beliefs over school quality, even in an education market, where parents typically make an explicit choice of school

Table 5. *Parental Response to Information on School Quality, Marginal Effects.*

	Probability of increasing parental time investment (1)	Probability parental time investment unchanged (2)	Probability of decreasing parental time investment (3)
T-C | good news ($\beta_0 + \beta_1$)	−0.147**	−0.086**	0.234**
	(0.071)	(0.038)	(0.099)
T-C | bad news ($\beta_0 + \beta_2$)	0.042	0.012	−0.055
	(0.061)	(0.020)	(0.079)
Diff-in-diff-in-diff ($\beta_1 - \beta_2$)	−0.190**	−0.099**	0.288**
	(0.095)	(0.042)	(0.127)
Forecast Ofsted rating	Yes	Yes	Yes
Child characteristics	Yes	Yes	Yes
Parent characteristics	Yes	Yes	Yes
School characteristics	Yes	Yes	Yes
Observations	621	621	621

Notes: The table shows ordered probit marginal effect estimates. Bootstrapped standard errors, clustered by local authority, are shown in parentheses. Here ** denotes significance 5%. The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and 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 the Ofsted inspection. The outcome variable is the change in parental help with homework between survey waves 3 and 1, or between survey waves 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 if there is 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.

based on expected quality (Burgess *et al.,* 2015; Agarwal and Somaini, 2018; Beuermann *et al.*, 2018).[30]

In Appendix C 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 information (Appendix Tables E8 to E10, Appendix Figures D4 and D5).

For completeness, we also 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 $\overline{S}$, or then having different forecast models. These results are summarised in Appendix Figures D6 and D7, which show marginal effect estimates from the ordered probit model. We find that the differential response to good and bad news is driven by higher educated households, non-white households, for boys and among children that are below median ability (as measured in administrative test score data).[31]

[30] 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.

[31] 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 Weinstein, 2008), or because it is more costly for them to act on their preferences. Del Boca *et al.* (2013) presented evidence from a dynamic structural model of child development suggesting ambiguous impacts of household

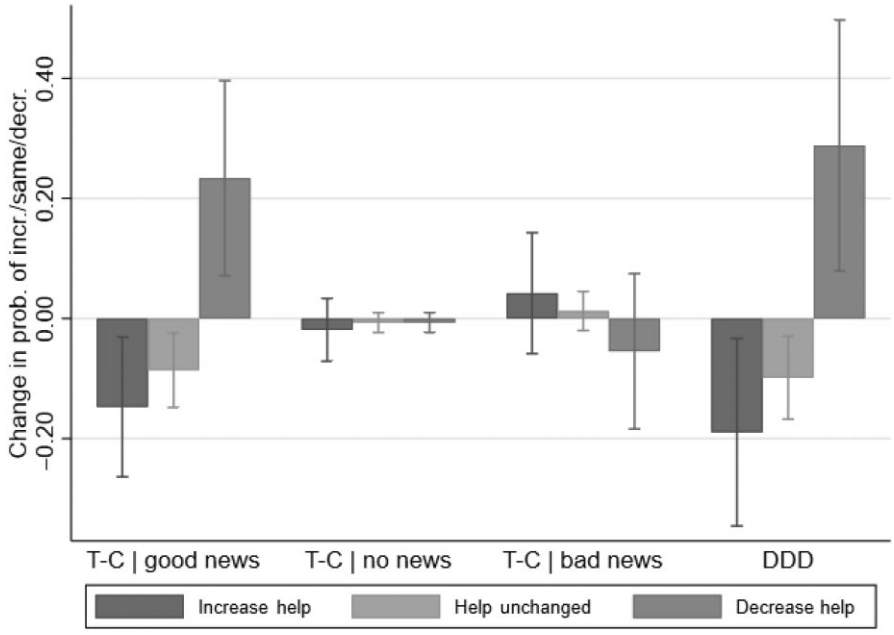


Fig. 5. *Marginal Impacts of Information on School Quality on Parental Investment.*
Notes: The sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question. Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of the Ofsted inspection. The outcome variable is the change in parental help with homework between survey waves 3 and 1, or between survey waves 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 if there is 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.

4.2. *Distributional Impacts*

Our results show that parents respond to news about school quality. A key consequence is that the inspection regime has distributional effects, 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 a by documenting $prob(news_{\sigma t}|rating_{\sigma t-1} = a)$ for each value of $news_{\sigma t}$ and school rating a. Columns (1) to (3) of Table 6 show these descriptives. Column (1) shows that schools previously rated as outstanding are more likely than other schools to receive a positive news shock: $prob(news_{\sigma t} = good_{\sigma t}|rating_{\sigma t-1} = \text{outstanding}) = 0.28$. We note that $prob(news_{\sigma t} = good_{\sigma t}|rating_{\sigma t-1} = a)$ is monotonically decreasing in initial school rating (column (1)). This suggests some high

income on child development. The reason is that a higher income often means greater labour supply and reductions in time investments into children. In line with our results, Autor *et al.* (2016) used data on opposite-sex siblings attending Florida public schools to document how boys benefit more from cumulative exposure to higher quality schools.

Table 6. *The Distribution of News and Parental Time Investments, by Previous Inspection Rating.*

Ofsted rating from earlier inspection ($rating\sigma_{t-1} = a$)	News			Share of schools	Parental time investment
	$P(news\sigma_t = good\sigma_t \mid rating\sigma_{t-1} = a)$ (1)	$P(news\sigma_t = nonews\sigma_t \mid rating\sigma_{t-1} = a)$ (2)	$P(news\sigma_t = bad\sigma_t \mid rating\sigma_{t-1} = a)$ (3)		Unconditional (4)
Outstanding	0.28	0.50	0.22	0.18	0.84
Good	0.18	0.60	0.22	0.41	0.82
Requires improvement	0.16	0.66	0.19	0.38	0.80
Inadequate	0.11	0.83	0.06	0.04	0.65
Share of schools	0.19	0.61	0.20		
Observations	93	303	100		2,529

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 shocks (good news/no news/zero news) in columns, by the schools' Ofsted grade in the previous inspection cycle, in each row. The sample in column (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 the 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).

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 $prob(news_{\sigma t} = bad_{\sigma t} | rating_{\sigma t-1} = a)$, column (3) shows that schools previously rated outstanding or good are most likely to receive bad news (with probabilities around 0.22), but this is closely followed by schools previously rated as requiring improvement (0.19). Schools previously rated inadequate are least likely to receive bad news (0.06). This is in line with Hussain (2015), who found 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 in all schools pre-inspection. 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 pre-treatment (period $t - 2$) levels of parental input. We find a positive gradient of parental input with regard to school ratings: in outstanding schools, 84% of parents provide high levels of input into their children, and this falls to 65% in the lowest rated schools.

Third, we combine these probabilities with other estimates from the data to calibrate implied impacts on (*i*) expected parental inputs, $E[Y]$; (*ii*) pre-inspection between-school inequality in parental inputs between high and low inspection ratings (s_L, s_H), denoted Q, where a low rating s_L corresponds to inadequate, and a high rating s_H corresponds to outstanding; (*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

$$E[Y] = \sum_a prob(rating_{\sigma t-1} = a) E[Y | rating_{\sigma t-1} = a],$$

$$Q = E[Y | rating_{\sigma t-1} = s_H] - E[Y | rating_{\sigma t-1} = s_L]. \tag{11}$$

Pre-inspection, the expected parental input is $E[Y] = 0.82$ and the inequality in parents' investments is $Q = 0.19$ (see Table 6). The treatment (T) effects on these are

$$\begin{aligned} \frac{\partial E[Y]}{\partial T} &= \sum_a prob(rating_{\sigma t-1} = a) \frac{\partial E[Y | rating_{\sigma t-1} = a]}{\partial T} \\ &= \sum_a prob(rating_{\sigma t-1} = a) \\ &\quad \times \left\{ \sum_j prob(news_{\sigma t} = j | rating_{\sigma t-1} = a) \frac{\partial E[Y | news_{\sigma t} = j]}{\partial T} \right\}, \end{aligned} \tag{12}$$

$$\frac{\partial Q}{\partial T} = \frac{\partial E[Y | rating_{\sigma t-1} = s_H]}{\partial T} - \frac{\partial E[Y | rating_{\sigma t-1} = s_L]}{\partial T}, \tag{13}$$

where $rating_{\sigma t-1} = a = \{$outstanding, good, requires improvement, inadequate$\}$ and $news_{\sigma t} = j = \{$good, zero, bad$\}$.[32] Each element in (12) can be substituted in for either using the evidence in Table 6 ($prob(rating_{\sigma t-1} = a)$, $prob(news_{\sigma t} = j | rating_{\sigma t-1} = a)$) or from other estimates from our working sample ($\partial E[Y | news_{\sigma t} = j] / \partial T$).

[32] Our research design assumed that the response to news was homogeneous across schools, so $\partial E[Y | rating_{\sigma t-1} = a, news_{\sigma t} = j] / \partial T = \partial E[Y | news_{\sigma t} = j] / \partial T$. Appendix Figure D8 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.

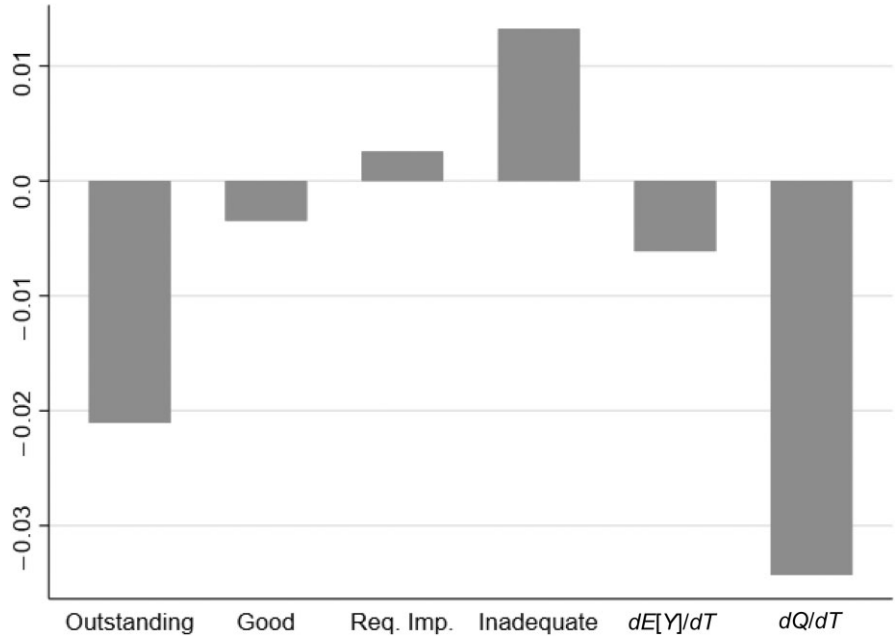


Fig. 6. *Distributional Impacts of the Schools Inspection Regime on Parental Investment.*
Notes: The figure shows the overall effect of the inspection regime on parents' investments, on the level of investment (first five bars) and inequality of investments (final bar). First, we define a high level of parental input ($Y = 1$) if the frequency of help with children's homework is almost every day or at least once a week. Second, we calculate the difference in Y between the treated and control households for each news shock. Third, we calculate the expected treatment effect for households in schools with each previous inspection rating. The first four bars show this: the treatment effect on parents' investments, given each previous inspection rating ('Outstanding' to 'Inadequate'). The final bar shows the weighted average of these: the overall treatment effect on the level of investment (equivalent to (12)). The final bar shows the treatment effect on the inequality in investments between those in schools previously rated as 'Outstanding' and 'Inadequate'. This is equivalent to (13).

Doing so yields the following calibration, summarised in Figure 6. The first four bars show, for schools of pre-inspection rating $rating_{\sigma t-1} = a$, the unconditional treatment effect on parental inputs. For outstanding and good schools, parental inputs fall overall in response to the inspections regime, with the largest falls occurring in schools ranked outstanding. This is consistent with the pattern that outstanding schools are most likely to receive a positive shock, and our main finding that parental investments decrease in response to a positive shock. For requires improvement and inadequate schools, parental investments increase, with the largest increase in schools ranked inadequate. Again, this is consistent with the patterns in our results. Aggregate parental inputs then fall marginally overall (as shown in the fifth bar on $\partial E[Y]/\partial T$). As inputs fall more in the highest ranked schools ($rating_{\sigma t-1} =$ outstanding) than in the lowest ranked schools ($rating_{\sigma t-1} =$ inadequate), input inequality also falls with the inspection regime using these to define the range Q. Just comparing differences between outstanding and inadequate schools we find that $Q = 0.19$ and $\partial Q/\partial T = -0.034$, so across-school inequality in parental inputs falls by 18% because of the information generated by the inspections regime.

5. Test Score Impacts

We now build on our core results to understand the test score impacts of news on school quality. Following the framework laid out in Section 1.3, parental time is not the only margin of response to such news. The UKHLS allows us to study a range of behaviour, 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 summarised in Figure 7, which shows marginal 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. These changes 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.[33]

Panel (b) shows changes in the amount of time children themselves report spending on their homework, so mapping to I^C in the conceptual framework.[34] Children's time inputs are less responsive than parents: there are no statistically significant changes in children's time investments in response to news. When a household receives good news about school quality, the estimates have the opposite sign to parents, however. This is suggestive that children partly compensate for the loss of parental inputs by increasing their own time spent on homework. When a household receives bad news about school quality, the point estimates suggest that children are more likely to increase the time spent on homework (in the same direction as parents), but this is estimated imprecisely.[35]

5.1. *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 NPD, which records high-stake nationwide exam scores, taken at ages 11 and 16. We focus on students enrolled in schools inspected in the 2011, 2012 or 2013 academic year, and who were taking high-stake General Certificate of Secondary Education (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 inspections.[36]

We estimate a value added model for test scores:

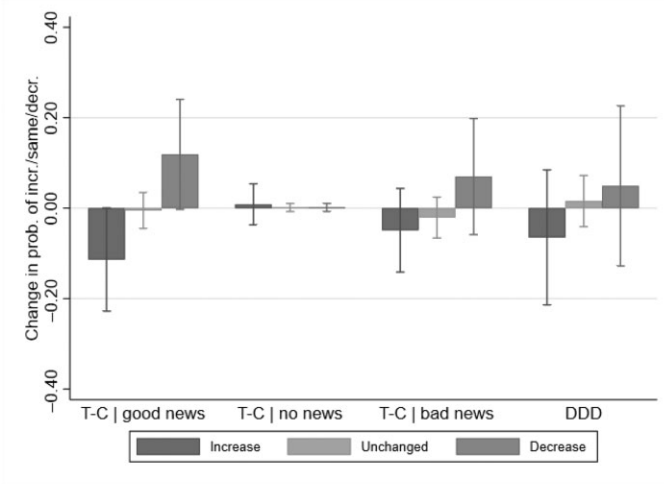
$$y_{i\sigma t} = \rho y_{it-1} + \beta_0 T_{\sigma t} + \beta_1 [T_{\sigma t}.good_{\sigma t}] + \beta_2 [T_{\sigma t}.bad_{\sigma t}] + \delta_1 good_{\sigma t} + \delta_2 bad_{\sigma t} + \mu_t + \gamma_0 X_{i\sigma t} + \gamma_1 Z_{\sigma t} + \varepsilon_{\sigma t}.$$

[33] We note that we can also examine additional margins of parental-child interactions. We find, for example, that changes in parental beliefs over whether they think A-levels (high-stake nationwide exams taken at age 18) are important, are not impacted by news on school quality, nor are changes in how interested parents report being in how well the child does at school. We do find some evidence that parents become less likely to attend parent evenings at school if they receive bad news on school quality.

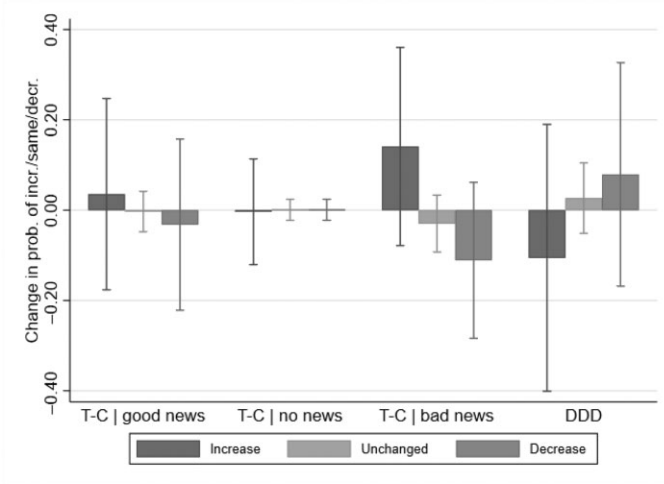
[34] The change in hours the child spends doing homework is derived from the following question in waves 2, 4 and 6: 'When you do homework on a week-day evening during term time, how many hours do you usually spend doing your homework?' ($N = 252$). We convert ΔY_{ist} into an increase, decrease or no change in time between waves.

[35] De Fraja *et al.* (2010) studied 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 found them to be complements.

[36] 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).



(a) Child talks with parents about things that matter



(b) Child own time investment: hours spent doing homework

Fig. 7. *Academic Responses to School Quality Information.*
Notes: Treated (control) households are defined as those whose UKHLS interview occurs after (before) the date of the 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. The change in how often a 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?' ($N = 664$). In panel (b) the sample is based on UKHLS households with an Ofsted school inspection in the same academic year as the interview, and with a non-missing outcome variable collected from the young person at waves 2, 4 and 6. 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?' ($N = 252$). 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.

Table 7. *Test Score Impacts of the School Inspection Regime.*

	Student level	School level	
	Standardised GCSE average point score (1)	SD (2)	Interquartile range (3)
T-C | good news ($\beta_0 + \beta_1$)	−0.083*	0.008	0.014
	(0.043)	(0.017)	(0.029)
	[−0.153, −0.012]	[−0.020, 0.037]	[−0.035, 0.062]
T-C | bad news ($\beta_0 + \beta_2$)	−0.041	0.009	0.004
	(0.058)	(0.015)	(0.027)
	[−0.137, 0.054]	[−0.016, 0.033]	[−0.041, 0.048]
Forecast Ofsted rating	Yes	Yes	Yes
School characteristics	Yes	Yes	Yes
Pupil characteristics	Yes	Yes	Yes
Number of schools	1,143	1,133	1,133
Number of pupils	203,503	–	–

Notes: The table shows OLS regression estimates. Bootstrapped standard errors, clustered by local authority, are shown in parentheses. The 90% confidence interval is shown in brackets. Here denotes significance at 10%. The sample comprises schools inspected by Ofsted in the 2011, 2012 or 2013 academic year. 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 standardised average point score on the age-16 GCSE exam. In columns (2) and (3) the dependent variables are the school-level SD in test scores and the interquartile range, respectively. For column (1), the regression includes students' lagged test scores (student's age-11 Key Stage 2 performance). For columns (2) and (3), regressions include the lagged school-level SD 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 the type of school (community, academy, voluntary aided, etc.), the school's religious status, the school's admission policy, single-sex entry, the percentage of students eligible for free school meals, the percentage of students speaking English as an additional language and total enrolment. Column (1) also includes student's eligibility for a free lunch, an 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). In columns (2) and (3), 10 schools without lagged test score information are dropped in the estimation sample; for robustness, we report a student-level analysis for this sample in Table E12 in Appendix E. 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.

Here $y_{i\sigma t}$ is student i's standardised average point score on the age-16 GCSE exams at the end of the academic year and 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 took place later in the year (January through April). Variables $good_{\sigma t}$ and $bad_{\sigma t}$ are as previously defined; μ_t is an 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.[37]

The results are in Table 7, where we show the DD coefficients of interest ($\beta_0 + \beta_1$, $\beta_0 + \beta_2$) and 90% confidence intervals on each. As with our earlier research design, these identify causal impacts of news if treatment assignment ($T_{\sigma t}$) is as good as random.

[37] 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.

Column (1) shows that, 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-stake exams ($\widehat{\theta_1} = -0.083sd$). The 90% confidence interval rules out any impact larger than $-0.008sd$. Bad news has no significant impact on test scores.

Recall that the earlier findings suggested that good news causes parents to reduce their time input and children to (marginally) increase their time input. Using the framework above and condition (6), the negative net impact on test scores ($dA/d\mu < 0$) suggests that 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 measured the relative productivity of parental and child investments (Del Boca *et al.*, 2017; Caetano *et al.*, 2019).[38]

Using school level outcomes in the SD 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.[39] Given the earlier results on the distributional impacts on parental inputs of the inspections regime, this suggests that, among those students whose schools are inspected early in the academic year of their high-stake exams, the inspection regime slightly decreases educational attainment overall and decreases inequality in test scores between high- and low-quality schools (but not within a school).

A key lesson is that children's behaviour, independent of their parents, can complicate the analysis of household-level behaviour and drivers of children's test scores. The fall in test scores as a result of parents and children receiving good news over school quality, is hard to reconcile with a unitary household model in which parents' and children's interests are perfectly aligned. This can be explained as a result of imperfect information of parents and children, so their combined responses to news on school quality can potentially lead them to make mistakes that reduce children's human development (at least in the short run as our design allows us to measure).

For example, parents might be imperfectly informed about the skills of their own child, or the marginal productivity of the child's own time investment ($\partial g/\partial I^C$). A growing evidence base suggests that parental investments into children are related to their beliefs over child skill and the productivity of various inputs into the production function for a child's human capital (De Fraja *et al.*, 2010; Boneva and Rauh, 2018; Dizon-Ross, 2019; Attanasio *et al.*, 2020a; 2020b). These typically find that parents have upwards biased beliefs about their children's skills or academic performance (Dizon-Ross, 2019; Bergman, 2021; Kinsler and Pavan, 2021).

Alternatively, if inspection ratings reflect a broad measure of school quality (as is the aim of the Ofsted inspection body), then a fall in test scores may be mitigated by a rise in non-cognitive outcomes that are also valued by parents and children (Beuermann *et al.*, 2018). An alternative possibility is that raised by MacLeod and Urquiola (2015), in that school reputation matters (as a signal to universities or employers). Hence, if school quality is better than expected, there may be scope for families to shave on the test score margin. Relatedly, if a school is performing better than expected on the national distribution of school quality, then parents may feel that they can reduce investments to generate test scores (Kinsler and Pavan, 2021).

[38] Del Boca *et al.* (2017) and Caetano *et al.* (2019) examined 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) found that child time investments are more productive than maternal time investments; Caetano *et al.* (2019) found them to be equally productive (with grandparents active time investment being the most productive input).

[39] 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 (Neal and Schanzenbach, 2010; Feng *et al.*, 2018).

Moreover, we reiterate that our research design allows us to only estimate impacts within the academic year. We cannot rule out that, over years, parents (and children) update and adjust their investments further, to leave test scores unchanged or improved in the long run. For example, Pop-Eleches and Urquiola (2013) showed 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 (Figlio and Loeb, 2011; Feng *et al.*, 2018; Dizon-Ross, 2020). Hence, any longer-term analysis would have to distinguish between household and school responses to information across academic years.

6. Conclusion

We extend the voluminous literature studying family and school inputs to investigate interactions between family inputs and parental *beliefs* over school quality. We do so by identifying the causal impact of exogenously released new information on school quality, on parental time investments into their children, on children's own time investments and the ultimate impact these multiple household responses have on the high-stake test scores of 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 is consistent with beliefs over school quality and parental inputs being substitutes in the production function for child human capital. In our setting, we have focused on parental time investments. Future work can try to exploit a richer array of parental inputs, to understand whether they all respond to news on school quality in the same way.

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 organised and financed (Albornoz *et al.*, 2017). These biases arise irrespective of whether estimates are based on experimental or observational data (Fu and Mehta, 2018). If behavioural responses of families to the same change in school inputs vary across contexts, this limits the external validity of any given study, and leads to conflicting results in a given literature.

Indeed, as Pop-Eleches and Urquiola (2013) and Albornoz *et al.* (2017) reviewed, there is an extensive literature examining the impact of school quality on test scores, but this has produced mixed findings.[40] The insight that interactions between parental beliefs about school quality and

[40] 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 Kirabo Jackson (2010). These found that marginal

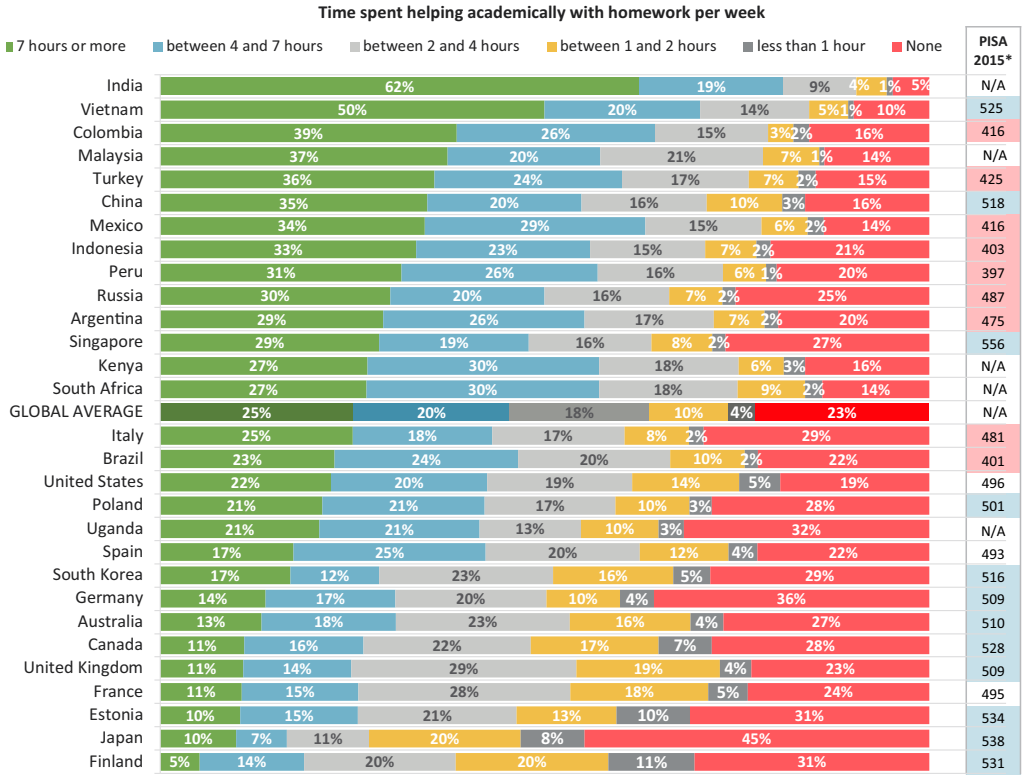


Fig. 8. *Global Survey Of Parents. 'On average, how much time, if any, do you personally spend helping your child academically with their education per week?'*

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 the figure are weighted by age, gender and region of the child and corrected for the gender of the 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 are representative of the urban online population.

investments in their children are important offers the possibility of reconciling a disparate set of results across the literature. The critical issue is variation in household behavioural responses to school quality across different contexts.

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 suggest that the kinds of issue we document in the English context will be even more relevant in these new settings. Figure 8 shows evidence

students just gaining admission to high achievement educational institutions have better academic and labour market outcomes. Other papers, however, found weaker evidence that school quality matters, including Cullen *et al.* (2006), Dobbie and Fryer (2011), Duflo *et al.* (2011), Abdulkadiroğlu *et al.* (2014), and Clark and Del Bono (2016).

from a global survey of parents conducted in 2017/18. Across countries at various stages of economic development, parents in lower income countries provide more time inputs into their children. This suggests that household responses to information about school quality may be even larger in low-income settings.[41]

In advancing understanding of the role of parent and child beliefs in the production of human capital among school age children, our results open up a broad agenda to study the framing, targeting and specifics of information provision about schools, with the ultimate aim of increasing efficiency in education markets, and fostering the development and well being of adolescents.

Appendix A. Derivations

To derive the first order condition (3), we start with the Lagrangian for the parental optimisation problem:

$$L = \theta \ln(C) + (1-\theta)\ln(H) + \lambda(w - wI - C).$$

This maximisation problem yields the first-order conditions

$$\frac{\theta}{C} = \lambda, \qquad \frac{(1-\theta)}{H}\frac{\partial H}{\partial I} = \lambda w.$$

Substituting H into the latter expression yields

$$(1-\theta)\frac{1}{a}[\gamma(\overline{S}+\mu)^{\rho} + (1-\gamma)I^{\rho}]^{-1/\rho}\frac{1}{\rho}[\gamma(\overline{S}+\mu)^{\rho} + (1-\gamma)I^{\rho}]^{1/\rho-1}\rho(1-\gamma)I^{\rho-1} = \lambda w.$$

Substituting out the Lagrange multiplier from the two first-order conditions yields

$$\frac{(1-\theta)(1-\gamma)I^{\rho-1}}{a[\gamma(\overline{S}+\mu)^{\rho} + (1-\gamma)I^{\rho}]} = \frac{\theta}{C}w.$$

By substituting C using the budget constraint, we obtain the expression in (3):

$$\frac{(1-\theta)(1-\gamma)I^{\rho-1}}{a[\gamma(\overline{S}+\mu)^{\rho} + (1-\gamma)I^{\rho}]} = \frac{\theta}{1-I}.$$

To derive the impact of a school quality information shock, $\partial I/\partial\mu$, we begin by rearranging the first-order condition,

$$a[\gamma(\overline{S}+\mu)^{\rho} + (1-\gamma)I^{\rho}] = \frac{(1-\theta)(1-\gamma)}{\theta}(I^{\rho-1} - I^{\rho}),$$

and then taking the partial differential of this expression with respect to the school quality information shock, μ:

$$\frac{(1-\theta)(1-\gamma)}{\theta}\left[(\rho-1)I^{\rho-2}\frac{\partial I}{\partial\mu} - \rho I^{\rho-1}\frac{\partial I}{\partial\mu}\right] = a\rho\gamma(\overline{S}+\mu)^{\rho-1} + a\rho(1-\gamma)I^{\rho-1}\frac{\partial I}{\partial\mu}.$$

[41] 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.

From the first-order condition, we have

$$\frac{(1-\theta)(1-\gamma)}{\theta} = \frac{a[\gamma(\overline{S}+\mu)^{\rho} + (1-\gamma)I^{\rho}]}{I^{\rho-1} - I^{\rho}}.$$

Thus,

$$a\rho\gamma(\overline{S}+\mu)^{\rho-1}(I^{\rho-1} - I^{\rho}) + a\rho(1-\gamma)I^{\rho-1}(I^{\rho-1} - I^{\rho})\frac{\partial I}{\partial \mu}$$

$$= a(\rho-1)I^{\rho-2}[\gamma(\overline{S}+\mu)^{\rho} + (1-\gamma)I^{\rho}]\frac{\partial I}{\partial \mu}$$

$$- a\rho I^{\rho-1}[\gamma(\overline{S}+\mu)^{\rho} + (1-\gamma)I^{\rho}]\frac{\partial I}{\partial \mu},$$

which simplifies to the expression given in (4):

$$\frac{\partial I}{\partial \mu} = \frac{-\rho\gamma(\overline{S}+\mu)^{-1}I^{\rho-1}(1-I)}{\gamma + (1-\gamma)(\overline{S}+\mu)^{-\rho}I^{\rho} - \rho\gamma(1-I)}.$$

Appendix B. Evidence in Support of the Identifying Assumptions

B.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 (10). Column (1) of Table E8 in Appendix E shows that the results are robust to the inclusion of month of interview dummies, and these dummies are not jointly significant ($p = 0.964$). One limitation of this check is that there is not a complete overlap in the month of interview between the control and treatment groups (as Figure 2(c) already showed). To probe this further, our next check then also includes schools without inspections in order for these month effects to be more precisely identified. The result in column (2) is in line with the earlier check, with the month dummies not being jointly significant ($p = 0.966$).

Third, 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 (3) shows that these future inspection ratings have no relationship with changes in parental investment the year before.

B.2. Within-Year School Responses

We now provide evidence in relation to assumption (*iv*), that there are no within-year school responses to Ofsted ratings. We first reiterate that Hussain (2015) showed 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 (4) of Table E8 in Appendix E shows that our core results are robust to dropping them.

Schools still might be able to adjust on various other margins in the short run. No data on fine-grained adjustments in secondary schools exist for England. To thus shed light on the issue,

we use the MCS, a panel of children tracked since birth in 2000/01, 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 (MCS, 2017).

This linkage covers MCS schools with an Ofsted inspection in academic years 2007/08 or 2008/09. Our working sample comprises 734 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 87, as primary schools are smaller than secondary schools from the UKHLS data. Nineteen percent 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 the 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 before that estimate a school's rating as a function of its past rating, school characteristics and past performance. We construct $news_{st}$ as in (7).

Finally, we estimate a specification analogous to (10) 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 E9 in Appendix E 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.

To create an underlying measure of teacher effort that combines the indicators to reduce measurement error and improve the power of the test, the outcomes in columns (7) and (14) are indices composed of similar dimensions of teacher response. For example, the time use index shown in column (7) combines outcomes from columns (1), (4), (5) and (6). Each index is standard normalised, and so the coefficients can be easily interpreted as effect sizes. We continue to find null impacts of news on school quality on these indices of short-run teacher responses related to time allocations or teaching practices.

Appendix C. Robustness Checks

We present a battery of robustness checks on our core result. To begin with, column (5) of Table E8 in Appendix E examines 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 that 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 (10); 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.

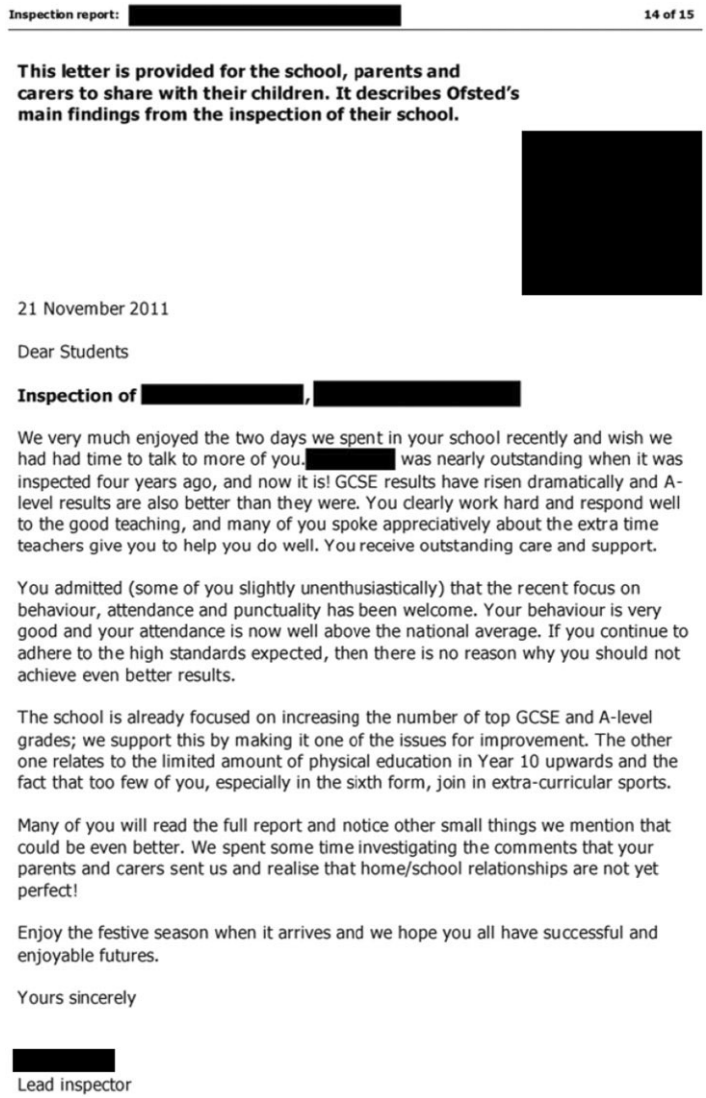
Table E10 in Appendix E then probes the robustness of the core result to using an alternative econometric approach. More precisely, we use a linear probability model for two outcomes: (*i*) whether the frequency of parental help with homework increases between $t-2$ and t (panel A); (*ii*) whether the 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 in the production function for children's human capital.

We next examine how our results are impacted if we utilise the full range of information available on parental time investments or on the extent of good and bad news parents receive on school quality. More precisely, on the former, we can define the change in parental time investment, $\Delta I_t = I_t - I_{t-2}$, as the difference in the two five-point Likert scales measured for the household over time (either between waves 1 and 3, or between waves 3 and 5). The resulting findings are shown in panel A of Figure D4(a): we again see that the response to good news is generally to significantly decrease parental time investments.

On the latter check, we move away from the definition of $news_{\sigma t}$ given in (9) and just use the full range of the (actual rating − predicted $\text{rating}_{\sigma t}$). As documented earlier, this ranges from −2 to +1, and so we can now refer to parents receiving bad news or very bad news. The resulting findings are shown in panel B of Figure D4(b): we again see that the response to good news is generally to significantly decrease parental time investments. The point estimates on responses to bad news or to very bad news are similar, although we lose precision with this finer definition of news.

Finally, we examine the sensitivity of the results to alternative forecasting models, thus allowing the assumed underlying information set parents use to vary. Table E11 in Appendix E presents these results, and Figure D5 plots the corresponding sets of marginal effects from each model. In column (1) we assume that parents use an AR(1) model that only conditions on past rating. Column (2) adds school characteristics, and column (3) adds school performance measures (our baseline specification). Column (4) 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 that our core result is robust to any small misspecification in the forecasting model.

Appendix D. Appendix Figures



Inspection report: [redacted] 14 of 15

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.

[redacted]

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.

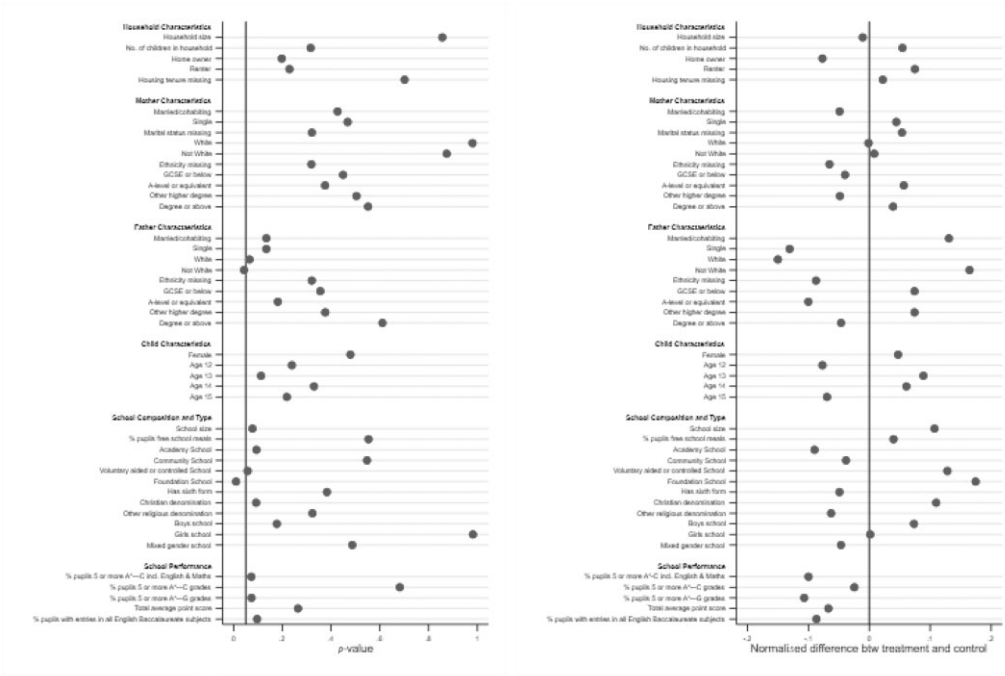
Many of you will read the full report and notice other small things we mention that could be even better. We spent some time investigating the comments that your parents and carers sent us and realise that home/school relationships are not yet perfect!

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

Yours sincerely

[redacted]
Lead inspector

Fig. D1. *Example Ofsted Letter to Parents.*



(a) Balance (b) Normalised differences

Fig. D2. *Balance and Normalised 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 the Ofsted inspection. The left-hand panel shows *p*-values for the test on the equality of household, child, mother, father and school characteristics across the treated and control households. These are derived by regressing characteristics on the treatment dummy and clustering standard errors by local authority. The vertical line indicates a *p*-value of 0.05. The right-hand panel rows display normalised difference of the means of household, child, mother, father and school characteristics between the treatment and control groups, derived by dividing the raw mean difference by the square root of the sum of the variances.

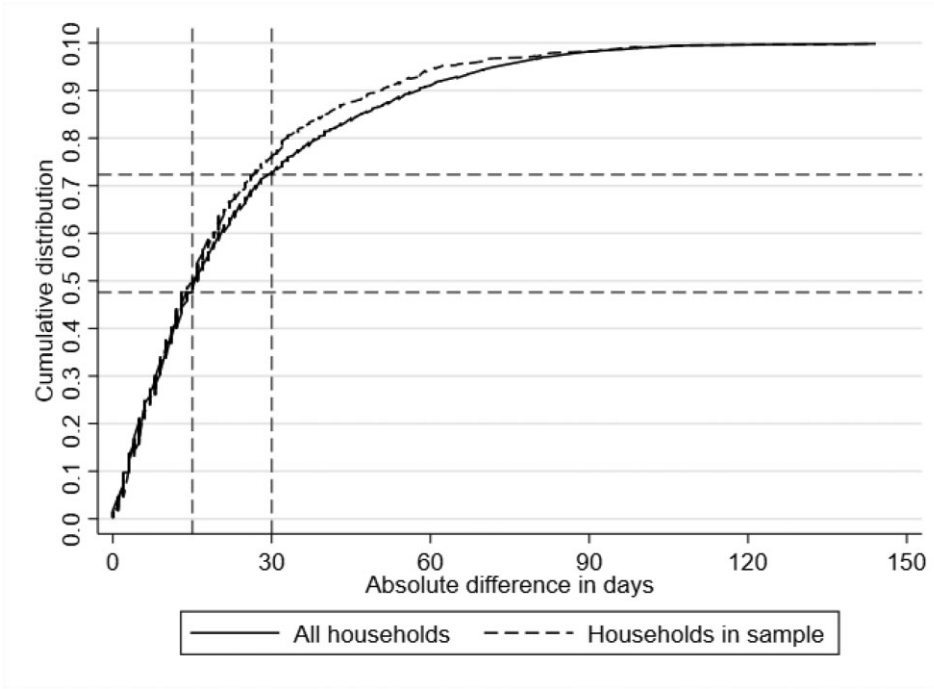
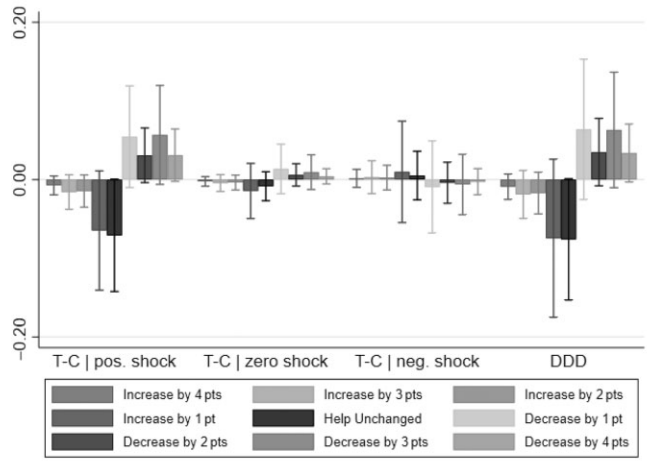
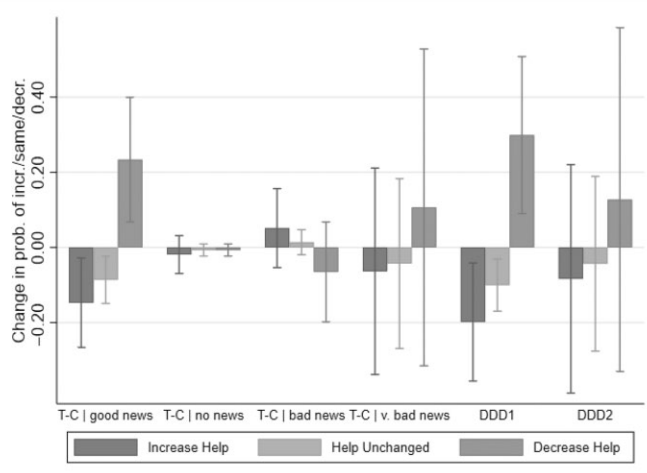


Fig. D3. *Absolute Difference in Interview Dates between Survey Waves.*
Notes: Two household samples are used. The first are all those observed in the UKHLS across consecutive waves in waves 1, 3 and 5 ($N = 4{,}660$). The second sample comprises UKHLS households that have an Ofsted school inspection during the academic year of interview, and have non-missing changes in the parental help with homework question ($N = 621$). The figure shows the cumulative distribution in the absolute difference in the date of interview at waves 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.



(a) Using the full range of change in parental help with homework



(b) Using the full range of news shocks on school quality

Fig. D4. *Using the Full Range of Variation in Help with Homework and News Shocks.*
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 the Ofsted inspection. The outcome variable in panel (a) is the change in parental help with homework between survey waves 3 and 1, or between survey waves 5 and 3: this is measured on a five-point Likert scale and can decrease by up to four points if parents decrease their help with homework, or increase by up to four points if they increase their help. Panel (b) codes the change in parental help as -1 if there is any decrease in parental help, 0 if there is no change in parental help and $+1$ if there is any increase in parental help with homework. Panel (b) uses the full range of the news shock, where good news is an inspection outcome one Ofsted grade better than expected, bad news is one Ofsted grade worse than expected and very bad news is two Ofsted grades worse than expected. 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.

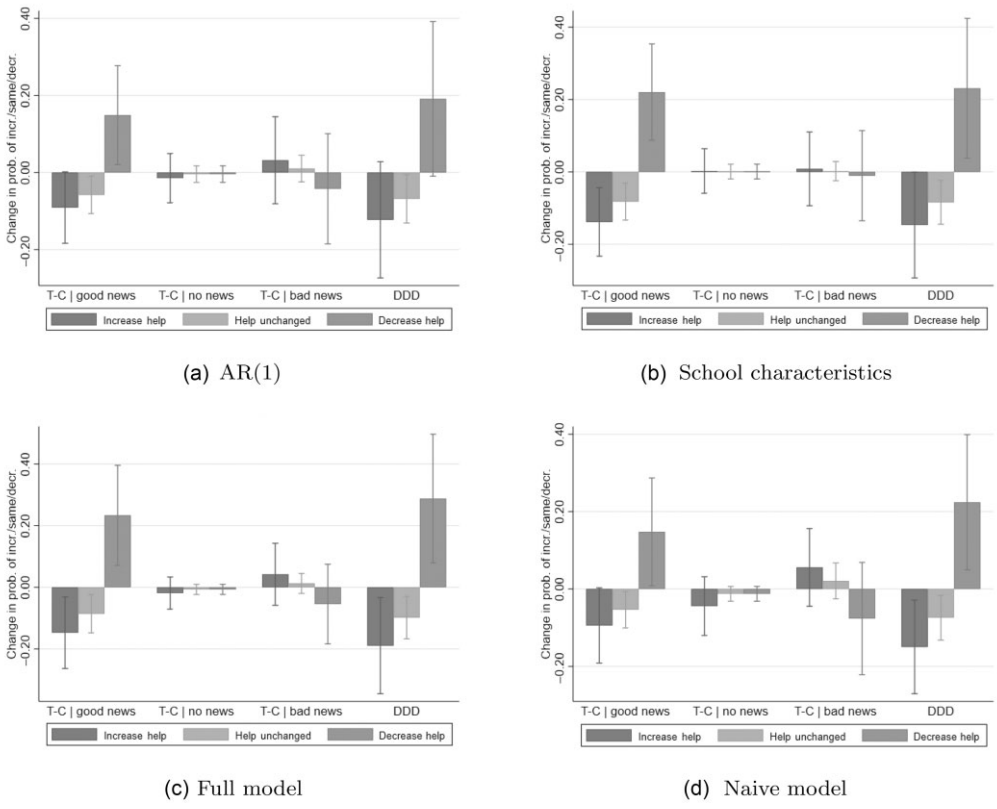


Fig. D5. *Marginal Impacts of Information on School Quality on Parental Investment, by Forecasting Model.*

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 the Ofsted inspection. The outcome variable is the change in parental help with homework between survey waves 3 and 1, or between survey waves 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 if there is an increase in parental help with homework. The panels vary in the underlying forecasting model used. Panel (a) uses a separate ordered probit regression for each past Ofsted inspection grade to forecast the Ofsted rating and controls for local authority fixed effects. Panel (b) additionally controls for schools characteristics; 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.

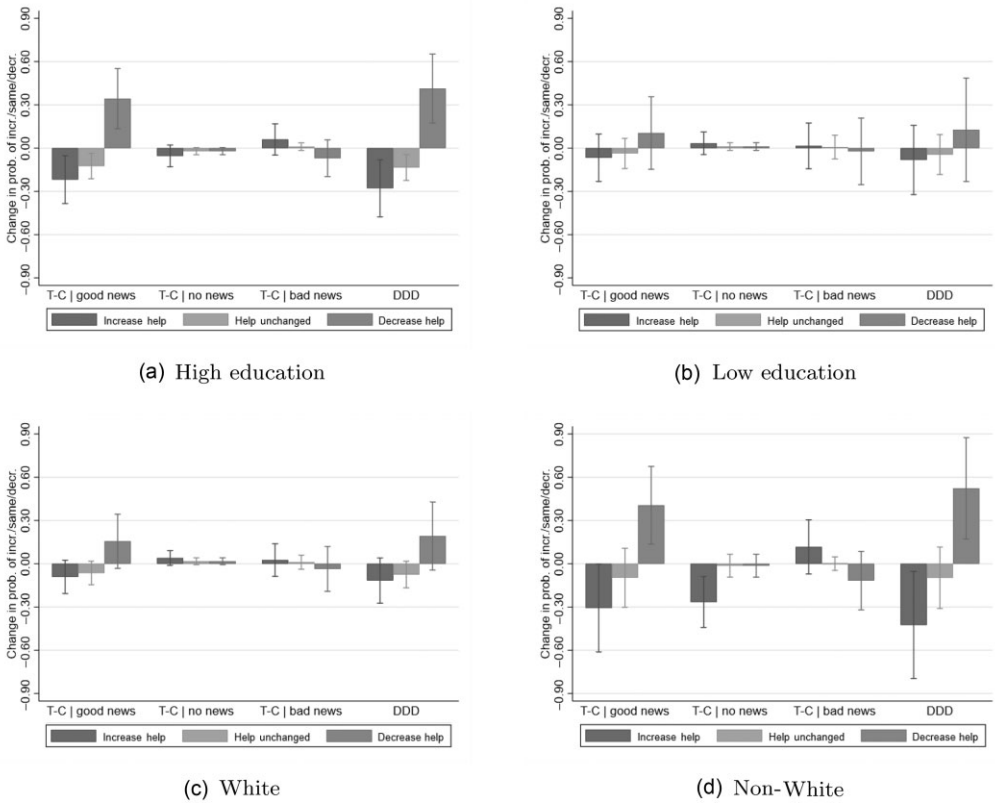


Fig. D6. *Marginal Impacts of Information on School Quality on Parental Investment, by Family Characteristic.*

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 the Ofsted inspection. The outcome variable is the change in parental help with homework between survey waves 3 and 1, or between survey waves 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 $+1$ if there is an increase in parental help with homework. Panels (a) and (b) split the sample into households where parents are highly educated, defined as having an A-level or higher education (panel (a), $N = 359$) or low educated, defined as having GCSEs or no qualification (panel (b), $N = 262$). Panels (c) and (d) split households by ethnicity (White, $N = 462$, Non-White, $N = 159$). 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.

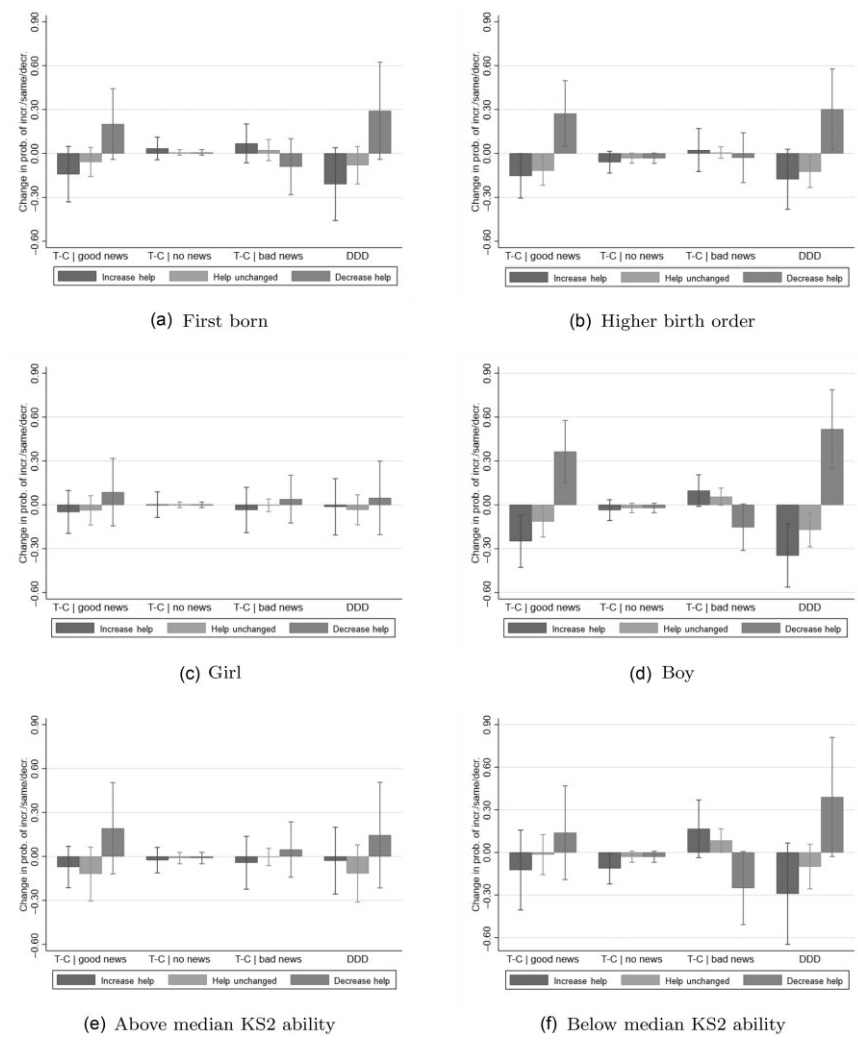
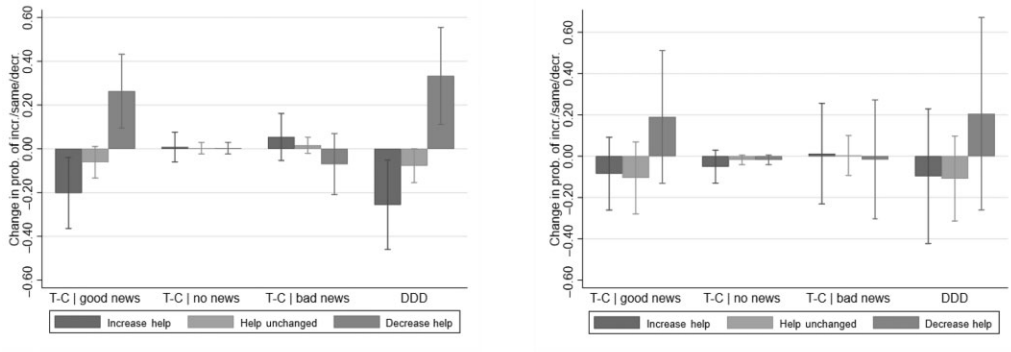


Fig. D7. *Marginal Impacts of Information on School Quality on Parental Investment, by Child Characteristic.*

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 the Ofsted inspection. The outcome variable is the change in parental help with homework between survey waves 3 and 1, or between survey waves 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 $+1$ if there is an increase in parental help with homework. Panels (a) and (b) split the sample by whether the child is a first born ($N = 294$) or a higher-order child ($N = 327$). Panels (c) and (d) split the sample by whether the child is a girl ($N = 313$) or a boy ($N = 308$). Panels (e) and (f) split the sample by whether the child had above median KS2 ability, measured as the average of maths and English fine points in national curriculum tests ($N = 176$) or below median KS2 ability ($N = 199$). 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.



(a) Past Ofsted rating outstanding/good

(b) Past Ofsted rating requires improvement/inadequate

Fig. D8. *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 the Ofsted inspection. The outcome variable is the change in parental help with homework between survey waves 3 and 1, or between survey waves 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 $+1$ if there is 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), $N = 363$) or were last rated as requires improvement or fail (panel (b), $N = 258$). 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.

Appendix E. Appendix Tables

Table E1. *Key Ofsted 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.

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 cyberbullying 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: https://dera.ioe.ac.uk/14076/1/The_evaluation_schedule_for_school_inspections_from_January_2012%5B1%5D.pdf.

Table E2. *Key Ofsted Grade Descriptors.*

Rating	Achievement of pupils at the school	Quality of teaching in the school
Outstanding	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, of any group of pupils are below those of all pupils nationally, the gap is closing dramatically over a period of time, as shown by a wide range of attainment indicators.	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.

Table E2. *Continued*

Rating	Achievement of pupils at the school	Quality of teaching in the school
Good	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, cyberbullying 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 channelling 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.

Table E2. *Continued*

Rating	Achievement of pupils at the school	Quality of teaching in the school
Requires improvement	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, cyberbullying 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.

Table E2. *Continued*

Rating	Achievement of pupils at the school	Quality of teaching in the school
Inadequate	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, cyberbullying 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.

Table E2. *Continued*

Rating	Behaviour and safety of pupils at the school	Quality of leadership in and management of the school
Outstanding	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, cyberbullying 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.	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 effectively check pupils' understanding throughout lessons, anticipating where they may need to intervene and doing so with notable impact on the quality of learning. Time is used very well and every opportunity is taken to successfully develop crucial skills, including being able to use their literacy and numeracy skills in other subjects. Appropriate and regular homework contributes very well to pupils' learning.

Table E2. *Continued*

Rating	Behaviour and safety of pupils at the school	Quality of leadership in and management of the school
Good	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 nationally over a sustained period.	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 skilfully question groups of pupils and individuals during lessons in order to reshape tasks and explanations to improve learning. Teaching consistently deepens pupils' knowledge and understanding and allows them to develop a range of skills, including communication, reading and writing and mathematics, across the curriculum. Appropriate and regular homework contributes well to pupils' learning. Teachers assess pupils' progress regularly and accurately and discuss assessments with them so that pupils know how well they have done and what they need to do to improve.

Table E2. *Continued*

Rating	Behaviour and safety of pupils at the school	Quality of leadership in and management of the school
Requires improvement	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 sustained period.	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.

Table E2. *Continued*

Rating	Behaviour and safety of pupils at the school	Quality of leadership in and management of the school
Inadequate	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 cannot 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.

Rating	Overall effectiveness
Outstanding	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.

Table E2. *Continued*

Rating	Overall effectiveness
Good	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	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	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.

Notes: For 'overall effectiveness', inspectors must consider the evidence gathered in support of their evaluations of the four key judgements. *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: https://dera.ioe.ac.uk/14076/1/The_evaluation_schedule_for_school_inspections_from_January_2012%5B1%5D.pdf.

Table E3. *Sample Selection of Households.*

	Children aged 10−15 (England) (1)	Homework variable can be constructed across waves (2)	Non-missing school code (3)	School inspected in academic year of interview (4)	Final sample (5)
Sample size (children):	14,080	4,660	2,664	665	621
Household char.					
Household size	4.51	4.46	4.21	4.15	4.11
	(1.47)	(1.40)	(1.31)	(1.32)	(1.30)
Home owner	0.61	0.64	0.65	0.64	0.65
Mother char.					
Married/cohabiting	0.76	0.75	0.74	0.72	0.72
White ethnicity	0.70	0.72	0.73	0.74	0.74
Education GCSE or below	0.45	0.43	0.43	0.43	0.42
Father characteristics					
Married/cohabiting	0.97	0.97	0.96	0.95	0.95
White ethnicity	0.70	0.73	0.75	0.79	0.79
Education GCSE or below	0.42	0.38	0.40	0.43	0.43

Notes: The table shows means, with SDs in parentheses. The sample is from pooling across survey waves 1, 3 and 5. 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 E4. *Sample Selection of Schools.*

	School inspected by Ofsted		School not inspected by Ofsted	
	Wave 3 Jan 2011−Dec 2012 (1)	Wave 5 Jan 2013−Dec 2014 (2)	Wave 3 reference year: 2011/12 (3)	Wave 5 reference year: 2013/14 (4)
Number of schools	2,060	2,356	1,682	1,477
School composition				
School size: number of pupils	816.12	788.25	887.87	871.45
	(496.1)	(488.68)	(490.37)	(504.00)
% Pupils FSM	21.81	22.5	18.29	18.45
	(15.71)	(15.23)	(15.19)	(14.98)
% Pupils EAL	12.24	13.71	13.15	13.74
	(18.25)	(19.44)	(19.1)	(18.73)
School type				
Academy school	0.25	0.13	0.37	0.11
Community school	0.28	0.20	0.23	0.13
Voluntary aided or controlled school	0.11	0.09	0.13	0.09
Foundation school	0.14	0.09	0.10	0.06
Special school	0.21	0.19	0.16	0.13
School performance				
% Pupils 5 or more A*–C grades incl. English & maths	0.46	0.44	0.54	0.54
	(0.26)	(0.24)	(0.27)	(0.28)
% Pupils 5 or more A*–C grades	0.68	0.51	0.74	0.61
	(0.33)	(0.27)	(0.32)	(0.3)
% Pupils 5 or more A*–G grades	0.81	0.78	0.85	0.83
	(0.34)	(0.35)	(0.32)	(0.33)
Total average point score	291.64	256.69	313.37	286.24
	(111.24)	(109.46)	(106.89)	(113.92)
% Pupils with entries in all English Baccalaureate subjects	0.11	0.17	0.18	0.26
	(0.13)	(0.15)	(0.21)	(0.23)

Notes: The table shows means, with SDs in parentheses. 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 E5. *Predicting Ofsted Inspection Ratings.*

Past grade:	Outstanding (1)	Good (2)	Requires improvement (3)	Inadequate (4)
School size: number of pupils	−0.000	0.000	−0.000**	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)
% Pupils FSM	0.006*	0.008***	0.012***	−0.002
	(0.004)	(0.002)	(0.004)	(0.007)
Academy school	−0.711**	−1.572***	−1.894***	−1.213***
	(0.323)	(0.253)	(0.390)	(0.459)
Community school	−0.818**	−1.679***	−1.999***	−1.637***
	(0.326)	(0.251)	(0.374)	(0.449)
Voluntary aided or controlled school	−0.785**	−1.451***	−1.828***	−1.540***
	(0.347)	(0.290)	(0.386)	(0.544)
Foundation school	−0.910***	−1.382***	−1.974***	−1.558***
	(0.345)	(0.263)	(0.392)	(0.447)
Has sixth form	0.074	−0.010	−0.090	−0.074
	(0.126)	(0.061)	(0.067)	(0.160)
Christian denomination	−0.145	−0.066	−0.105	−0.121
	(0.199)	(0.129)	(0.121)	(0.289)
Other religious denomination	−0.552	−0.875	−7.532***	
	(0.514)	(0.608)	(0.288)	
Mixed gender school	−0.142	−0.106	−0.100	−1.350**
	(0.155)	(0.134)	(0.185)	(0.629)
Boys' school	−0.256	−0.154	−0.260	−1.804***
	(0.263)	(0.152)	(0.210)	(0.579)
% Pupils 5 or more A*–C grades	−0.390	0.229	1.381***	0.933
	(0.674)	(0.361)	(0.374)	(0.688)
% Pupils 5 or more A*–C grades incl. English & maths	1.108	1.638***	2.083***	2.351
	(0.999)	(0.563)	(0.652)	(1.663)
% Pupils 5 or more A*–G grades	−2.712***	−3.829***	−5.013***	−2.782**
	(0.686)	(0.515)	(0.630)	(1.141)
% Expected progress in English	1.132*	1.420***	2.717***	0.261
	(0.625)	(0.383)	(0.439)	(0.859)
% Expected progress in maths	1.796***	1.847***	2.559***	1.995*
	(0.631)	(0.347)	(0.450)	(1.212)
% Pupils achieving English Bacc.	−0.189	0.156	−0.855**	0.728
	(0.553)	(0.356)	(0.390)	(0.932)
Total average point score	0.002	0.005***	0.004***	0.004***
	(0.002)	(0.001)	(0.001)	(0.001)
Expected grade	3.380	2.915	2.394	2.041
25th percentile of modal pred. prob.	0.512	0.481	0.520	0.691
Median modal pred. prob.	0.542	0.559	0.559	0.751
75th percentile of modal pred. prob.	0.557	0.594	0.608	0.778
Number of LAs	138	151	146	106
Number of inspections	626	1,762	1,715	316
Number of schools	614	1,593	1,333	297

Notes: The dependent variable is the Ofsted grade. Results are from ordered probit regressions, with standard errors clustered by local authority. Here ***, ** and * denote significance at 1%, 5% and 10%, respectively. The sample comprises schools that are inspected during the survey period covering waves 1, 3 and 5 of UKHLS. 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). The columns show the results of separate ordered probit regressions for the school's Ofsted grade in the previous inspection cycle. Column (1) shows results for schools rated outstanding at the past inspection, columns (2)−(4) for schools rated good, requires improvement and inadequate, respectively. All columns control for the school, pupil and performance characteristics shown and cluster standard errors at the local education 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. At the foot of each column we report, for each past Ofsted grade, statistics for the predicted expected Ofsted grade. The first row shows the mean predicted Ofsted grade (the sum of the proportion predicted each Ofsted grade multiplied by Ofsted grade). The following three rows show statistics for the modal predicted value only (the most likely predicted grade).

Table E6. *Balance, by News Shock.*

	Good news			No news			Bad news		
	Treated: interviewed after inspection (1)	Control: interviewed before inspection (2)	Test of equality [*p*-value] (3)	Treated: interviewed after inspection (4)	Control: interviewed before inspection (5)	Test of equality [*p*-value] (6)	Treated: interviewed after inspection (7)	Control: interviewed before inspection (8)	Test of equality [*p*-value] (9)
Number of children	69	46		220	159		78	49	
Panel A: school characteristics									
School size: number of pupils	1,111.2	1,136.0	[0.705]	1,169.1	1,079.5	[0.068]	1,109.8	1,085.9	[0.747]
	(372.1)	(353.4)		(375.0)	(331.7)		(374.9)	(392.1)	
% Pupils FSM	17.21	17.11	[0.979]	17.93	16.16	[0.097]	16.54	18.34	[0.464]
	(10.83)	(16.54)		(13.75)	(13.35)		(10.00)	(14.13)	
Academy school	0.16	0.30	[0.082]	0.25	0.26	[0.857]	0.14	0.24	[0.147]
Boys' school	0.06	0.02	[0.370]	0.05	0.03	[0.425]	0.06	0.06	[0.957]
% Pupils 5 or more A*−C grades	0.80	0.78	[0.598]	0.78	0.79	[0.596]	0.71	0.72	[0.748]
	(0.18)	(0.14)		(0.14)	(0.14)		(0.22)	(0.15)	
Total average GCSE point score	333.32	336.78	[0.672]	335.50	337.25	[0.641]	317.01	325.86	[0.330]
	(54.25)	(27.42)		(33.77)	(34.04)		(63.56)	(37.97)	
Panel B: household characteristics									
Household size	3.91	4.50	[0.007]	4.20	4.08	[0.338]	4.00	3.92	[0.740]
	(1.12)	(1.13)		(1.37)	(1.31)		(1.33)	(1.19)	
Home owner	0.61	0.59	[0.840]	0.63	0.75	[0.021]	0.68	0.59	[0.334]
Panel C: child characteristics									
Female	0.51	0.48	[0.778]	0.56	0.48	[0.175]	0.40	0.51	[0.252]
Age	13.72	13.70	[0.899]	13.4	13.5	[0.353]	13.45	13.12	[0.096]
	(1.24)	(1.19)		(1.07)	(1.18)		(0.99)	(1.05)	
Panel D: mother characteristics									
Married/cohabiting	0.67	0.69	[0.810]	0.73	0.79	[0.236]	0.65	0.60	[0.547]
White ethnicity	0.73	0.69	[0.734]	0.72	0.79	[0.092]	0.80	0.62	[0.018]
Education GCSE or below	0.35	0.52	[0.045]	0.43	0.38	[0.294]	0.40	0.53	[0.175]
Panel E: father characteristics									
Married/cohabiting	0.91	1.00	[0.085]	0.98	0.93	[0.083]	1.00	0.87	[0.089]
White ethnicity	0.70	0.82	[0.294]	0.79	0.87	[0.095]	0.66	0.70	[0.736]
Education GCSE or below	0.64	0.36	[0.071]	0.42	0.44	[0.695]	0.41	0.30	[0.453]

Notes: The table shows means, with SDs in parentheses and *p*-values in brackets. 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 the Ofsted inspection. For schools that receive a positive news shock, columns (1) and (2) show means and SDs 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 E7. *Parental Response to Information on School Quality.*

	Forecast, unconditional (1)	Plus child characteristics (2)	Plus parent characteristics (3)	Plus school characteristics (4)
Treated (β_0)	−0.067	−0.085	−0.071	−0.072
	(0.118)	(0.119)	(0.120)	(0.120)
Treated × good news (β_1)	−0.569**	−0.538*	−0.565*	−0.564*
	(0.278)	(0.276)	(0.309)	(0.308)
Treated × bad news (β_2)	0.191	0.231	0.214	0.227
	(0.257)	(0.259)	(0.243)	(0.248)
Good news (δ_1)	0.148	0.101	0.159	0.141
	(0.251)	(0.253)	(0.266)	(0.262)
Bad news (δ_2)	−0.105	−0.112	−0.055	−0.071
	(0.219)	(0.221)	(0.210)	(0.216)
T-C | good news ($\beta_0 + \beta_1$)	−0.636**	−0.623**	−0.636**	−0.636**
	(0.251)	(0.251)	(0.289)	(0.289)
T-C | bad news ($\beta_0 + \beta_2$)	0.124	0.146	0.142	0.155
	(0.232)	(0.233)	(0.227)	(0.230)
Diff-in-diff-in-diff ($\beta_1 - \beta_2$)	−0.760**	−0.769**	−0.778**	−0.791**
	(0.356)	(0.355)	(0.369)	(0.370)
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	621	621	621	621

Notes: The table shows ordered probit regression estimates. Bootstrapped standard errors, clustered by local authority, are shown in parentheses. Here ** and * denote significance at 5% and 10%, respectively. 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 the Ofsted inspection. The outcome variable is the change in parental help with homework between survey waves 3 and 1, or between survey waves 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 if there is 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 E8. *Further Support for Identifying Assumptions and Robustness Checks.*

	Controlling for month of interview (1)	Controlling for month of interview (2)	Placebo: next year's inspections (3)	Drop inadequate schools (4)	Omitting obs 2 weeks post-inspection (5)	Omitting obs 3 weeks post-inspection (6)	Omitting obs 4 weeks post-inspection (7)	Controlling for full set of school characteristics (8)	Controlling for previous Ofsted rating (9)	Dropping 12 year olds (10)
T-C | good news ($\beta_0 + \beta_1$)	−0.698**	−0.605**	0.165	−0.636**	−0.682**	−0.705***	−0.720***	−0.606**	−0.638**	−0.601*
	(0.290)	(0.255)	(0.224)	(0.291)	(0.285)	(0.262)	(0.264)	(0.288)	(0.288)	(0.333)
T-C | bad news ($\beta_0 + \beta_2$)	0.042	0.116	0.154	0.146	0.170	0.171	0.130	0.165	0.161	0.111
	(0.277)	(0.239)	(0.217)	(0.240)	(0.233)	(0.246)	(0.244)	(0.234)	(0.230)	(0.318)
Diff-in-diff-in-diff ($\beta_1 - \beta_2$)	−0.740*	−0.721**	0.011	−0.783**	−0.852**	−0.876**	−0.850**	−0.771**	−0.798**	−0.712
	(0.395)	(0.353)	(0.315)	(0.368)	(0.367)	(0.352)	(0.352)	(0.370)	(0.369)	(0.461)
F-test: month dummies [*p*-value]	6.434 [0.778]	3.542 [0.966]								
Forecast Ofsted rating	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Child char.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parent char.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School char.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	621	2,103	513	575	598	590	584	620	621	453

Notes: The table shows ordered probit regression estimates. Bootstrapped standard errors, clustered by local authority, are shown in parentheses. Here ***, ** and * denote significance at 1%, 5% and 10%, respectively. 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 the Ofsted inspection. The outcome variable is the change in parental help with homework between survey waves 3 and 1, or between survey waves 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 if there is 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. Column (2) controls for month of household interview and includes schools without an inspection in the sample. At the foot of columns (1) and (2) we report the *F*-statistic (and *p*-value) on the null that these month of interview dummies are jointly insignificant. The sample in column (3) 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 (4) drops failing schools (those ranked as 1 = inadequate by Ofsted in year *t*). Columns (5) to (7) omit households interviewed within 2, 3 and 4 weeks of an Ofsted inspection. Column (8) additionally controls for a full set of school characteristics. Column (9) additionally controls for the previous Ofsted rating. Column (10) drops 12 year olds (who 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 education authority level and shown in parentheses.

Table E9. *Short-Run School Responses to Information on School Quality.*

	Hours children expected to do homework (1)	No. class support teachers (2)	Days used supply teachers (3)	Time teaching numeracy and literacy (4)	% Time in numeracy and literacy groups (5)	% Time numeracy and literacy with individuals (6)	Time use index based on (1), (4)−(6) (7)
T-C | good news ($\beta_0 + \beta_1$)	−6.523	−0.205	−2.168	−0.203	2.661	0.008	0.073
	(8.849)	(0.222)	(1.623)	(0.579)	(1.849)	(3.871)	(0.267)
T-C | bad news ($\beta_0 + \beta_2$)	−8.653	0.248	−1.514	0.292	−0.661	−0.196	0.041
	(7.735)	(0.223)	(1.479)	(0.482)	(2.130)	(3.932)	(0.209)
Diff-in-diff-in-diff ($\beta_1 - \beta_2$)	2.13	−0.453	−0.655	−0.495	3.321	0.203	0.032
	(9.630)	(0.274)	(2.146)	(0.715)	(2.733)	(4.989)	(0.289)
Mean dep. var.	47.7	1.40	3.91	10.78	32.13	35.26	0.00
Forecast Ofsted rating	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School char.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,198	1,132	1,138	893	1,081	1,076	774

	Use streaming (8)	Use sets: literacy (9)	Use sets: numeracy (10)	Use ability grouping (11)	Use subject groups: literacy (12)	Use subject groups: numeracy (13)	Teaching practice index based on (8)−(13) (14)
T-C | good news ($\beta_0 + \beta_1$)	−0.122	−0.197	−0.173	0.129	0.089	0.142	−0.004
	(0.112)	(0.130)	(0.144)	(0.108)	(0.087)	(0.089)	(0.062)
T-C | bad news ($\beta_0 + \beta_2$)	−0.092	−0.119	−0.096	0.002	−0.012	0.037	−0.039
	(0.108)	(0.130)	(0.138)	(0.093)	(0.090)	(0.091)	(0.053)
Diff-in-diff-in-diff ($\beta_1 - \beta_2$)	−0.03	−0.078	−0.077	0.128	0.102	0.104	0.035
	(0.142)	(0.174)	(0.187)	(0.134)	(0.120)	(0.114)	(0.073)
Mean dep. var.	0.20	0.32	0.37	0.83	0.91	0.89	0.59
Forecast Ofsted rating	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School char.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,258	1,229	1,218	1,277	1,277	1,275	1,172

Notes: The table shows linear regression estimates. Bootstrapped standard errors, clustered by local authority, are shown in parentheses. 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 the 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. Column (7) is a time use index derived from standardised measures of columns (1) and (4)–(6). Streaming in column (8) is the practice of dividing a class into hierarchical ability groups according to overall ability. Setting (columns (9) and (10)) is ability grouping by subject. Column (14) is an index of teaching practices derived from columns (8)–(13). 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 E10. *Linear Probability Model Estimates of the Parental Response to Information on School Quality.*

	Panel A: frequency of help increases (0/1)			Panel B: frequency of help decreases (0/1)		
	Child char. (1)	Plus parent char. (2)	Plus school char. (3)	Child char. (4)	Plus parent char. (5)	Plus school char. (6)
T-C | good news	−0.227***	−0.227**	−0.227**	0.154	0.157	0.154
$(\beta_0 + \beta_1)$	(0.082)	(0.089)	(0.089)	(0.096)	(0.105)	(0.105)
T-C | bad news	−0.073	−0.073	−0.069	−0.157*	−0.157*	−0.161*
$(\beta_0 + \beta_2)$	(0.079)	(0.078)	(0.079)	(0.090)	(0.088)	(0.087)
Diff-in-diff-in-diff	−0.155	−0.154	−0.158	0.311**	0.313**	0.315**
$(\beta_1 - \beta_2)$	(0.118)	(0.119)	(0.120)	(0.139)	(0.143)	(0.141)
Forecast Ofsted rating	Yes	Yes	Yes	Yes	Yes	Yes
Child char.	Yes	Yes	Yes	Yes	Yes	Yes
Parent char.	No	Yes	Yes	No	Yes	Yes
School char.	No	No	Yes	No	No	Yes
Adjusted R^2	0.022	0.024	0.024	0.024	0.034	0.042
Observations	621	621	621	621	621	621

Notes: The table shows linear regression estimates. Bootstrapped standard errors, clustered by local authority, are shown in parentheses. Here ***, ** and * denote significance at 1%, 5% and 10%, respectively. 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 the Ofsted inspection. The outcome variable is the change in parental help with homework between survey waves 3 and 1, or between survey waves 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 if there is 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 wave 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 E11. *Alternative Forecasting Models.*

	AR(1) (1)	School characteristics (2)	Full model (3)	Naïve model (4)
T-C | good news ($\beta_0 + \beta_1$)	−0.401*	−0.600**	−0.636**	−0.238
	(0.227)	(0.236)	(0.289)	(0.239)
T-C | bad news ($\beta_0 + \beta_2$)	0.119	0.030	0.155	0.375
	(0.255)	(0.225)	(0.230)	
	(0.242)			
Diff-in-diff-in-diff ($\beta_1 - \beta_2$)	−0.520	−0.630*	−0.791**	−0.613
	(0.355)	(0.343)	(0.370)	(0.294)
Forecast Ofsted rating	Yes	Yes	Yes	No
Child characteristics	Yes	Yes	Yes	Yes
Parent characteristics	Yes	Yes	Yes	Yes
School characteristics	Yes	Yes	Yes	Yes
Observations	621	621	621	621

Notes: The table shows ordered probit regression estimates. Bootstrapped standard errors, clustered by local authority, are shown in parentheses. Here ** and * denote significance at 5% and 10%, respectively. 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 the Ofsted inspection. The outcome variable is the change in parental help with homework between survey waves 3 and 1, or between survey waves 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 if there is 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 separate ordered probit regression for each past Ofsted inspection grade to forecast the Ofsted rating and controls for local authority fixed effects. Column (2) additionally controls for school characteristics; column (3) additionally controls for school characteristics plus all performance measures. Column (4) uses the past Ofsted grade as the forecast grade to derive the news shock variable. 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 education authority level and shown in parentheses.

Table E12. *Test Score Impacts of the School Inspection Regime: Robustness Check, Excluding Schools without Lagged Test Score Information.*

	Student level Standardised GCSE average point score (1)
T-C | good news ($\beta_0 + \beta_1$)	−0.094*
	(0.046)
	[−0.169, −0.018]
T-C | bad news ($\beta_0 + \beta_2$)	−0.037
	(0.032)
	[−0.089, 0.014]
Forecast Ofsted rating	Yes
School characteristics	Yes
Pupil characteristics	Yes
Number of schools	1,133
Number of pupils	202,408

Notes: See the notes to Table 7. This table replicates the analysis in column (1) of Table 7 for the sample of students in schools where lagged test score information is available (the sample of schools is 1,133, consistent with that in columns (2) and (3) of Table 7).

University of Bristol & The Institute for Fiscal Studies, UK
University of Sussex, UK
University of Essex, UK
University College London & The Institute for Fiscal Studies, UK

Additional Supporting Information may be found in the online version of this article:

Replication Package

References

Abdulkadiroğlu, A., Angrist, J. and Pathak, P. (2014). 'The elite illusion: Achievement effects at Boston and New York exam schools', *Econometrica*, vol. 82(1), pp. 137–96.
Abdulkadiroğlu, A., Pathak, P.A., Schellenberg, J. and Walters, C.R. (2020). 'Do parents value school effectiveness?', *American Economic Review*, vol. 110(5), pp. 1502–39.
Agarwal, N. and Somaini, P. (2018). 'Demand analysis using strategic reports: An application to a school choice mechanism', *Econometrica*, vol. 86(2), pp. 391–444.
Ainsworth, R., Dehejia, R., Pop-Eleches, C. and Urquiola, M. (2020). 'Information, preferences, and household demand for school value added', Working paper, National Bureau of Economic Research.
Ajayi, K.F., Friedman, W.H. and Lucas, A.M. (2017). 'The importance of information targeting for school choice', *American Economic Review*, vol. 107(5), pp. 638–43.
Albornoz, F., Berlinski, S. and Cabrales, A. (2017). 'Motivation, resources, and the organization of the school system', *Journal of the European Economic Association*, vol. 16(1), pp. 199–231.
Andrabi, T., Das, J. and Khwaja, A.I. (2017). 'Report cards: The impact of providing school and child test scores on educational markets', *American Economic Review*, vol. 107(6), pp. 1535–63.
Attanasio, O., Boneva, T. and Rauh, C. (2020a). 'Parental beliefs about returns to different types of investments in school children', *Journal of Human Resources*, vol. 57(6), pp. 1789–825.
Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C. and Rubio-Codina, M. (2020b). 'Estimating the production function for human capital: Results from a randomized controlled trial in Colombia', *American Economic Review*, vol. 110(1), pp. 48–85.
Autor, D., Figlio, D., Karbownik, K., Roth, J. and Wasserman, M. (2016). 'School quality and the gender gap in educational achievement', *American Economic Review*, vol. 106(5), pp. 289–95.
Avvisati, F., Gurgand, M., Guyon, N. and Maurin, E. (2013). 'Getting parents involved: A field experiment in deprived schools', *The Review of Economic Studies*, vol. 81(1), pp. 57–83.
Becker, G.S. and Tomes, N. (1976). 'Child endowments and the quantity and quality of children', *Journal of Political Economy*, vol. 84(4, Part 2), pp. S143–62.
Behrman, J.R., Pollak, R.A. and Taubman, P. (1982). 'Parental preferences and provision for progeny', *Journal of Political Economy*, vol. 90(1), pp. 52–73.
Bergman, P. (2021). 'Parent-child information frictions and human capital investment: Evidence from a field experiment', *Journal of Political Economy*, vol. 129(1), pp. 286–322.
Beuermann, D., Jackson, C.K., Navarro-Sola, L. and Pardo, F. (2018). 'What is a good school, and can parents tell? Evidence on the multidimensionality of school output', Working paper, National Bureau of Economic Research.
Boneva, T. and Rauh, C. (2018). 'Parental beliefs about returns to educational investments–The later the better?', *Journal of the European Economic Association*, vol. 16(6), pp. 1669–711.
Bono, E.D., Francesconi, M., Kelly, Y. and Sacker, A. (2016). 'Early maternal time investment and early child outcomes', Economic Journal, vol. 126(596), pp. F96–135.
Burgess, S., Greaves, E., Vignoles, A. and Wilson, D. (2015). 'What parents want: School preferences and school choice', Economic Journal, vol. 125(587), pp. 1262–89.
Burgess, S., Wilson, D. and Worth, J. (2013). 'A natural experiment in school accountability: The impact of school performance information on pupil progress', *Journal of Public Economics*, vol. 106, pp. 57–67.
Bursztyn, L. and Coffman, L.C. (2012). 'The schooling decision: Family preferences, intergenerational conflict, and moral hazard in the Brazilian favelas', *Journal of Political Economy*, vol. 120(3), pp. 359–97.
Burton, P., Phipps, S. and Curtis, L. (2002). 'All in the family: A simultaneous model of parenting style and child conduct', *American Economic Review*, vol. 92(2), pp. 368–72.
Caetano, G., Kinsler, J. and Teng, H. (2019). 'Towards causal estimates of children's time allocation on skill development', *Journal of Applied Econometrics*, vol. 34(4), pp. 588–605.
Carneiro, P., Løken, K.V. and Salvanes, K.G. (2015). 'A flying start? Maternity leave benefits and long-run outcomes of children', *Journal of Political Economy*, vol. 123(2), pp. 365–412.
Clark, D. and Del Bono, E. (2016). 'The long-run effects of attending an elite school: Evidence from the United Kingdom', *American Economic Journal: Applied Economics*, vol. 8(1), pp. 150–76.

Cullen, J.B., Jacob, B.A. and Levitt, S. (2006). 'The effect of school choice on participants: Evidence from randomized lotteries', *Econometrica*, vol. 74(5), pp. 1191–230.
Cunha, F. and Heckman, J. (2007). 'The technology of skill formation', *American Economic Review*, vol. 97(2), pp. 31–47.
Cunha, F. and Heckman, J.J. (2009). 'The economics and psychology of inequality and human development', *Journal of the European Economic Association*, vol. 7(2–3), pp. 320–64.
Cunha, F., Heckman, J.J. and Schennach, S.M. (2010). 'Estimating the technology of cognitive and noncognitive skill formation', *Econometrica*, vol. 78(3), pp. 883–931.
Currie, J. and Almond, D. (2011). 'Human capital development before age five', in (D. Card and O. Ashenfelter, eds.), *Handbook of Labor Economics*, vol. 4, pp. 1315–486, Amsterdam: Elsevier.
Dale, S.B. and Krueger, A.B. (2002). 'Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables', *The Quarterly Journal of Economics*, vol. 117(4), pp. 1491–527.
Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidharan, K. and Sundararaman, V. (2013). 'School inputs, household substitution, and test scores', *American Economic Journal: Applied Economics*, vol. 5(2), pp. 29–57.
Datar, A. and Mason, B. (2008). 'Do reductions in class size "crowd out" parental investment in education?', *Economics of Education Review*, vol. 27(6), pp. 712–23.
De Fraja, G., Oliveira, T. and Zanchi, L. (2010). 'Must try harder: Evaluating the role of effort in educational attainment', *The Review of Economics and Statistics*, vol. 92(3), pp. 577–97.
Del Boca, D., Flinn, C. and Wiswall, M. (2013). 'Household choices and child development', *The Review of Economic Studies*, vol. 81(1), pp. 137–85.
Del Boca, D., Monfardini, C. and Nicoletti, C. (2017). 'Parental and child time investments and the cognitive development of adolescents', *Journal of Labor Economics*, vol. 35(2), pp. 565–608.
Deming, D.J., Cohodes, S., Jennings, J. and Jencks, C. (2016). 'School accountability, postsecondary attainment, and earnings', *The Review of Economics and Statistics*, vol. 98(5), pp. 848–62.
Dizon-Ross, R. (2019). 'Parents' beliefs about their children's academic ability: Implications for educational investments', *American Economic Review*, vol. 109(8), pp. 2728–65.
Dizon-Ross, R. (2020). 'How does school accountability affect teachers?: Evidence from New York City', *Journal of Human Resources*, vol. 55(1), pp. 76–118.
Dobbie, W. and Fryer, R.G. (2011). 'Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem children's zone', *American Economic Journal: Applied Economics*, vol. 3(3), pp. 158–87.
Doepke, M., Sorrenti, G. and Zilibotti, F. (2019). 'The economics of parenting', *Annual Review of Economics*, vol. 11(1), pp. 55–84.
Dranove, D. and Jin, G.Z. (2010). 'Quality disclosure and certification: Theory and practice', *Journal of Economic Literature*, vol. 48(4), pp. 935–63.
Duflo, E., Dupas, P. and Kremer, M. (2011). 'Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya', *American Economic Review*, vol. 101(5), pp. 1739–74.
Feng, L., Figlio, D. and Sass, T. (2018). 'School accountability and teacher mobility', *Journal of Urban Economics*, vol. 103, pp. 1–17.
Figlio, D. and Loeb, S. (2011). 'School accountability', in (E.A. Hanushek, S. Machin and L. Woessmann, eds.), *Handbook of the Economics of Education*, vol. 3, pp. 383–421, Amsterdam: Elsevier.
Figlio, D.N. and Kenny, L.W. (2009). 'Public sector performance measurement and stakeholder support', *Journal of Public Economics*, vol. 93(9), pp. 1069–77.
Figlio, D.N. and Lucas, M.E. (2004). 'What's in a grade? School report cards and the housing market', *American Economic Review*, vol. 94(3), pp. 591–604.
Fiorini, M. and Keane, M.P. (2014). 'How the allocation of children's time affects cognitive and noncognitive development', *Journal of Labor Economics*, vol. 32(4), pp. 787–836.
Fredriksson, P., Öckert, B. and Oosterbeek, H. (2016). 'Parental responses to public investments in children: Evidence from a maximum class size rule', *Journal of Human Resources*, vol. 51(4), pp. 832–68.
Fu, C. and Mehta, N. (2018). 'Ability tracking, school and parental effort, and student achievement: A structural model and estimation', *Journal of Labor Economics*, vol. 36(4), pp. 923–79.
Greenwood, J. (2019). *Evolving Households: The Imprint of Technology on Life*, Cambridge, MA: MIT Press.
Hanushek, E.A. and Raymond, M.E. (2005). 'Does school accountability lead to improved student performance?', *Journal of Policy Analysis and Management*, vol. 24(2), pp. 297–327.
Hao, L., Hotz, V.J. and Jin, G.Z. (2008). 'Games parents and adolescents play: Risky behaviour, parental reputation and strategic transfers', Economic Journal, vol. 118(528), pp. 515–55.
Hastings, J., Kane, T.J. and Staiger, D.O. (2009). 'Heterogeneous preferences and the efficacy of public school choice', Working Paper 2145, National Bureau of Economic Research.
Hastings, J.S. and Weinstein, J.M. (2008). 'Information, school choice, and academic achievement: Evidence from two experiments', *The Quarterly Journal of Economics*, vol. 123(4), pp. 1373–414.
Heckman, J.J. and Mosso, S. (2014). 'The economics of human development and social mobility', *Annual Review of Economics*, vol. 6(1), pp. 689–733.
Hoekstra, M. (2009). 'The effect of attending the flagship state university on earnings: A discontinuity-based approach', *The Review of Economics and Statistics*, vol. 91(4), pp. 717–24.

Houtenville, A.J. and Conway, K.S. (2008). 'Parental effort, school resources, and student achievement', *Journal of Human Resources*, vol. 43(2), pp. 437–53.
Hoxby, C.M. and Avery, C. (2012). 'The missing "one-offs": The hidden supply of high-achieving, low income students', Working paper, National Bureau of Economic Research.
Hussain, I. (2015). 'Subjective performance evaluation in the public sector: Evidence from school inspections', *Journal of Human Resources*, vol. 50(1), pp. 189–221.
Hussain, I. (2020). 'Housing market and school choice response to school quality information shocks', Unpublished Manuscript.
Jacob, B.A. (2005). 'Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago public schools', *Journal of public Economics*, vol. 89(5–6), pp. 761–96.
Jacob, B.A. and Lefgren, L. (2007). 'What do parents value in education? An empirical investigation of parents' revealed preferences for teachers', *The Quarterly Journal of Economics*, vol. 122(4), pp. 1603–37.
Jacob, B.A. and Levitt, S.D. (2003). 'Rotten apples: An investigation of the prevalence and predictors of teacher cheating', *The Quarterly Journal of Economics*, vol. 118(3), pp. 843–77.
Jensen, R. (2010). 'The (perceived) returns to education and the demand for schooling', *The Quarterly Journal of Economics*, vol. 125(2), pp. 515–48.
Johnson, P. (2004). 'Education policy in England', *Oxford Review of Economic Policy*, vol. 20(2), pp. 173–97.
Kane, T.J. and Staiger, D.O. (2002). 'The promise and pitfalls of using imprecise school accountability measures', *Journal of Economic Perspectives*, vol. 16(4), pp. 91–114.
Kinsler, J. and Pavan, R. (2021). 'Local distortions in parental beliefs over child skill', *Journal of Political Economy*, vol. 129(1), pp. 81–100.
Kirabo Jackson, C. (2010). 'Do students benefit from attending better schools? Evidence from rule–based student assignments in Trinidad and Tobago', Economic Journal, vol. 120(549), pp. 1399–429.
Lucas, A.M. and Mbiti, I.M. (2012). 'The determinants and consequences of school choice errors in Kenya', *American Economic Review*, vol. 102(3), pp. 283–8.
MacLeod, W.B. and Urquiola, M. (2015). 'Reputation and school competition', *American Economic Review*, vol. 105(11), pp. 3471–88.
MacLeod, W.B. and Urquiola, M. (2019). 'Is education consumption or investment? Implications for school competition', *Annual Review of Economics*, vol. 11(1), pp. 563–89.
Maldonado, J.E., De Witte, K. and Declercq, K. (2021). 'The effects of parental involvement in homework: Two randomised controlled trials in financial education', *Empirical Economics*, vol. 62(3), pp. 1–26.
MCS. (2017). 'Millennium cohort study: Ofsted and teacher survey linked data, 2008: Secure access', UK Data Service, SN: 8227, http://dx.doi.org/10.5255/UKDA-SN-8227-1.
Murphy, K.M. and Topel, R.H. (2002). 'Estimation and inference in two-step econometric models', *Journal of Business and Economic Statistics*, vol. 20(1), pp. 88–97.
Neal, D. and Schanzenbach, D.W. (2010). 'Left behind by design: Proficiency counts and test-based accountability', *The Review of Economics and Statistics*, vol. 92(2), pp. 263–83.
Neilson, C., Allende, C. and Gallego, F. (2019). 'Approximating the equilibrium effects of informed school choice', Working paper, Princeton University.
Pathak, P.A. and Sönmez, T. (2008). 'Leveling the playing field: Sincere and sophisticated players in the Boston mechanism', *American Economic Review*, vol. 98(4), pp. 1636–52.
Pop-Eleches, C. and Urquiola, M. (2013). 'Going to a better school: Effects and behavioral responses', *American Economic Review*, vol. 103(4), pp. 1289–324.
Rouse, C.E., Hannaway, J., Goldhaber, D. and Figlio, D. (2013). 'Feeling the Florida heat? How low-performing schools respond to voucher and accountability pressure', *American Economic Journal: Economic Policy*, vol. 5(2), pp. 251–81.
Todd, P.E. and Wolpin, K.I. (2003). 'On the specification and estimation of the production function for cognitive achievement', Economic Journal, vol. 113(485), pp. F3–33.
UKHLS. (2018). 'Understanding Society: Waves 1-10, 2009-2019 and Harmonised BHPS: Waves 1-18, 1991-2009', 11 edn., UK Data Service, SN: 6931.
Weinberg, B.A. (2001). 'An incentive model of the effect of parental income on children', *Journal of Political Economy*, vol. 109(2), pp. 266–80.
Yi, J., Heckman, J.J., Zhang, J. and Conti, G. (2015). 'Early health shocks, intra-household resource allocation and child outcomes', Economic Journal, vol. 125(588), pp. F347–71.