



Education  
Endowment  
Foundation

# Magic Breakfast

## Evaluation report and executive summary

Originally published: November 2016

Updated: December 2019

### **Independent evaluators:**

Claire Crawford, Amy Edwards, Christine Farquharson, Ellen Greaves, Grace Trevelyan, Emma Wallace, and Clarissa White



Institute for  
Fiscal Studies



The Education Endowment Foundation (EEF) is an independent grant-making charity dedicated to breaking the link between family income and educational achievement, ensuring that children from all backgrounds can fulfil their potential and make the most of their talents.

The EEF aims to raise the attainment of children facing disadvantage by:

- identifying promising educational innovations that address the needs of disadvantaged children in primary and secondary schools in England;
- evaluating these innovations to extend and secure the evidence on what works and can be made to work at scale; and
- encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus Trust (now part of Impetus - Private Equity Foundation) and received a founding £125m grant from the Department for Education.

Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.



**For more information about the EEF or this report please contact:**

**Jonathan Kay**  
Education Endowment Foundation  
9th Floor, Millbank Tower  
21–24 Millbank  
SW1P 4QP

**p:** 020 7802 1679  
**e:** [jonathan.kay@eefoundation.org.uk](mailto:jonathan.kay@eefoundation.org.uk)  
**w:** [www.educationendowmentfoundation.org.uk](http://www.educationendowmentfoundation.org.uk)

## About the evaluator

The project was independently evaluated by a team from the Institute for Fiscal Studies and the National Children's Bureau: Claire Crawford, Amy Edwards, Christine Farquharson, Ellen Greaves, Grace Trevelyan, Emma Wallace, and Clarissa White.

Ellen Greaves led the impact evaluation, working with Christine Farquharson and Claire Crawford, and with support from Laura Westwood. The process evaluation was carried out by a team from the National Children's Bureau (NCB) Research Centre: Emma Wallace and Clarissa White oversaw the design and delivery of the process evaluation and edited the report. Grace Trevelyan and Amy Edwards carried out the case studies, analysis and reporting; Jo Lea was involved in the early stages of the research and helped to design the topic guides used to explore breakfasts with young people.

The lead evaluator was Ellen Greaves.

### **Contact details:**

Ellen Greaves  
IFS, 7 Ridgmount Street, London, WC1E 7AE.

Tel: 020 7291 4800

Email: [ellen\\_g@ifs.org.uk](mailto:ellen_g@ifs.org.uk)

## Contents

|   |     |
|---|-----|
| About the evaluator .....   | 2   |
| Revised report .....  | 5   |
| Updated executive summary .....                                     | 6   |
| Introduction.....   | 9   |
| Methods.....  | 10  |
| Impact evaluation .....   | 20  |
| Conclusion .....  | 24  |
| Revised Appendix B: Security classification of trial findings ..... | 28  |
| Revised Appendix H: Analysis code.....                              | 32  |
| Appendix X: Documenting changes to original report.....             | 35  |
| Appendix Y: Local Education Authority (LEA) codes .....             | 37  |
| Appendix Z: Additional robustness checks .....                      | 38  |
| ORIGINAL REPORT AS PUBLISHED NOVEMBER 2016 .....                    | 55  |
| Executive summary .....   | 55  |
| Introduction.....   | 57  |
| Methods.....  | 63  |
| Impact evaluation .....   | 73  |
| Process evaluation.....   | 103 |
| Conclusion .....  | 110 |
| References .....  | 114 |
| Appendix A: EEF cost rating .....                                   | 116 |
| Appendix B: Security classification of trial findings.....          | 117 |
| Appendix C: Consent form and Memorandum of Understanding.....       | 118 |
| Appendix D: Pupil, teacher and headteacher surveys .....            | 122 |

|   |            |
|---|------------|
| <b>Appendix E: Stata code for randomisation .....</b>   | <b>140</b> |
| <b>Appendix F: Distribution of attainment for those with and without demographic information in NPD .....</b> | <b>141</b> |
| <b>Appendix G: Achieved sample for case studies .....</b>   | <b>142</b> |
| <b>Appendix H: Analysis code .....</b>  | <b>143</b> |
| <b>Appendix I: Pupil Hunger Sensitivity Analysis .....</b>  | <b>146</b> |
| <b>Appendix J: Cost distribution .....</b>  | <b>147</b> |

## Revised report

The Magic Breakfast project provides schools with support and resources to offer a free, universal, before-school breakfast club. The purpose of this project was to evaluate the impact that breakfast clubs—and, specifically, Magic Breakfast—have on academic attainment and other outcomes at Key Stage 1 and Key Stage 2. In our original report, 'Magic Breakfast: Evaluation report and executive summary', we reported evidence from a randomised controlled trial (RCT) that this programme improved attainment at Key Stage 1 by about two months' expected progress (with smaller, and not statistically significant, results at Key Stage 2).

Unfortunately, there was a methodological error in how schools were allocated to intervention and control groups meaning that the breakfast clubs were not randomly allocated to schools. This means that the trial cannot be considered an RCT. Instead, the project has a comparison group design. This mistake was made, discovered, and reported to the EEF by the principal investigator for the impact evaluation. More information on this error can be found in a statement by IFS published alongside this revised report.

We have therefore re-analysed the data to take account of this non-random assignment and apply a more appropriate methodology. This revised report presents full details on revised results on the effectiveness of breakfast club provision. We find that the original conclusions remain broadly the same at Key Stage 1 (age 7): offering schools support to establish a free, universal breakfast club boosts attainment at Key Stage 1 by around two months' expected progress, although this result now has a lower security rating. There is no impact at Key Stage 2 (Year 6 pupils). The original estimate for Key Stage 2 was also not statistically significantly different from zero, though the new estimate is smaller and less secure.

The error in the randomisation code meant that schools in London and other urban areas were disproportionately assigned to the intervention group, which has the potential to bias the original results. For example, Greaves, Macmillan, and Sibieta (2014) show that London schools are more effective than schools in other regions at improving attainment at primary level. To the extent that this overall 'London effect' is relevant in our sample of schools, analysis that does not take school location into account will attribute these general differences to the impact of the intervention. Since our intervention schools are disproportionately in London, this would tend to overstate the impact of Magic Breakfast. In our revised results, we therefore explicitly account for (a) whether the school is inside or outside London and (b) for measures of previous school-level attainment in addition to other pupil and school-level factors that were accounted for in the original analysis. We have conducted multiple checks to assess the sensitivity of these results, which have been peer-reviewed by two independent reviewers. The revised report presents only the results from the preferred specification for clarity, but all robustness checks are available in Appendix Z.

In this revised report, we first present sections that have been revised, which supersede the relevant sections in the original report. These are the executive summary, introduction, methods (trial design and analysis sub-sections), impact evaluation (academic attainment sub-section for primary outcomes), and conclusion. Second, we briefly describe additional findings from the original report, commenting on whether the results are likely to hold using the preferred estimation strategy. Third, we present revised appendices: these are Appendix B (security classification of trial findings) and Appendix H (analysis code). Additional appendices (X, Y and Z) document all changes made to the revised sections of the report and provide additional information about the non-random assignment. Finally, the original report is replicated for ease of comparison.

The principal investigator takes full responsibility for the error and is grateful for the help of colleagues at IFS and the EEF and the independent peer reviewers for assistance in establishing the reliability of the revised results.

# Updated executive summary

## The project

The Magic Breakfast project provided 106 schools with support and resources to offer a free, universal, before-school breakfast club, including to all Year 2 and Year 6 pupils. The aim of the project was to improve attainment outcomes by increasing the number of children who ate a healthy breakfast. The schools in the project were schools in England with a relatively high proportion of disadvantaged pupils. The project ran between September 2014 and July 2015. Schools were provided with free food, support from a Magic Breakfast school change leader, and a £300 grant towards up-front costs. The intervention itself was delivered by school staff and volunteers.

The impact of the project was evaluated using a comparison group design involving around 8,600 pupils. The process evaluation involved qualitative research with four case study schools. The project was jointly funded by the Department for Education and the Education Endowment Foundation and delivered by the charity Magic Breakfast.

### Key conclusions

1. Year 2 children in breakfast club schools made the equivalent of two months' additional progress compared to Year 2 children in the business as usual control group. This result has a low to moderate security rating. These results are similar to the original results, although they are now less secure.
2. There is no evidence that breakfast clubs had an impact on Year 6 pupil outcomes. This result has moderate to high security. Compared to the original results, the effect size for Year 6 pupils is lower and less secure.
3. The findings suggest that, where improvements are seen, it is not just eating breakfast that delivers improvements but attending a breakfast club. This could be due to the content of the breakfast itself or to other social or educational benefits of the club.
4. Pupil behaviour, as measured by a teacher survey, improved in breakfast club schools. This is interesting because it shows that breakfast clubs may improve outcomes for children who do not even attend breakfast club by improving classroom environments. This key conclusion is unchanged from the original report.
5. Activities thought to increase take-up of the breakfast provision included promoting it to parents and encouraging all children to attend while sensitively targeting pupils most likely to benefit. The project required additional staff time which some schools found difficult to provide without charging for breakfast. This key conclusion is unchanged from the original report. All findings from the process evaluation are unaffected by the error in the impact evaluation.

## How secure are the findings?

The Key Stage 1 findings have low to moderate security, whilst the Key Stage 2 findings have moderate to high security. The project was an effectiveness trial, which tested whether the intervention worked under everyday condition in a large number of schools. It was evaluated using a comparison group design that compared the progress of pupils in the breakfast club schools to that of a control group of pupils receiving 'business as usual'.

There are several reasons why the comparison between intervention and control group schools is valid:

- Students in the intervention and comparison groups had *similar observable characteristics* (aside from their location). Of the 31 dimensions that we tested, the intervention and comparison groups were well-matched ('balanced') on 29. These characteristics include important drivers of attainment at Key Stage 1 and Key Stage 2, such as eligibility for free school meals, special educational needs, and the

school's Ofsted rating and level of disadvantage. They also include baseline characteristics like eating habits, absences the year before the trial, and prior attainment. The measure of prior attainment for Key Stage 1 pupils was the Foundation Stage Profile (FSP), whilst for Year 6 pupils it was Key Stage 1 outcomes. As the FSP is a less sensitive measure of prior attainment we can be less secure in the Key Stage 1 outcomes and these have therefore been awarded two padlocks.

- In addition—and unusually for a matched-comparison analysis—*the schools in this project are likely to be well-matched in important unobservable ways*. Typically in matched studies, factors like enthusiasm are unobservable and therefore impossible to match on. In this case, however, all schools were willing to participate in the evaluation, and baseline surveys of headteachers reveal that intervention and comparison group schools had similar reasons for joining the project. This means that the schools in our analysis are likely to be better matched on this dimension than is the case in many other comparison group studies.
- This was a *large-scale project* involving around 8,600 pupils in 106 schools. By including a wide range of students and schools in our analysis, we can get a better picture of the true effect of the programme on students' outcomes. In addition, there was very low attrition from outcome testing.

However, there are also important limitations to this analysis. The most important is that findings from a comparison group design cannot be as secure as those based on a randomised controlled trial. Despite the well-balanced intervention and comparison groups, and additional school-level characteristics accounted for in the revised report, it is not possible to conclusively rule out that intervention schools were systematically different to control schools before the project began. Another limitation is that around 40% of comparison group schools established some form of breakfast club provision. If the outcomes of the control group improved in response to this provision, the estimate of impact for breakfast clubs will be lower than if the control group had not established breakfast club provision.

## What are the findings?

The provision of a breakfast club led to an improvement in Key Stage 1 outcomes of around two months' progress, roughly equal to the effect of providing universal free school meals in two pilot areas in 2011 (Brown *et al.*, 2012) that led to the roll-out of that programme in infant schools. For Key Stage 2 (KS2) assessments in reading and maths the impact was close to zero. The results at KS1 are very similar in size to those in the original report (see Table 1).

The results at KS2 are smaller and close to zero. Like in the original report, these results are not statistically different from zero and so suggest no impact for Year 6 pupils.

Findings from the original report showed that teacher perceptions of classroom behaviour and concentration indicate an improvement in the breakfast club schools relative to the comparison schools over the course of the school year. These improvements mean that breakfast club provision can potentially have benefits even for children who do not attend by improving their classroom learning environment. Attendance at school also improved for children in breakfast club schools resulting in about 26 fewer half-days of absence per year for a class of 30. For these secondary outcomes, the original analysis controlled for teacher- and pupil-level baseline measures (of behaviour and absences, respectively). This means that these results look at changes within the school year and so will be less sensitive to the randomisation error and the disproportionate number of urban schools in the intervention group.

There was no evidence of an impact on the body mass index of Year 6 students. (It was not possible to estimate the effect on Year 2 students' body mass index as this year group is not part of the National Child Measurement Programme.)

Key factors for successful implementation of the breakfast clubs were (a) communication with parents to encourage take-up, (b) an established school breakfast routine, and (c) a well-functioning delivery team supported by the wider school. The main challenges were compensating staff for additional hours of work and balancing the supply of, and demand for, food. In some schools, barriers to take-up included earlier start times for pupils, breakfast charges, and a lack of ongoing promotion from the school. These results from the process evaluation are not affected by the change in methodology for the impact evaluation.



## How much does it cost?

The cost per pupil per year over three years is £11.86 (in 2014/2015 prices), averaged across all pupils in the breakfast club schools. The total cost was, on average, £4,462.11 per school. In addition, schools used 820 person-hours per year to deliver the intervention. On average, this included 87 teacher hours, 449 teaching assistant hours, 164 support staff hours, and 100 volunteer hours over the year.

**Table 1: Summary of impact on primary outcomes**

| Group & outcome         | No. of schools | Effect size<br>(95% confidence interval) | Estimated months' progress | Security rating | EEF cost rating |
|-------------------------|----------------|--|----------------------------|-----------------|-----------------|
| <b>Revised results</b>  |                |  |                            |                 |                 |
| Year 2: KS1 maths       | 102            | 0.141 (0.036; 0.245)                     | + 2                        | 🔒🔒🔒🔒🔒           | £ £ £ £ £       |
| Year 2: KS1 reading     | 102            | 0.117 (0.008; 0.225)                     | + 2                        | 🔒🔒🔒🔒🔒           | £ £ £ £ £       |
| Year 2: KS1 writing     | 102            | 0.159 (0.042; 0.275)                     | + 2                        | 🔒🔒🔒🔒🔒           | £ £ £ £ £       |
| Year 6: KS2 reading     | 98             | -0.030 (-0.209; 0.149)                   | + 0                        | 🔒🔒🔒🔒🔒           | £ £ £ £ £       |
| Year 6: KS2 maths       | 98             | 0.005 (-0.152; 0.162)                    | + 0                        | 🔒🔒🔒🔒🔒           | £ £ £ £ £       |
| <b>Original results</b> |                |  |                            |                 |                 |
| Year 2: KS1 maths       | 102            | 0.149 (0.051; 0.248)                     | + 2                        | 🔒🔒🔒🔒🔒           | £ £ £ £ £       |
| Year 2: KS1 reading     | 102            | 0.104 (0.012; 0.196)                     | + 2                        | 🔒🔒🔒🔒🔒           | £ £ £ £ £       |
| Year 2: KS1 writing     | 102            | 0.138 (0.038; 0.239)                     | + 2                        | 🔒🔒🔒🔒🔒           | £ £ £ £ £       |
| Year 6: KS2 reading     | 98             | 0.103 (-0.056; 0.262)                    | + 2                        | 🔒🔒🔒🔒🔒           | £ £ £ £ £       |
| Year 6: KS2 maths       | 98             | 0.075 (-0.060; 0.210)                    | + 1                        | 🔒🔒🔒🔒🔒           | £ £ £ £ £       |

# Introduction

## Intervention

Magic Breakfast is a charity that supports the provision of breakfast clubs in disadvantaged schools in England, defined as those as having at least 35% of pupils eligible for free school meals or 50% of pupils having been eligible in the previous six years. This evaluation tested the impact of providing primary schools with resources and support to introduce a breakfast club which is before-school, universal, and free over one academic year from September 2014 to July 2015.

The rationale for this intervention is that addressing pupil hunger may lead to improvements in concentration and behaviour in the classroom, fewer absences from school, and ultimately improvements in pupil attainment.

The breakfast club in each school was supported by Magic Breakfast through the provision of free healthy food (as much as required), £300 in capital funding, and advice about establishing and successfully delivering a breakfast club (including the staffing arrangements for supervision of pupils during the breakfast club).

The control condition was 'business as usual' with a 'waitlist' design where schools allocated to the control group were helped to establish their breakfast club the following academic year. Around 40% of schools in the control group that responded to the follow-up headteacher survey adopted some form of breakfast club throughout the year. In this evaluation, we assess the effect of providing schools with resources and support to introduce a universal, free, before-school breakfast club, rather than the effect of schools actually doing so.<sup>1</sup> This means that our estimates are based on all schools taking part in the project, including those that received breakfast club support but did not establish a universal, free, before-school club, and those that did not receive support but established breakfast provision anyway.

The main motivation for schools joining the intervention, reported in the headteacher survey before randomisation and during case study visits, was a concern about pupils arriving at school hungry and the impact that this might have on students' wellbeing, behaviour, and experiences in the classroom. The introduction of Magic Breakfast provision was seen by staff as a way to tackle a number of pre-existing problems:

- that of pupil hunger and, by virtue of this, improve children's wellbeing, concentration, and behaviour in class;
- to improve attendance and punctuality by creating an added incentive for children to arrive at school on time;
- to support parents by reducing the stress of morning routines and providing early morning childcare; and
- to establish and reinforce a welcoming and inclusive community ethos within the school.

## Evaluation objectives

The primary objective of the impact evaluation was to measure the impact of breakfast club provision on pupils' academic attainment by comparing pupil outcomes in the intervention group with a 'business as usual' control group. The secondary objective was to explore the mechanisms through which provision of a breakfast club improves academic attainment, and variation in effectiveness across subgroups of pupils.<sup>2</sup>

---

<sup>1</sup> More detail on the fidelity to the programme (defined by the provision of a universal, free, before-school breakfast club) and take-up of the offer is given in the 'Outcomes and Analysis' section in the original evaluation report.

<sup>2</sup> Further details about primary and secondary outcome analysis can be found in the evaluation protocol (available at: [https://educationendowmentfoundation.org.uk/public/files/Projects/EEF\\_Project\\_Protocol\\_MagicBreakfast.pdf](https://educationendowmentfoundation.org.uk/public/files/Projects/EEF_Project_Protocol_MagicBreakfast.pdf)) and in the original evaluation report.

## Methods

This section outlines the design of the project with a focus on how schools were assigned to treatment and comparison groups and how we have adapted the analysis to account for this non-random assignment. We also provide an overview of participant selection, data collection, and the choice of outcome measures; further details can be found in the original report.

### Study design

The project used a comparison group design. Disadvantaged primary schools from around England signed up to participate in the evaluation. Researchers assigned schools—rather than pupils—to intervention and comparison groups at the outset of the project. The intervention group received support and resources to establish a universal free breakfast club before school in the academic year 2014/2015. The comparison group was assigned to a ‘business as usual’ condition in 2014/2015 and so received none of this support; however, they received the breakfast club support and resources for the two following academic years (2015/2016 and 2016/2017). This ‘waitlist’ design was chosen to minimise attrition from the trial and maximise the survey response rate.

The package of support for breakfast clubs included:

- free breakfast foods (as much as requested) provided by Magic Breakfast; these included cereals, porridge, bagels, wheat biscuits, and juice;
- support from a dedicated school change leader, who offered advice and guidance on establishing school breakfast provision; and
- a £300 capital grant to offset the upfront costs of expanding provision, such as buying new toasters or improving the school cafeteria.

Schools were required to make their own arrangements to cover any additional financial costs. Importantly, schools were also responsible for arranging (and, if necessary, paying for) supervision during the breakfast club. While schools were instructed to use the resources to open a universal, before-school, free breakfast club, in practice many schools instead adapted this model to their own context.<sup>3</sup>

### Intervention and comparison group assignment

The allocation to intervention and comparison groups was intended to be random within strata defined by the school's type (infant or primary/junior), school-level prior attainment (at Key Stage 2 for primary and junior schools and at Key Stage 1 for infant schools), and the share of students with English as an Additional Language (EAL).<sup>4</sup> We defined seven strata (the number of schools in each is reported in parentheses):

- infant, below median KS1 (4);
- infant, above median KS1 (3);<sup>5</sup>
- junior/primary, missing data/small schools (13);

---

<sup>3</sup> See the original report's Methods section for a discussion of the alternative models of provision that were considered.

<sup>4</sup> The trial protocol pre-specified that either the percentage of EAL pupils or the percentage of students eligible for free school meals would be used. The percentage with EAL was chosen as there was more variation in this school characteristic in the sample of schools recruited. Further details are available in the original report.

<sup>5</sup> Publicly-available data collected pre-randomisation based on 2012 information indicated that seven of the schools in our sample were infant schools. However, one school also admitted Year 6 pupils in 2014/2015, so we use a figure of six infant schools elsewhere in the report.

- junior/primary, below median KS2 and below median EAL (18);
- junior/primary, below median KS2 and above median EAL (23);
- junior/primary, above median KS2 and below median EAL (21); and
- junior/primary, above median KS2 and above median EAL (24).

Unfortunately, an error in the randomisation process meant that the assignment to intervention and comparison groups was not actually random. This error was made and discovered by the principal investigator for the impact evaluation.<sup>6</sup> Rather than being assigned to intervention and control group by a randomly generated number within each stratum, schools with a low school code (LAESTAB number) were assigned to the intervention group. The LAESTAB number begins with a three-digit code for the Local Education Authority in which a school is located. These LEA codes are assigned first to schools in Inner London, followed by those in Outer London, Greater Birmingham, Greater Liverpool, and Greater Manchester (see Appendix Y for a list of LEA codes). The effect of this is that a disproportionate number of schools in London and other urban areas were allocated to the intervention group. Table 2 shows the number of intervention and control schools inside and outside London, and by urban/rural classification. This shows that 23 of the 32 schools inside London (72%) that signed up to the trial were allocated to the intervention group, compared to 30 of the 74 schools outside London (41%). Outside London, schools in urban areas were disproportionately assigned to the intervention group, for example 40 of the 59 schools in ‘major urban’ areas (68%).

**Table 2: Geography of schools assigned to intervention and control groups**

| Geography                 | Control   | Intervention | Total      |
|---------------------------|-----------|--------------|------------|
| <b>London</b>             |           |              |            |
| LEA outside London        | 44        | 30           | 74         |
| LEA inside London         | 9         | 23           | 32         |
| <b>Total</b>              | <b>53</b> | <b>53</b>    | <b>106</b> |
|                           |           |              |            |
| <b>Urbanicity in 2011</b> |           |              |            |
| Rural                     | 16        | 3            | 19         |
| Semi-urban                | 18        | 10           | 28         |
| Major Urban               | 19        | 40           | 59         |
| <b>Total</b>              | <b>53</b> | <b>53</b>    | <b>106</b> |

Note: Urbanicity is based on the 2011 six-category urban-rural classification. We define three groups: ‘rural’ (Rural-80, Rural-50, and Significant Rural), ‘semi-urban’ (Other Urban, Large Urban), and ‘major urban’ (Major Urban).

### The comparison group design

In a randomised controlled trial, the allocation to intervention and comparison groups is random. This means that the allocation is completely unaffected by any characteristics of the schools involved. In turn, this means that these two groups should be very similar on average in all their characteristics (both those observed by the research team and those that are not captured in the data). This lets researchers evaluate the impact of a policy by simply comparing the outcomes of intervention and comparison group schools (which do a good job of reflecting what would have happened had the intervention schools not received the breakfast club programme).

Since the original assignment to intervention and comparison groups was not actually random, this argument no longer holds. Instead, the extent to which we can attribute differences in outcomes between intervention and comparison group schools depends on how well matched the two groups are. The more similar they are in

<sup>6</sup> More information can be found in the statement published by the IFS. <https://www.ifs.org.uk/MBevaluation/Statement>

observable and unobservable ways, the more accurate a picture the comparison group's outcomes paint of what would have happened in the intervention schools in the absence of breakfast club provision.

In the original report, we assessed how well matched the two groups were on a range of observable characteristics, including school-level traits such as Ofsted rating, neighbourhood deprivation score, and pupil-level characteristics like gender, ethnicity, and free school meals eligibility. We also tested for balance at the baseline on several of our outcomes of interest, including pupils' eating habits, pupil attainment at the prior Key Stage, and pupil absences in the previous academic year. All in all, we tested for balance on 31 dimensions; the intervention and comparison groups were well-matched on 29 of these—defined as no statistically significant differences between the intervention and control groups. Table 6 from the original report is reproduced (Table 3 below) for further detail. Unfortunately, we are unable to reproduce this table in line with current EEF requirements as we no longer have access to the data required. In particular, we are unable to discuss the level of balance through effect sizes.

**Table 3: Baseline comparison**

| Variable   | Intervention group         |                    | Control group |                    |             |
|--|----------------------------|--------------------|---------------|--------------------|-------------|
|  | School-level (categorical) | n/N (missing)      | Percentage    | n/N (missing)      | Percentage  |
| <b>Religiously affiliated<sup>a</sup></b>                                    |                            | 11/53 (0)          | 20.8          | 11/53 (0)          | 20.8        |
| <b>Community school</b>  |                            | 30/53 (0)          | 56.6          | 32/53 (0)          | 60.4        |
| <b>Academy</b>   |                            | 12/53 (0)          | 22.6          | 10/53 (0)          | 18.9        |
| <b>Voluntary or Foundation school</b>  |                            | 11/53 (0)          | 20.8          | 11/53 (0)          | 20.8        |
| <b>Ofsted: Outstanding<sup>b</sup></b>                                       |                            | 4/53 (0)           | 7.5           | 4/53 (0)           | 7.5         |
| <b>Ofsted: Good<sup>b</sup></b>  |                            | 30/53 (0)          | 56.6          | 33/53 (0)          | 62.3        |
| <b>Ofsted: Satisfactory<sup>b</sup></b>                                      |                            | 12/53 (0)          | 22.6          | 10/53 (0)          | 18.9        |
| <b>Ofsted: Inadequate<sup>b</sup></b>  |                            | 0/53 (0)           | 0.0           | 2/53 (0)           | 3.8         |
| <b>School in urban area<sup>c***</sup></b>                                   |                            | 40/53 (0)          | 75.5          | 19/53 (0)          | 35.8        |
| <b>School in rural area<sup>c***</sup></b>                                   |                            | 3/53 (0)           | 5.7           | 16/53 (0)          | 30.2        |
| <b>'Some' or 'most' pupils badly behaved</b>                                 |                            | 17/53 (0)          | 32.1          | 16/53 (0)          | 30.2        |
| <b>Main motivation for joining trial: improve pupil health and wellbeing</b> |                            | 26/53 (0)          | 49.1          | 25/53 (0)          | 47.2        |
| <b>Main motivation for joining trial: reduce pupil hunger</b>                |                            | 18/53 (0)          | 34.0          | 14/53 (0)          | 26.4        |
| <b>School offers some ad hoc provision before trial</b>                      |                            | 22/53 (0)          | 41.5          | 25/53 (0)          | 47.2        |
| <b>School-level (continuous)</b>   |                            | <b>n (missing)</b> | <b>Mean</b>   | <b>n (missing)</b> | <b>Mean</b> |
| <b>Number of pupils</b>  |                            | 53 (0)             | 333.9         | 53 (0)             | 291.5       |
| <b>% Free School Meals</b>   |                            | 53 (0)             | 43.8          | 53 (0)             | 44.1        |
| <b>% English as Additional Language</b>                                      |                            | 50 (3)             | 35.5          | 47 (6)             | 31.3        |

|  |                      |                   |                      |                   |
|--|----------------------|-------------------|----------------------|-------------------|
| <b>KS2 Average Point Score (overall)<sup>d</sup></b>                   | 46 (7)               | 27.5              | 48 (5)               | 27.4              |
| <b>KS2 Overall Value-Added Measure<sup>d</sup></b>                     | 46 (7)               | 100.2             | 48 (5)               | 100.4             |
| <b>School LSOA's IMD 2010 rank<sup>e</sup></b>                         | 53 (0)               | 15.5              | 53 (0)               | 19.6              |
| <b>School LSOA's IMD 2010 score<sup>e</sup></b>                        | 53 (0)               | 42.5              | 53 (0)               | 41.3              |
| <b>% of pupils with healthy weight<sup>**</sup></b>                    | 44 (9)               | 57.0              | 41 (12)              | 61.2              |
| <b>Pupil-level (categorical)</b>                                       | <b>n/N (missing)</b> | <b>Percentage</b> | <b>n/N (missing)</b> | <b>Percentage</b> |
| <b>Female<sup>f</sup></b>  | 2283/4609 (0)        | 49.5              | 1969/4038 (0)        | 48.8              |
| <b>Ethnicity: White<sup>f</sup></b>                                    | 2505/4546 (63)       | 55.1              | 2228/4000 (38)       | 55.7              |
| <b>Ethnicity: Black<sup>f</sup></b>                                    | 961/4546 (63)        | 21.1              | 653/4000 (38)        | 16.3              |
| <b>Ethnicity: Asian<sup>f</sup></b>                                    | 478/4546 (63)        | 10.5              | 640/4000 (38)        | 16.0              |
| <b>Ethnicity: Mixed<sup>f</sup></b>                                    | 342/4546 (63)        | 7.5               | 252/4000 (38)        | 6.3               |
| <b>FSM: Currently Eligible<sup>f</sup></b>                             | 1553/4574 (35)       | 34.0              | 1430/4009 (29)       | 35.7              |
| <b>FSM: Ever Eligible<sup>f</sup></b>                                  | 2478/4574 (35)       | 54.2              | 2300/4009 (29)       | 57.4              |
| <b>SEN: Any recorded<sup>f</sup></b>                                   | 983/4574 (35)        | 21.5              | 954/4009 (29)        | 23.8              |
| <b>English as Additional Language<sup>f</sup></b>                      | 1786/4568 (41)       | 39.1              | 1645/4003 (35)       | 41.1              |
| <b>Ate breakfast today</b>   | 3211/3526 (1611)     | 91.1              | 2326/2556 (1887)     | 91.0              |
| <b>Ate breakfast at school</b>   | 346/3506 (1631)      | 9.9               | 269/2536 (1907)      | 10.6              |
| <b>Hungry at start of day<sup>*</sup></b>                              | 1230/3212 (1925)     | 38.3              | 786/2356 (2087)      | 33.4              |
| <b>Will eat lunch today</b>  | 3449/3538 (1599)     | 97.5              | 2483/2564 (1879)     | 96.8              |
| <b>Good level of development at FSP<sup>g</sup></b>                    | 1059/2376 (139)      | 44.6              | 838/2012 (130)       | 41.7              |
| <b>Pupil-level (continuous)</b>  | <b>n (missing)</b>   | <b>Mean</b>       | <b>n (missing)</b>   | <b>Mean</b>       |
| <b>Total half-sessions absent in 2013/2014<sup>h</sup></b>             | 4376 (233)           | 13.0              | 3792 (246)           | 13.8              |
| <b>Total authorised half-sessions absent in 2013/2014<sup>h*</sup></b> | 4376 (233)           | 9.5               | 3792 (246)           | 10.3              |

|  |            |      |            |      |
|--|------------|------|------------|------|
| <b>Total unauthorised half-sessions absence in 2013/2014<sup>h</sup></b> | 4376 (233) | 3.6  | 3792 (246) | 3.5  |
| <b>Total half-sessions late in 2013/2014<sup>h</sup></b>                 | 4609 (233) | 0.3  | 4038 (246) | 0.4  |
| <b>Reading points at KS1<sup>i</sup></b>                                 | 1914 (180) | 14.5 | 1698 (197) | 14.4 |
| <b>Writing points at KS1<sup>i</sup></b>                                 | 1914 (180) | 13.3 | 1698 (197) | 13.2 |
| <b>Maths points at KS1<sup>i</sup></b>                                   | 1913 (181) | 14.8 | 1697 (198) | 14.6 |

Notes:

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*)

Pupil-level variables account for the clustering of pupils within schools.

<sup>a</sup> Religious affiliation includes both Church of England and Catholic schools.

<sup>b</sup> Ofsted ratings are the most recent overall effectiveness ratings available as at 2014.

<sup>c</sup> Urban-rural classification is from the six-point 2011 DEFRA classifications of U.K. local authority districts. 'Rural' refers to areas in the bottom three tiers; 'urban' refers to those in Tier 6, 'Major Urban' settlements.

<sup>d</sup> KS2 average point score and overall value-added measures are measured for the previous cohort since test results for the cohort of interest are not available for the year before the intervention.

<sup>e</sup> Index of Multiple Deprivation rank is from the 2010 Department for Communities and Local Government's overall LSOA-level IMD. Lower ranks and higher scores imply greater deprivation.

<sup>f</sup> Innate demographic characteristics (ethnicity, FSM, EAL, etc.) are measured on the cohorts of interest.

<sup>g</sup> A good level of development at Foundation Stage Profile is checked for balance only among Year 2 pupils as the relevant measure of prior attainment for this group.

<sup>h</sup> Absences are measured for the cohorts of interest but prior to the start of the intervention.

<sup>i</sup> KS1 point scores are measured only for pupils in Year 6, who wrote the KS1 tests three years prior to the intervention.

Unusually for a matched comparison design, the schools in our intervention and comparison groups are also well matched on one important unobservable characteristic: all the schools in this project expressed a willingness to participate in the evaluation and so had similar levels of motivation to establish breakfast club provision. In a typical matched comparison design, for example, comparing schools that established a breakfast club with those that did not, the motivation and need for breakfast club provision is highly likely to be different between the two groups. We also surveyed headteachers at the outset of the project, asking about their motivations to join the project. We find that these were very similar across the intervention and comparison groups. For example, in the intervention group, 49.1% of headteachers reported that the main motivation for joining the trial was to improve pupil health and wellbeing, compared to 47.2% in the control group. In the intervention group, 34% of headteachers reported that the main motivation was to reduce pupil hunger, compared to 26.4% in the control group.

We have adapted our analysis methodology to reflect the change from a randomised trial to a comparison group design, and we have accounted for additional characteristics to improve the statistical match between the treatment and comparison groups. We provide more details in this in the 'Analysis' sub-section.

## Participant selection

Schools were eligible for the trial if at least 35% of pupils were currently FSM-eligible and they had no existing breakfast club provision or an existing provision usually attended by 6% of pupils or less. Of the 1,765 primary, infant, and junior schools that were approached that met the criteria for pupil disadvantage, only 374 (21%) met the existing provision eligibility criterion. This meant that 374 of the 1,765 schools that were initially approached were eligible. In addition, schools were required to have agreed to the conditions of the project and the evaluation by signing the Memorandum of Understanding shown in Appendix C in the original report. All pupils in selected schools were eligible for the trial as the breakfast club offer was universal and free. Schools were recruited by the Magic Breakfast project team and the process included ample opportunity to discuss in detail

how the partnership between the school and Magic Breakfast would work, and discuss the challenges faced by the school, such as the perceived impact of hunger.

Parent/guardian opt-out consent for participation in the research analysis and linking between pupil survey and administrative data in the National Pupil Database (NPD) was sought after randomisation to intervention or control group.

Consent forms are available in Appendix C in the original report.

## **Data collection**

The data collected for this evaluation included administrative data (for pupils of parents/guardians that granted opt-out consent) and online and paper surveys of teachers, headteachers, and pupils. We summarise these different measures here, but full details are available in the original report in the Methods section and in Appendix D. Survey response and the level of missing data is shown in Figure 2 of the original report.

### **National Pupil Database**

The results of both teacher- and externally-marked national assessments for all pupils in state-funded schools were collected from the National Pupil Database (NPD). This data was linked to individual students' prior attainment (at the previous Key Stage), their demographics (for example, gender, ethnicity, and FSM eligibility), and their absences in the trial year (2014/2015) as well as the previous year.

### **School performance tables**

The additional school-level continuous measures of prior attainment at KS2 used in the revised specification are collected from school performance tables (available from the Department for Education website), carefully accounting for changes in school codes over time. Prior school performance measures were observed for all schools in 2012, but not observed for four non-infant schools in 2014 and nine non-infant schools in 2013. We include binary variables to indicate missing scores in our preferred specification to include all schools in the analysis.

### **National Child Measurement Programme (NCMP)**

School-level data on average Body Mass Index (BMI) and the share of students at a healthy weight was collected for Year 6 pupils in each school (which were labelled as intervention or comparison group, but not further identified).

### **Pupil survey**

The baseline pupil survey was completed by pupils in Year 2 and Year 6 in intervention and control schools around September–October 2014. Pupils were asked about their eating habits. The survey was repeated at endline, in June–July 2015.

### **Teacher survey**

The baseline teacher survey was completed by teachers of Year 2 and Year 6 classes in intervention and control schools around September–October 2014. The survey measured teachers' perceptions of the typical level of pupil concentration and behaviour in their class at the time of response. The survey was repeated at endline in June–July 2015.

### **Follow-up teacher survey**

The questions and format of the baseline teacher survey were repeated in the follow-up teacher survey completed by teachers of Year 2 and Year 6 teachers in June–July 2015.

### **Baseline headteacher survey**



The baseline headteacher survey was completed by headteachers in intervention and control schools as a condition of taking part in the project. The baseline survey was typically completed between June and August 2014. The baseline survey asked about the motivations for signing up to the Magic Breakfast intervention as well as administrative details.

### **Follow-up headteacher survey**

The follow-up headteacher survey was also completed online using Google Forms, in June–July 2015. For headteachers of comparison schools, the questions included information about whether a breakfast club had been established and, if so, the format of the club. For headteachers of intervention schools, the questions also included information about the format, coverage, and costs of the breakfast club that was implemented as well as the headteacher's opinion of its success.

### **Food orders data**

Schools placed orders with Magic Breakfast, which delivered the requested food free of charge each week. Magic Breakfast provided us with the order sheets for intervention schools during two representative weeks in the spring term (weeks beginning 2 and 9 March 2015) as well as information on the termly porridge orders for the spring term.

### **Outcome measures**

The primary outcome measures are national assessments in English and maths taken by all pupils in state-funded schools in England. These assessments are held at the end of each Key Stage in primary schools when children are Year 2 (aged 6–7, KS1) and Year 6 (aged 10–11, KS2). These assessments are done exclusively by the teacher at KS1 (known as 'teacher assessments') and by both teacher assessments and externally-marked tests at KS2. We use the teacher assessments in reading, writing, and maths at KS1 and the externally marked tests in English and maths at KS2 as our primary outcomes (see the original report for a more detailed discussion of why we chose the externally marked tests as primary outcomes for KS2). In particular, we use the point scores obtained in KS1 and the fine point scores awarded by external markers on the KS2 tests. We standardise each academic outcome within the sample to have a mean of zero and a standard deviation of one. This means that the coefficients derived from the analysis are expressed in terms of standard deviations and are comparable across the academic, behaviour, and concentration outcomes within this evaluation. It should be noted that these standard deviations are related to the sample presented here and are not comparable with standard deviations derived from other samples.

We also identified a set of secondary academic outcomes which are shown in the original evaluation report. For these we use additional teacher-assessed outcomes at KS1 (speaking and listening) and teacher-assessed outcomes at KS2: reading, writing, English, maths, and science.

For all teacher-assessed outcomes, teachers assess students against levels of achievement, which we convert to numeric points using the DfE's 'KS1 and KS2 test and examination point scores used in the 2015 school performance tables' guidance, then standardise as discussed above (DfE, 2015a).

For our primary outcomes we used externally marked tests rather than teacher assessed measures, where available. This is because teacher assessments have the potential to be affected by teachers' knowledge of the intervention or to conflate improvements in mechanisms (such as pupil behaviour) with improvements in attainment. They are also less finely scored than the externally marked tests. For these reasons, they are considered less robust outcome measures than externally marked tests. However, the teacher assessed

measures we use are moderated and used for national performance data and therefore still provide sufficiently reliable outcome measures.<sup>7</sup>

In addition to supplementary attainment measures, other secondary outcomes measures (shown in the original evaluation report and summarised in this report) were chosen to show the mechanisms through which any impact on attainment occurs. These included breakfast eating habits and student hunger, teachers' perceptions of student behaviour and concentration, absences (overall and those likely related to poor health), and late arrivals at school. Information on child health, proxied by healthy weight, was collected from the NCMP.

## Sample size

The target for recruitment of schools was 50 schools per group. This was decided through sample size calculations, shown in the original report.

## Analysis

In the original report, we conducted the analysis based on the random assignment of schools to the intervention or comparison group. We therefore conducted an analysis of Magic Breakfast's effect sizes, calculated as the average difference in the standardised outcome between intervention and control groups. Our original analysis also augmented this by accounting for the randomisation strata, the relevant baseline measure, pupil demographics, and school characteristics. The decision to include additional covariates (over and above randomisation strata) was to improve the precision of the estimates and account for any pre-existing differences in observable characteristics between the intervention and control groups. In detail, these were:

- strata (entered as a series of binary variables);
- relevant baseline measure—
  - for KS1, binary pupil-level variables for whether development measured at the Foundation Stage Profile was at the expected level or missing (relative to not at the expected level);
  - for KS2, binary pupil-level variables for categories of average attainment on KS1 tests: 11–14.99 points (did not achieve grade level), 15–16.99 points (at grade level), 17+ points (above grade level), or missing (that is, fewer than 11 points—did not achieve any level 2);
  - for pupils' breakfast consumption and hunger: pupil-level breakfast consumption and hunger measured in the baseline pupil survey;
  - for teachers' perceptions of pupil behaviour and concentration: classroom-level pupil behaviour and concentration measured in the baseline teacher survey;
  - for attendance: a continuous pupil-level measure of sessions lost in the previous academic year (2013/2014); and
  - for late arrival: a continuous pupil-level measure of late arrivals to sessions in the previous academic year (2013/2014);
- pupil demographics—binary variables for female, ever eligible for FSM, any special educational need, English as an Additional Language, and major ethnic group (Black, Asian, Mixed, and Other, with White as the reference group); and

---

<sup>7</sup> For some subjects our data allow us to compare the impact of the intervention on both teacher assessments and test scores (for example, for KS2 maths and reading). In these cases results are similar which suggests that any bias introduced by teacher assessment is minimal. See the original report for details.

- school characteristics—binary variables for the latest measure of Ofsted effectiveness prior to the intervention ('good', 'requires improvement', 'inadequate', and 'missing', with 'outstanding' as the reference group), rural-urban category ('semi-urban' and 'major urban', with 'rural' as a reference category), continuous variables for percentile of the school's Lower Super Output Area rank on the Index of Multiple Deprivation, and the total number of students in the school.

In light of the actual, non-random assignment of schools to intervention and control groups, there are two important features of this original approach. First, we originally accounted for the relevant baseline measure. This helps to ensure that the estimates represent the value added by the trial (from baseline to the end of the trial) rather than pre-existing differences between intervention and control groups. Second, we included control variables for urbanicity. This will partially account for the disproportionate number of intervention schools in urban areas, though it will not fully capture the additional disproportionate share of intervention schools in London.

In this revised report, we present the results from an updated specification that accounts for two additional characteristics to reflect the original non-random assignment. These are:

- An indicator for whether a school is in London. This means that we are effectively comparing intervention schools to comparison schools in the same type of area (London, other urban, semi-urban, or rural).
- Three continuous variables measuring school-level KS2 attainment in 2012, 2013, and 2014. These prior attainment measures help to account for differences in pre-trial trends in KS2 attainment.<sup>8</sup>

These two additions are particularly important for our primary (academic) outcome variables. This is because the relevant baseline measures for these outcomes are measured several years before the trial (in 2012/2013 for KS1 results, and in 2010/2011 for KS2 results). This means that there is more scope for differences between London and non-London schools to have had an effect, over and above the baseline measure.

By contrast, the baseline measures for behaviour and eating outcomes were measured at the start of the 2014/2015 school year and the baseline measures for absences and late arrivals were measured in the previous school year. This means that the specifications for these outcomes will capture the change over the 2014/2015 school year, comparing pupils in intervention and comparison group schools. Over this short period, there is much less scope for any systematic differences in trends between London and non-London schools to have an effect.

Other than the addition of these two new control variables, the technical details of our analysis are identical to those reported in the original report. These are summarised here, but further details are available in the original report.

All analysis is on an 'intention-to-treat' basis, which means that all schools allocated to the intervention and control groups are used to estimate the effect of the intervention independent of their level of engagement with the intervention, control group activity, or continued participation. The effect size is computed through ordinary least squares (OLS) regression for all outcomes except those relating to pupil hunger (reported in the original report) where we report average marginal effects after logistic regression.

---

<sup>8</sup> This is one of the main concerns with the non-random allocation of schools. Since previous research suggests that attainment in London schools has improved faster than elsewhere in the country (Greaves, Macmillan, and Sibieta, 2014), there is a risk that analysis that does not account for this will attribute this more general 'London effect' to the influence of the Magic Breakfast programme. These value-added control variables mitigate this risk by explicitly accounting for schools' records in improving attainment.

Clustering of pupils within schools is taken into account using robust standard errors (clustered sandwich estimator) that allow observations within schools to be correlated. This approach was used rather than multi-level modelling (or random effects) to account for the clustering of pupils within schools because the model does not impose the assumption that the school-level effect and pupil- and school-level covariates in the model are independent.<sup>9</sup> Stratification is accounted for by including the variables used to stratify as regressors in the OLS regression.

Data on BMI and healthy weight (presented in the original evaluation report) is available at the school level rather than the pupil level. The relevant baseline measure (BMI data for the Year 6 pupils in the year before the trial) is also at the school level under the assumption that the health of pupils in adjacent cohorts within schools is positively correlated.

The small number of observations with missing demographic covariates in the NPD was not included in the final sample. This is 62 pupils at KS1 and 65 pupils at KS2. Appendix F of the original report discusses this sample restriction.

Subgroup analysis (presented in the original evaluation report) is conducted according to the evaluation protocol. Of particular interest is the impact for pupils that have ever been eligible for free school meals, a group important for the EEF and Magic Breakfast. Other subgroups of interest are boys versus girls, students with low prior attainment versus those who achieved the expected level of attainment or better in their last assessment, and students who did not report having breakfast in the baseline surveys versus those who did.

Additional analysis (presented in the original evaluation report) not pre-specified in the evaluation protocol relates to the results for the sub-set of schools where no students mentioned observing Ramadan in the follow-up pupil survey, and the sub-set of schools in the intervention group whose breakfast offer was universal and free (two of the components of the intended model as described in the Trial Design section).

---

<sup>9</sup> Subsequent analysis confirms that our original results are robust to specifications using multi-level modelling and Bayesian analysis.

## Impact evaluation

In this section, we present the revised impacts based on the analytical approach discussed in the previous section. We present these in their own right but also discuss how they compare to the findings reported in our original evaluation report.

Our focus in this revised report is on the primary outcome measures; as discussed in the previous section, these are the results likely to be most affected by the randomisation error. We provide a brief summary of the conclusions from the analysis of secondary outcomes but refer the reader to the original report for a more detailed discussion.

### Academic outcomes

Table 4 reports the primary outcomes: the effect of breakfast club provision on pupils' attainment at KS1 and KS2.

Breakfast club provision has positive, statistically significant impacts on KS1 scores in maths, reading, and writing. Progress in maths and writing was moderately enhanced by the intervention; scores in intervention schools were on average 0.14–0.16 standard deviations higher, equivalent to about two months' progress. The impact for KS1 reading was a slightly smaller, though still significant, at 0.12 standard deviations.

The magnitude of the results at KS1 are very similar to those in the original report. The point estimates for KS1 maths in the revised and original reports are 0.141 and 0.149, respectively, with confidence intervals of 0.036 to 0.245 and 0.051 to 0.248. For KS1 reading, the equivalent point estimates are 0.117 and 0.104, with confidence intervals of 0.008 to 0.225 and 0.012 to 0.196. For KS1 writing, the point estimates are 0.159 and 0.138, with confidence intervals of 0.042 to 0.275 and 0.038 to 0.239. All results in the revised and original report are statistically significant, meaning that we can reject the null hypothesis that the impact is zero.

The magnitude is similar to the estimated effect on KS1 attainment of providing universal free school meals in two relatively disadvantaged pilot areas (Brown *et al.*, 2012). Taken together, these results suggest that breakfast and lunch consumption may have additive effects as the improvements in attainment from breakfast club provision in trial occurred even after the introduction of universal free school meals for all pupils in Year 2.

At KS2, the effects of intervention on primary outcomes are very close to zero and not statistically significant. The magnitude of the results at KS2 are different between the revised and original report. The point estimates for KS2 reading in the revised and original report are -0.030 and 0.103, respectively, with confidence intervals of -0.209 to 0.149 and -0.056 to 0.262. The point estimates for KS2 maths are 0.005 and 0.075, with confidence intervals of -0.152 to 0.162 and -0.060 to 0.210. All results in the revised and original report are not statistically significant, meaning that we can't reject the null hypothesis that the impact is zero.

This is in contrast to the evaluation of universal free school meals provision where universal provision was found to improve academic attainment *more* for KS2 than KS1 pupils (Brown *et al.*, 2012). Possible reasons include:

- the fact that 91% of control schools (that responded to our follow-up survey) offered large-scale breakfast provision to Year 6 pupils during the week of KS2 tests;
- higher breakfast consumption prior to the intervention for Year 6 pupils (as reported in the baseline pupil survey);
- Year 6 pupils being more likely to fast during Ramadan, which occurred during the school year;
- the fact that the KS1 primary outcomes were teacher-assessed and the KS2 primary outcomes were not (see the Outcome Measures section in the original report for more detail on this); and

- in our preferred specification we account for school-level prior attainment at KS2 rather than KS1 as measures for the latter age group are not publicly available. This means we will likely account for pre-existing levels and trends in academic performance better for KS2 outcomes than KS1 outcomes.

In relation to the first point above, both the process evaluation and the headteacher surveys indicate that Year 6 pupils were actively encouraged to take up the breakfast club offer during the KS2 testing period. About 95% of intervention and 91% of control schools that responded to the survey provided breakfast for KS2 students for this short period of the year. If the effects of breakfast provision are mostly immediate (affecting students' ability to focus on the test that day, for example) rather than cumulative (allowing students to learn more each day during the year), the high levels of provision in KS2 control schools could also attenuate the impact of the intervention on academic outcomes assessed by these tests.

In relation to the penultimate point above, it might be possible that the intervention improves behaviour and causes teachers to perceive their class more positively, which affects the teacher-assessed outcomes at KS1 (but not the KS2 tests). However, Table 8 in the original report indicates that the intervention has similar effects on test and teacher assessment scores for reading and maths at KS2, which are assessed in both ways. This suggests that teachers' judgments are not substantially affected by the presence of a breakfast club and are a reliable measure of attainment in this context. Unfortunately, it is not possible to produce histograms of these outcome variables, in line with current EEF guidelines, as we no longer have access to the data required to do so.

**Table 4 (Table 7 in the original report): Primary outcomes—academic attainment**

| Outcome                | Raw means: standardised variables |                          |                |                           | Effect size                              |                           |          |
|------------------------|-----------------------------------|--------------------------|----------------|---------------------------|--|---------------------------|----------|
|                        | Intervention group                |                          | Control group  |                           | n in model<br>(intervention;<br>control) | Effect size<br>(95% CI)   | p-value  |
|                        | n<br>(missing)                    | Mean<br>(95% CI)         | n<br>(missing) | Mean<br>(95% CI)          |  |                           |          |
| <b>KS1<br/>maths</b>   | 2,473<br>(42)                     | 0.204<br>(0.167; 0.242)  | 2113<br>(29)   | 0.053<br>(0.011; 0.095)   | 102<br>(51; 51)                          | 0.141<br>(0.036; 0.245)   | 0.001*** |
| <b>KS1<br/>reading</b> | 2,473<br>(42)                     | 0.206<br>(0.169; 0.242)  | 2113<br>(29)   | 0.075<br>(0.033; 0.116)   | 102<br>(51; 51)                          | 0.117<br>(0.008; 0.225)   | 0.008*** |
| <b>KS1<br/>writing</b> | 2,473<br>(42)                     | 0.228<br>(0.191; 0.266)  | 2113<br>(29)   | 0.097<br>(0.056; 0.138)   | 102<br>(51; 51)                          | 0.159<br>(0.042; 0.275)   | 0.000*** |
| <b>KS2<br/>reading</b> | 2,050<br>(44)                     | 0.047<br>(0.006; 0.089)  | 1857<br>(38)   | -0.015<br>(-0.061; 0.031) | 98<br>(48; 50)                           | -0.030<br>(-0.209; 0.149) | 0.892    |
| <b>KS2<br/>maths</b>   | 2,050<br>(44)                     | 0.008<br>(-0.035; 0.051) | 1857<br>(38)   | 0.027<br>(-0.018; 0.072)  | 98<br>(48; 50)                           | 0.005<br>(-0.152; 0.162)  | 0.909    |

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*).

Standard errors clustered at school level.

Outcomes are the standardised fine point scores for the relevant tests. KS1 outcomes reported in the NPD are teacher-assessed, informed in part by the pupil's test results.

Controls for KS1 outcomes: randomisation strata; Foundation Stage attainment; demographics (sex, ever FSM, ethnic group, SEN, EAL); pre-intervention school characteristics (Ofsted rating, IMD rank, urban-rural, class size).

Controls for KS2 outcomes are the same, but control for prior attainment at KS1 rather than for Foundation Stage attainment. This is done through binary variables for categories of average attainment on KS1 tests: fewer than 11 points (did not achieve any level 2); 11–14.99 points (did not achieve grade level); 15–16.99 points (at grade level); 17+ points (above grade level); missing.

## Secondary outcomes

This section describes the findings from the original evaluation report for secondary outcomes. Unfortunately, it was not possible to repeat all analysis required to replicate the entire report with the preferred specification before the licence to the necessary data expired. However, several of the secondary outcomes will be less sensitive to the change in methodology and so the results from these outcomes are still informative. This is because the original analysis was able to account for relevant and informative baseline characteristics, and so the estimated impact of the trial is over the year of the trial (from baseline to follow-up) rather than from an unknown starting point. These less sensitive secondary outcomes are:

- pupils' breakfast consumption and self-reported hunger;
- pupils' absence and punctuality; and
- teachers' perceptions of the typical level of pupil concentration and behaviour in their class at the time of response.

For other secondary outcomes and related subgroup analysis, we cannot be as sure that the findings are the same when using the preferred (revised) specification. This is because the relevant baseline measure is taken from before the start of the trial, and there may be systematic differences in progress (for example, in teacher-assessed Key Stage 1 to Key Stage 2 scores) between schools allocated to intervention and control groups. These secondary outcomes are limited to alternative academic outcomes at Key Stage 1 and Key Stage 2 and related subgroups analysis.

The results for Body Mass Index were less robust than other secondary outcomes in the original evaluation report as the data is at the school level and the relevant baseline control is the previous academic cohort in the same school. We therefore omit a discussion of these results.

The process evaluation is unaffected by the error in the impact evaluation and therefore all findings in the original report are relevant and informative.

### **Breakfast consumption and hunger**

There was a high level of breakfast consumption before the intervention, which means that any increase in consumption due to the intervention is therefore meaningful. The intervention was found to shift breakfast consumption from home to school as there was a marginal increase in overall breakfast consumption but a 15-percentage point increase in eating breakfast at school. The original evaluation therefore concluded that any direct effect from breakfast consumption on pupil attainment is therefore likely to be due to changing the content and context of breakfast rather than whether any breakfast is consumed.

### **Absence and punctuality**

The results suggest that the intervention led to small reductions in late arrivals and absences, particularly authorised absences (which includes absence due to illness). Punctuality appears to have improved only marginally, although the four schools in the process evaluation did perceive a decrease in the proportion of pupils arriving late to class.

### **Behaviour and concentration**

The breakfast club intervention had a large positive effect on teacher-perceived student behaviour and concentration in the classroom. The trial increased positive behaviour in intervention schools by 48% of a standard deviation; this measure of behaviour incorporates teachers' perception of their classroom's learning environment, lost time due to disruption, disruptive noise, and lost time waiting for students to be quiet. The results for pupil concentration—teachers' perception of the share of their class that was ready to learn first thing in the morning and after lunch, and the share of students that had good and poor concentration in the mornings—were even larger; the intervention improved scores on the concentration index by 64% of a standard

deviation. For a class with an average score on the concentration index at baseline, this is roughly equivalent to moving from the 50th to the 74th percentile.

These findings are consistent with findings from the school case studies where respondents felt that children were more settled, less disruptive, and better able to concentrate as a result of eating breakfast at school.



## Conclusion

### Key conclusions

1. Year 2 children in breakfast club schools made the equivalent of two months' additional progress compared to Year 2 children in the business as usual control group. This result has a low to moderate security rating. These results are similar to the original results although they are now less secure.
2. There is no evidence that breakfast clubs had an impact on Year 6 pupil outcomes. This result has moderate to high security. Compared to the original results, the effect size for Year 6 pupils is lower and less secure.
3. The findings suggest that, where improvements are seen, it is not just eating breakfast that delivers improvements but attending a breakfast club. This could be due to the content of the breakfast itself or to other social or educational benefits of the club.
4. Pupil behaviour, as measured by a teacher survey, improved in breakfast club schools. This is interesting because it shows that breakfast clubs may improve outcomes for children who do not even attend breakfast club by improving classroom environments. This key conclusion is unchanged from the original report.
5. Activities thought to increase take-up of the breakfast provision included promoting it to parents and encouraging all children to attend while sensitively targeting pupils most likely to benefit. The project required additional staff time which some schools found difficult to provide without charging for breakfast. This key conclusion is unchanged from the original report. All findings from the process evaluation are unaffected by the error in the impact evaluation.

### Interpretation

Magic Breakfast was introduced in the context of concerns about pupils arriving at school hungry and the impact that this may be having on their wellbeing, behaviour, and experiences in the classroom. The introduction of Magic Breakfast provision was seen by staff as a way to tackle a number of pre-existing problems:

- to address problems of pupil hunger and, by virtue of this, improve children's wellbeing, concentration, and behaviour in class;
- to improve attendance and punctuality by creating an added incentive for children to arrive at school on time;
- to support parents by reducing the stress of morning routines and providing early morning childcare; and
- to establish and reinforce a welcoming and inclusive community ethos within the school.

At the end of the intervention around 70% of participating intervention schools (that responded to the follow-up headteacher survey) reported that they planned to continue breakfast provision, which reflects the positive experience of the intervention and perceived positive impacts. For example, the majority of headteachers felt that concentration, behaviour, attendance, and attainment had improved.

The impact evaluation findings are largely consistent with headteachers' perceptions. There was evidence to suggest that attainment at the end of Key Stage 1 improved more than would otherwise be expected by chance,

equivalent to around two months' progress. The estimated impact of the trial on attainment at the end of Key Stage 2 was, in contrast, close to zero.

There are multiple mechanisms through which this effect occurred that support the Theory of Change outlined in Figure 1 (in the original report). First, breakfast consumption at school increased, although breakfast consumption overall increased only marginally. This suggests that the school context (which was typically more social) or school food (which was potentially more nutritious) contributed to the improvement in attainment, rather than whether or not breakfast was eaten.

Second, there were large improvements in teacher-reported levels of concentration and behaviour in the classroom. This means that the classroom environment improved for teachers, and presumably pupils, which may have been beneficial for learning. Through this mechanism, even pupils who did not change their breakfast consumption patterns may have benefitted from the intervention.

Attendance and punctuality recorded in administrative data improved slightly, which could be a result of improved health (fewer days of sickness) or a greater incentive to arrive at school on time. An increase in time in school for pupils, and reduced class disruption due to lateness and absence, might therefore have also played a role in improving outcomes.

Finally, there was no evidence to suggest that Year 6 pupils' Body Mass Index (a proxy for health) was affected. This data is less reliable, however, as only school-level averages for Year 6 pupils for a limited set of schools were available, with no individual-level control for BMI before the intervention.

The effect sizes for Key Stage 1 attainment are of a similar magnitude to the effect size in the evaluation of universal free school meals in two pilot areas (Brown *et al.*, 2012) that led to a national roll-out of this policy. The breakfast club intervention occurred at a time when all KS1 pupils were eligible for free school lunches, which suggests that the provision of a school breakfast has an effect over and above that of provision of school lunch. This was particularly noticeable in teachers' reports of classroom behaviour and concentration, where impacts were larger for KS1 than for KS2, despite KS1 pupils already being eligible for a free school meal.

In contrast to the evaluation of universal free school meals, there was no impact on the academic outcomes of Year 6 pupils despite the intervention's significantly larger effect on school breakfast consumption at KS2. This may be because there was a high presence of breakfast clubs in control schools during the week of national tests (above 90% of schools), or because the improvement in concentration and behaviour translated less into KS2 externally marked assessments than KS1 teacher assessments. It is also possible that the slight imbalance in prior attainment for KS1 students (where 44.6% of pupils in the intervention group had a 'good' level of development in the Foundation Stage Profile when entering school compared to 41.7% of pupils in the comparison group) led to upwards bias in the KS1 results reported here.

The findings are broadly consistent with evidence from Wales where a national breakfast club policy was made a duty for maintained schools in 2013. Using an experimental design, Murphy *et al.* (2001) find positive effects for pupil diet, with more 'healthy' breakfast items consumed, but no effect on memory or behaviour. While there were positive but insignificant effects on a test of episodic memory, unfortunately the authors were not able to analyse any objectively reported academic outcomes. An early English randomised controlled trial (Shemilt *et al.*, 2004) found that breakfast provision reduces absences and improves student concentration three months after its introduction (though it also worsens conduct). However, the control group contamination was so severe that they were unable to complete a second follow-up as planned and instead analysed the one-year outcomes as observational data.

Outcomes discussed by case study participants were generally positive; for example, breakfast was perceived to have reduced hunger, increased exposure to new foods, and improved behaviour, attendance, and social skills. It was evident that schools in the case study and follow-up surveys adapted the model of breakfast club provision to their context. Access and funding considerations exerted the biggest influence over the design of

breakfast provision, with some schools deciding that the need to cover the cost of staff supervision outweighed access considerations for keeping the club free of charge.

Data from the four case study schools suggested that delivering breakfast effectively appeared to depend upon a proactive and innovative approach to getting children through the door, the establishment of a solid daily routine built around simple time-saving strategies, and the existence of a well-functioning, mutually-supportive team driving forward and championing the breakfast provision. Delivery challenges arose in the form of a number of different barriers to children accessing the provision, difficulties in balancing supply and demand, and problematic team dynamics. Overall, it was these team dynamics—particularly the extent to which the wider school community had bought in to the importance of Magic Breakfast—which appeared to be most instrumental in determining the success of the breakfast club because they set an overarching tone which either supported or undermined other aspects of delivery.

While commitment to the principle of giving children breakfast in school was unwavering, concerns were widely expressed about the feasibility of continuing the provision once the practical and financial support provided by Magic Breakfast was withdrawn. Specifically, there were concerns about how to remunerate staff and cover the cost of food without introducing a charge or raising the price of breakfast. It was acknowledged by some that an elevated charge would defeat the very purpose of providing breakfast in the first place by making it inaccessible to many of the children it was designed to support. Even where staff were reimbursed through time off in lieu, there was ambiguity over whether or not this arrangement could continue.

Other sustainability-related concerns were linked to the unique role played by individual 'breakfast champions' with fears being expressed that the momentum might be lost if particular individuals were to leave a school. Related to this were fears that schools might be veering towards an overdependence on the goodwill of delivery staff. Not only did this run the risk of causing staff to burn out, it would also be difficult to replicate following any changes in personnel.

Overall, however, schools appeared to be optimistic that they could find a way to ensure there was a continuing breakfast offer at their schools. Schools were open to novel ideas, different types of food, and alternative funders—for example, approaching local businesses for sponsorship.

## Limitations

The results of this evaluation are relevant to schools in a similar social and economic context to the schools in the trial—those that have a higher proportion of pupils eligible for free school meals than the national average and are more likely to be in urban areas. The estimated impacts are generalizable to schools in these similar circumstances as the trial was an effectiveness (or 'real world') trial.

Schools varied their approach to breakfast club provision depending on their constraints and objectives, despite signing up to provide a universal and free breakfast club, which means that fidelity to the tested model was sometimes limited.

The majority of outcomes are accurately and reliably recorded and available for almost all pupils of interest. Primary outcomes at Key Stage 1 are based on teacher assessments rather than externally marked tests, which could introduce some bias if teachers' perceptions are skewed by knowledge of participation in the trial. However, teacher assessments and externally marked tests both taken at Key Stage 2 show a similar effect of the breakfast club where it is possible to compare, which suggests that teacher assessments are reliable indicators of attainment in these schools. In addition, teachers in intervention and control schools should have similar practices for determining teacher-assessed levels; for example, all teacher assessments are moderated and follow the same criteria defined by the Department for Education.

The evaluation has not been able to adequately explore the effect of the intervention on health. The reduction in overall absences from school is consistent with fewer days of illness, but not conclusive. The measurement

of body mass index used in the evaluation is a school-level measure for Year 6 pupils only, where an individual baseline measure is not available, making the estimates more imprecise and subject to other variation.

The pupil survey had a lower response rate than the teacher and headteacher surveys and therefore has the most potential to be influenced by non-random response bias. For example, average points at both KS1 and KS2 were significantly lower among those students whose NPD records did not link to a pupil survey compared to those who could be linked. There is, however, little difference in the estimated effectiveness of the programme on academic outcomes among those who can and cannot be linked to a pupil survey; this provides some reassurance that the effects estimated using the pupil surveys should be reasonably representative of the effects amongst all trial pupils.

Finally, the trial was not implemented as a randomised controlled trial as intended. Instead, the trial has a comparison group design. The intervention group and comparison groups of schools are similar in all school- and pupil-level characteristics that we observe aside from location and the Body Mass Index of earlier cohorts of students, which suggests that the difference estimated between the two groups is a result of the trial. In addition, schools in both groups are the same in an important dimension that is typically not observable to researchers: willingness to participate in a trial and, in this case, desire to implement a breakfast club. These facts, together with the preferred specification which accounts for differences between the intervention group and control group in location in London and prior school attainment measures (in addition to pupil- and school-level characteristics accounted for in the original report) give confidence in the results. However, the lack of randomisation means that the trial has a lower security rating than it otherwise would have done.

## **Future research and publications**

A large number of schools in England provide breakfast clubs of some form, and the direction in U.K. policy seems to be for this to increase. Further research is therefore required to inform schools about the most effective ways of delivering breakfast provision and improving pupil attainment. Through case study visits, this evaluation has suggested that the school context is an important determinant of the feasible model of breakfast club provision, but further quantitative research should provide more evidence on the direct effects of different choices schools can make, for example whether to provide breakfast before school or during school, or with or without charge, for some or all pupils.

For transparency, subject to participating schools remaining anonymous, a dataset containing richer school characteristics used in the preferred specification in this report will be deposited for further external research use. The EEF will also appoint external evaluators to provide an independent assessment of the impact of the study. The IFS will provide all original and revised code to aid this process.

## Revised Appendix B: Security classification of trial findings

### OUTCOME: Key Stage 1 results

| Rating | Criteria for rating  | MDES        | Attrition | Initial score | Adjust  | Final score |
|--------|--|-------------|-----------|---------------|---|-------------|
| 5      | Design   |             |           |               |   |             |
|        | Randomised design  | <= 0.2      | 0-10%     |               |   |             |
| 4      | Design for comparison that considers some type of selection on unobservable characteristics (e.g. RDD, Diff-in-Diffs, Matched Diff-in-Diffs)                                       | 0.21 - 0.29 | 11-20%    |               |   |             |
| 3      | Design for comparison that considers selection on all relevant observable confounders (e.g. Matching or Regression Analysis with variables descriptive of the selection mechanism) | 0.30 - 0.39 | 21-30%    | 3             | Adjustment for threats to internal validity<br>[-1] |             |
| 2      | Design for comparison that considers selection only on some relevant confounders   | 0.40 - 0.49 | 31-40%    |               |   | 2           |
| 1      | Design for comparison that does not consider selection on any relevant confounders   | 0.50 - 0.59 | 41-50%    |               |   |             |
| 0      | No comparator  | >=0.6       | >50%      |               |   |             |

| Threats to validity                       | Threat to internal validity? | Comments   |
|---|------------------------------|--|
| <b>Threat 1: Confounding</b>              | Moderate                     | Study is a matched comparison with a coarse and binary prior attainment measure (reaching the expected level of development using the Foundation Stage Profile). There were some differences in prior attainment favouring the treatment group (44.6% of pupils in the intervention group had a 'good' level of development in the Foundation Stage Profile when entering school compared to 41.7% of pupils in the comparison group) but it is difficult to judge the implication of this due to the low quality nature of the measure. To the extent that differences in this measure are captured by prior attainment, this will be controlled for in the regression results; however, to the extent that this measure does not capture all relevant information, there may be an imbalance between treatment and control groups. |
| <b>Threat 2: Concurrent interventions</b> | No Information               | No information is available around other activities undertaken by schools that may have affected the impact of Magic Breakfast.  |
| <b>Threat 3: Experimental effects</b>     | Moderate                     | Around 40% of the control schools introduced a breakfast club which may reduce the observed impact of Magic Breakfast.   |

|  |          |  |
|--|----------|--|
| <b>Threat 4: Implementation fidelity</b> | Moderate | The intervention was not always implemented as planned as some schools charged for breakfast clubs or otherwise targeted them at particular groups of students.  |
| <b>Threat 5: Missing data</b>            | Low      | Missing data was minimal. No additional analyses were conducted.   |
| <b>Threat 6: Measurement of outcomes</b> | Low      | The study used national standardised assessment that were conducted independently from the study (KS1)   |
| <b>Threat 7: Selective reporting</b>     | Moderate | The original analysis was conducted according to the protocol and statistical analysis plan. Given the nature of the error observed, the analysis in this updated report was not pre-specified but it has been peer reviewed by EEF, independent peer reviewers and will be subject to a 'stress test' by external researchers as referred to in the accompanying statement. |

- **Initial padlock score:** 3 Padlocks—design for comparison that considers some relevant confounders.
- **Reason for adjustment for threats to validity:** 1 Padlock—the pre-test measure is of low-quality and there were small relevant differences in this measure that were controlled for. Some control schools implemented similar breakfast clubs and some treatment schools charged for the breakfast. Taken together, these warrant dropping one padlock.
- **Final padlock score:** initial score adjusted for threats to validity = 2 Padlocks.

## OUTCOME: Key Stage 2 results

| Rating | Criteria for rating  |             |                  | Initial score | Adjust   | Final score |
|--------|--|-------------|------------------|---------------|--|-------------|
|        | <b>Design</b>  | <b>MDES</b> | <b>Attrition</b> |               |  |             |
| 5      | Randomised design  | <= 0.2      | 0-10%            |               |  |             |
| 4      | Design for comparison that considers some type of selection on unobservable characteristics (e.g. RDD, Diff-in-Diffs, Matched Diff-in-Diffs)                                       | 0.21 - 0.29 | 11-20%           |               |  |             |
| 3      | Design for comparison that considers selection on all relevant observable confounders (e.g. Matching or Regression Analysis with variables descriptive of the selection mechanism) | 0.30 - 0.39 | 21-30%           | 3             | Adjustment for threats to internal validity<br>[0] | 3           |
| 2      | Design for comparison that considers selection only on some relevant confounders   | 0.40 - 0.49 | 31-40%           |               |  |             |
| 1      | Design for comparison that does not consider selection on any relevant confounders   | 0.50 - 0.59 | 41-50%           |               |  |             |
| 0      | No comparator  | >=0.6       | >50%             |               |  |             |

| Threats to validity                       | Threat to internal validity? | Comments   |
|---|------------------------------|--|
| <b>Threat 1: Confounding</b>              | Low                          | Study is a matched comparison with a high-quality prior attainment measure (Key Stage 1). There were some very small differences favouring the treatment group in terms of previous attainment but these were accounted for in the analytical model. |
| <b>Threat 2: Concurrent interventions</b> | No Information               | No information is available around other activities undertaken by schools that may have affected the impact of Magic Breakfast.  |
| <b>Threat 3: Experimental effects</b>     | Moderate                     | Some control schools introduced a breakfast club which could bias the estimate of impact downwards.  |
| <b>Threat 4: Implementation fidelity</b>  | Moderate                     | The intervention was not always implemented as planned as some schools charged for breakfast clubs or otherwise targeted them at particular groups of students..   |
| <b>Threat 5: Missing data</b>             | Low                          | Missing data was minimal. No additional analyses were conducted.   |
| <b>Threat 6: Measurement of outcomes</b>  | Low                          | The study used national standardised assessment that were conducted independently from the study (KS2)   |

|                                      |          |   |
|--------------------------------------|----------|---|
| <b>Threat 7: Selective reporting</b> | Moderate | The original analysis was conducted according to the protocol and statistical analysis plan. Given the nature of the error observed, the analysis in this updated report was not pre-specified but it has been peer reviewed by the EEF and independent peer reviewers and will be subject to a 'stress test' by external researchers as referred to in the accompanying statement. |
|--------------------------------------|----------|---|

- **Initial padlock score:** 3 Padlocks—design for comparison that considers all relevant confounders. Prior achievement is accounted for by Key Stage 1 results, which is a stronger measure.
- **Reason for adjustment for threats to validity:** 0 Padlocks—some control schools implemented similar breakfast clubs and some treatment schools charged for the breakfast, contravening the programme that stipulated breakfasts to be free. Taken together, these threats are insufficient to drop a padlock.
- **Final padlock score:** initial score adjusted for threats to validity = 3 Padlocks.



## Revised Appendix H: Analysis code

This appendix provides the analysis code to create Table 5, the headline table of results. For the code behind the rest of the tables, please contact the authors directly. Note that the updates to this code (for the revised specification) are highlighted in red text.

### \*\*\* TABLE 5 - PRIMARY ACADEMIC OUTCOMES \*\*\*

\*\*\*\*\*

```
tempname file
```

```
file open `file' using `"$ptables\Analysis 7 (NPD effect size)\reportable tables.csv"', write text replace
```

```
// Headline results - tests
```

```
file write `file' "Table A: Headline academic outcomes" _new
```

```
file write `file' " , Raw means, , , , Effect size, , " _new
```

```
file write `file' " , Intervention group, , Control group, , , , " _new
```

```
file write `file' "Outcome, n (missing), Mean (95% CI), n (missing), Mean (95% CI), n in model  
(intervention; control), Effect size (95% CI), p-value" _new
```

```
foreach yvar in ks1_readpoints_z ks1_writpoints_z ks1_matpoints_z ks2_readfine_z ks2_matfine_z  
{
```

```
    if "`yvar'" == "ks1_readpoints_z"    local name = "KS1 reading"
```

```
    if "`yvar'" == "ks1_writpoints_z"    local name = "KS1 writing"
```

```
    if "`yvar'" == "ks1_matpoints_z"    local name = "KS1 maths"
```

```
    if "`yvar'" == "ks2_readfine_z"    local name = "KS2 reading"
```

```
    if "`yvar'" == "ks2_matfine_z"    local name = "KS2 maths"
```

```
    if "`yvar'" == "ks1_readpoints_z" | "`yvar'" == "ks1_writpoints_z" | "`yvar'" ==  
"ks1_matpoints_z" {
```

```
        local ctrlattain = "ifspattain"
```

```
        local stage "ks1"
```

```
    }
```

```
    if "`yvar'" == "ks2_readfine_z" | "`yvar'" == "ks2_matfine_z" {
```

```
        local ctrlattain = "iks1attain"
```

```
        local stage "ks2"
```

```
    }
```

```
// Treatment Group raw means
```

```
sum `yvar' if sample == 1 & hasnpd == 1 & _treat == 1 & fullresults == 1 & hasdemo == 1 &
```

```
stage == "stage"
```

```
cii mean r(N) r(mean) r(sd)
```

```

local TN = trim("": display %6.0f r(N)""
local Tmean = trim("": display %6.3f r(mean)""
local Tcilo = trim("": display %6.3f r(lb)""
local Tcihi = trim("": display %6.3f r(ub)""

count if sample == 1 & hasnpd == 1 & stage == "`stage'" & _treat == 1
local Tmiss_temp = r(N) - `TN'
local Tmiss = trim("": display %6.0f `Tmiss_temp'""

// Control Group raw means
sum `yvar' if sample == 1 & hasnpd == 1 & _treat == 0 & fullresults == 1 & hasdemo == 1 &
stage == "`stage'"
cii mean r(N) r(mean) r(sd)
local CN = trim("": display %6.0f r(N)""
local Cmean = trim("": display %6.3f r(mean)""
local Ccilo = trim("": display %6.3f r(lb)""
local Ccihi = trim("": display %6.3f r(ub)""

count if sample == 1 & hasnpd == 1 & stage == "`stage'" & _treat == 0
local Cmiss_temp = r(N) - `CN'
local Cmiss = trim("": display %6.0f `Cmiss_temp'""

// Formatting
local paro (
local parc )
local semicn ;

local col2 `TN' `paro`Tmiss`parc'
local col3 `Tmean' `paro`Tcilo`semicn`Tcihi`parc'
local col4 `CN' `paro`Cmiss`parc'
local col5 `Cmean' `paro`Ccilo`semicn`Ccihi`parc'

// Number of schools
gen tempflagT = (sample == 1 & hasnpd == 1 & _treat == 1 & sample == 1 & stage == "`stage'"
& fullresults == 1 & hasdemo == 1 & !missing(`yvar'))
bys tempflagT urn: gen urnflagT = _n
count if urnflagT == 1 & tempflagT == 1
local NschIT = r(N)

gen tempflagC = (sample == 1 & hasnpd == 1 & _treat == 0 & sample == 1 & stage == "`stage'"
& fullresults == 1 & hasdemo == 1 & !missing(`yvar'))
bys tempflagC urn: gen urnflagC = _n
count if urnflagC == 1 & tempflagC == 1
local NschIC = r(N)

local Nschltot = `NschIT' + `NschIC'

local col6 `Nschltot' `paro`NschIT`semicn`NschIC`parc'
drop tempflagT tempflagC urnflagT urnflagC

// Effect size
reg `yvar' i._treat i.strat3 `ctrlattain' i.female i.everfsm_6_spr15 i.ethnicshort i.anysen i.eal
i.effectiveness_2014 imd2010rank i.urbrur2 i.london census_totpupsendn ///
taps2014 missing_2014 taps2013 missing_2013 taps2012 infant ///
if sample == 1 & stage == "`stage'" & fullresults == 1 & hasdemo == 1, cluster(urn)

```

```

assert `Nschltot' == e(N_clust)

local Ecoef = trim("` : display %6.3f _b[1._treat]'"")
local Ecilo = trim("` : display %6.3f _b[1._treat] - invttail(e(df_r),0.025) * _se[1._treat]'"")
local Ecihi = trim("` : display %6.3f _b[1._treat] + invttail(e(df_r),0.025) * _se[1._treat]'"")
local Epval = trim("` : display %6.3f 2*(1-normprob(abs(_b[1._treat]/_se[1._treat])))'"")

if      `Epval' <= 0.10 local sigstars = "*"
if      `Epval' <= 0.05 local sigstars = "***"
if      `Epval' <= 0.01 local sigstars = "****"
if      `Epval' > 0.10 local sigstars = ""

// Formatting
local col7 `Ecoef' `paro' `Ecilo' `semicn' `Ecihi' `parc'
local col8 `Epval' `sigstars'

file write `file' "` name" " , " `col2" " , " `col3" " , " `col4" " , " `col5" " , " `col6" " , " `col7" " , " `col8" _new
}

```

## Appendix X: Documenting changes to original report

This appendix documents changes made to the original report to produce the revised report. Page numbers refer to the revised report.

- Page 6:
  - Description of evaluation design changed from 'randomised controlled trial' to 'comparison group design'.
  - Updated key conclusion 1 and 2.
  - Change in padlock rating from 'moderate to high security' to 'low to moderate security' for KS1 and 'moderate to high security' for KS2.
  - Removed description of study as 'randomised controlled trial'.
  - Additional discussion of the validity of the comparison group design.
- Page 7:
  - Continued additional discussion of the validity of the comparison group design.
  - Updated discussion of the limitations of the study.
  - Summary of the findings at KS1 and KS2 in comparison to the original findings.
  - Removed of discussion of results for pupils eligible and not eligible for Free School Meals.
- Page 8:
  - Table 1 shows the original and revised estimates of impact at KS1 and KS2.
  - Table 1 omits estimates for pupils eligible and not eligible for Free School Meals.
- Page 10:
  - Removed discussion of theory of change.
  - Removed theory of change (Figure 1).
  - Removed discussion of choice of breakfast club delivery for the intervention.
  - Removed 'Background evidence' section.
  - Shortened description of the evaluation objectives.
  - Removed link to evaluation protocol.
- Page 11:
  - Removed 'Ethical review' section.
  - Removed 'Project team' section.
  - Removed 'Trial registration' section.
  - Revised 'Trial design' to 'Study design' and updated the text to reflect the revised design.

- The 'Participant selection' section is moved to later in the document (page 16) and slightly shortened.
- The 'Data collection' section is moved to later in the document (page 16) and shortened.
- The 'Outcome measures' section is moved to later in the document (page 17).
- The 'Sample size' section is moved to later in the document (page 18) and shortened to give the target number of schools recruited.
- Revised 'Randomisation' section to 'Intervention and comparison group assignment' section. This section describes the randomisation error in detail.
- The 'Analysis' section is moved to later in the document (page 18) and revised to describe the revised methodology.
- Removed 'Implementation and process evaluation' section.
- Removed 'Costs' section.
- Removed 'Timeline' section.
- Page 20:
  - Removed flow diagram for participants.
  - Removed Table 5 (minimum detectable effect size).
  - Removed 'Pupil characteristics' section. Table 6 (baseline comparison) is included in the new 'Intervention and comparison group assignment' (page 13).
  - Removed 'Sample sizes and missing data' section.
  - Removed 'Outcomes and analysis' section.
  - Updated the description of results at KS1 and KS2, giving comparison to the original results.
- Page 21:
  - Updated Table 7 to show revised results.
  - Created a new section to discuss the results for secondary outcomes. These results have not been updated. This section discusses the validity of these results.
  - Removed Table 9 (sensitivity analysis—academic outcomes) and shortened related discussion.
  - Removed Table 10 (secondary outcomes—breakfast consumption and hunger) and shortened related discussion.
  - Removed Table 11 (secondary outcomes—absence and late arrivals) and shortened related discussion.
  - Removed Table 12 (secondary outcomes—classroom behaviour and concentration) and shortened related discussion.
  - Removed Table 13 (secondary outcomes—Body Mass Index and healthy weight) and shortened related discussion.

- Removed 'Summary' section.
  - Removed 'Subgroup analysis' section.
  - Removed 'Cost' section.
  - Removed 'Process evaluation' section. This section is unaffected by the error in the randomisation code.
- Page 23:
    - Updated key conclusions 1 and 2.
  - Page 25:
    - Expanded discussion of previous evidence.
  - Page 26:
    - Additional limitation to the study described, that the study design was comparison group rather than randomised controlled trial.
    - Additional note in 'Future research and publications' that the dataset needed to re-analyse the study will be deposited.

## Appendix Y: Local Education Authority (LEA) codes

The school code (LAESTAB number) for each school begins with a three-digit code for the Local Education Authority in which a school is located. These LEA codes are assigned first to schools in Inner London, followed by those in Outer London, Greater Birmingham, Greater Liverpool, and Greater Manchester. Table Y1 outlines the geographical grouping of LEA codes.

**Table Y1: Local Education Authority codes and areas**

| Local Education Authority Code | Area   |
|--------------------------------|--|
| 201–213                        | Inner London   |
| 301–320                        | Outer London   |
| 330–336                        | Greater Birmingham   |
| 340–344                        | Greater Liverpool  |
| 350–359                        | Greater Manchester   |
| 370–373                        | South Yorkshire  |
| 380–384                        | West Yorkshire   |
| 390–394                        | Gateshead/Newcastle/Tyneside   |
| 420                            | Isles of Scilly  |
| 800s and 900s                  | Remaining unitary authorities and counties;<br>no clear geographic or urbanicity pattern |

## Appendix Z: Additional robustness checks

### Explaining the issue: non-random allocation to treatment group

Our analysis allocated treatment to schools within strata. Strata were defined by infant/junior+primary; primary school with missing data or small number of pupils; above/below median KS2; and above/below median EAL in the sample (Table Z1).

**Table Z1: Strata and treatment indicator**

| Strata   | Control   | Intervention | Total      |
|--|-----------|--------------|------------|
| 1. Infant: below median KS1                    | 2         | 2            | 4          |
| 2. Infant: above median KS1                    | 1         | 2            | 3          |
| 3. Primary: missing data/small school          | 7         | 6            | 13         |
| 4. Primary: below median KS2; below median EAL | 9         | 9            | 18         |
| 5. Primary: below median KS2; above median EAL | 11        | 12           | 23         |
| 6. Primary: above median KS2; below median EAL | 11        | 10           | 21         |
| 7. Primary: above median KS2; above median EAL | 12        | 12           | 24         |
| <b>Total</b>                                   | <b>53</b> | <b>53</b>    | <b>106</b> |

Given the definition of strata, some strata are almost universally in or outside London. For example, Table Z2 shows that those with below median EAL are all outside London. Those with above median KS2 and above median EAL are predominantly inside London.

**Table Z2: Strata and London**

| Strata   | LA outside London | LA inside London | Total      |
|--|-------------------|------------------|------------|
| 1. Infant: below median KS1                    | 3                 | 1                | 4          |
| 2. Infant: above median KS1                    | 3                 | 0                | 3          |
| 3. Primary: missing data/small school          | 13                | 0                | 13         |
| 4. Primary: below median KS2; below median EAL | 18                | 0                | 18         |
| 5. Primary: below median KS2; above median EAL | 13                | 10               | 23         |
| 6. Primary: above median KS2; below median EAL | 21                | 0                | 21         |
| 7. Primary: above median KS2; above median EAL | 3                 | 21               | 24         |
| <b>Total</b>                                   | <b>74</b>         | <b>32</b>        | <b>106</b> |

Unfortunately, rather than being randomly allocated within strata to treatment or control group as desired, schools were non-randomly sorted by their school code (LAESTAB). As London schools have the lowest LA codes (Inner London then Outer London), followed by Birmingham, this makes the treatment group disproportionately urban (Table Z3) and in London (Table Z4). Where there are London schools in the control group, it is only because the treatment group in that stratum is exclusively in London.

**Table Z3: Urban and allocation to intervention group**

| Urban/rural classification in 2011 | Control   | Intervention | Total      |
|------------------------------------|-----------|--------------|------------|
| Rural-80                           | 4         | 0            | 4          |
| Rural-50                           | 5         | 1            | 6          |
| Significant Rural                  | 7         | 2            | 9          |
| Other Urban                        | 14        | 5            | 19         |
| Large Urban                        | 4         | 5            | 9          |
| Major Urban                        | 19        | 40           | 59         |
| <b>Total</b>                       | <b>53</b> | <b>53</b>    | <b>106</b> |

**Table Z4: London and allocation to intervention group**

| London            | Control   | Intervention | Total      |
|-------------------|-----------|--------------|------------|
| LA outside London | 44        | 30           | 74         |
| LA inside London  | 9         | 23           | 32         |
| <b>Total</b>      | <b>53</b> | <b>53</b>    | <b>106</b> |

Comparing the outcomes for the treatment and control groups without other controls (for example prior attainment) would conflate the impact of the treatment with any existing differences between the two groups. For example, London schools tend to out-perform schools elsewhere in the country, so a treatment group that is predominantly London-based would tend to have higher test scores (which we would mistakenly interpret as an effect of the treatment).

In reality, our main results for the initial EEF evaluation report already included a range of controls aimed at accounting, to some extent, for the pre-existing differences in attainment between schools allocated to the treatment and control groups:

at school level: dummies for stratum; dummies for 2014 Ofsted effectiveness rating; IMD 2010 rank; dummies for urbanicity (rural/semi-urban/major urban); and number of pupils in the school; and

at pupil level: prior attainment; ever-FSM; ethnic group dummies; SEN; and EAL.

The following sections describe the robustness checks we have conducted to test whether the positive results reported in the evaluation report for the Magic Breakfast trial are likely to be attributable to the trial or to pre-existing characteristics of the schools assigned to control and treatment groups. The robustness checks are divided into pupil-level analysis, which augments the estimation used in the EEF evaluation report, and school-level analysis, which uses a difference-in-difference design.

The final section discusses further robustness checks that could be conducted with additional data from the Department for Education.



## Robustness checks: pupil-level analysis augmenting the estimation used in the EEF evaluation report

### Controlling for urbanicity

Our main results in the previous EEF evaluation report already controlled for urbanicity in three categories (rural, semi-urban, and major urban). However, given the differences between London and the rest of the country and the importance of London in our treatment allocation, our first robustness checks control explicitly for London using a binary indicator (equal to one if the school is in London and zero otherwise).

The headline results of this were that the improvement in attainment at KS1 was similar, but close to zero at KS2 (Table Z5).

**Table Z5: Adding London dummy to specification**

| Outcome            | Original EEF report      |          | Adding London dummy       |          |
|--------------------|--------------------------|----------|---------------------------|----------|
|                    | Effect size (95% CI)     | p-value  | Effect size (95% CI)      | p-value  |
| <b>KS1 reading</b> | 0.104<br>(0.012; 0.196)  | 0.025**  | 0.155<br>(0.059; 0.252)   | 0.001*** |
| <b>KS1 writing</b> | 0.138<br>(0.038; 0.239)  | 0.006*** | 0.147<br>(0.037; 0.257)   | 0.008*** |
| <b>KS1 maths</b>   | 0.149<br>(0.051; 0.248)  | 0.003*** | 0.197<br>(0.092; 0.303)   | 0.000*** |
| <b>KS2 reading</b> | 0.103<br>(-0.056; 0.262) | 0.198    | -0.012<br>(-0.195; 0.170) | 0.892    |
| <b>KS2 maths</b>   | 0.075<br>(-0.060; 0.210) | 0.27     | 0.009<br>(-0.142; 0.159)  | 0.909    |

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*). Standard errors clustered at school level.

This is also the case when including Inner London and London binary variables (Table Z6) although the estimated coefficients for KS2 are now negative (but not significantly different from zero).

**Table Z6: Adding London and Inner London dummy to specification**

| Outcome            | Original EEF report     |          | Adding London dummy     |          |
|--------------------|-------------------------|----------|-------------------------|----------|
|                    | Effect size (95% CI)    | p-value  | Effect size (95% CI)    | p-value  |
| <b>KS1 reading</b> | 0.104<br>(0.012; 0.196) | 0.025**  | 0.141<br>(0.048; 0.234) | 0.003*** |
| <b>KS1 writing</b> | 0.138<br>(0.038; 0.239) | 0.006*** | 0.135<br>(0.030; 0.241) | 0.011**  |
| <b>KS1 maths</b>   | 0.149<br>(0.051; 0.248) | 0.003*** | 0.139<br>(0.029; 0.249) | 0.012**  |

|                    |                          |       |                           |       |
|--------------------|--------------------------|-------|---------------------------|-------|
| <b>KS2 reading</b> | 0.103<br>(-0.056; 0.262) | 0.198 | -0.053<br>(-0.235; 0.129) | 0.563 |
| <b>KS2 maths</b>   | 0.075<br>(-0.060; 0.210) | 0.27  | -0.086<br>(-0.226; 0.054) | 0.221 |

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*).  
Standard errors clustered at school level.

### London/non-London subsamples

It is also possible that the effects of treatment or of other controls on attainment differ between London and the rest of the country. We therefore also ran the same specification as our headline EEF results for the sample of schools inside and outside London separately. This specification is more flexible but has lower power due to the smaller number of observations in each sub-sample. Results here are also in line with the EEF report for KS1 results, although not statistically significant for reading and writing for schools inside London (Table Z7). KS2 reading in London has a negative sign, although not statistically significant. The results outside London are more uniformly positive.

**Table Z7: London and non-London sub-samples**

| Outcome            | Original EEF report      |          | London                    |         | Non-London               |         |
|--------------------|--------------------------|----------|---------------------------|---------|--------------------------|---------|
|                    | Effect size<br>(95% CI)  | p-value  | Effect size<br>(95% CI)   | p-value | Effect size<br>(95% CI)  | p-value |
| <b>KS1 reading</b> | 0.104<br>(0.012; 0.196)  | 0.025**  | 0.088<br>(-0.070; 0.246)  | 0.256   | 0.093<br>(-0.000; 0.186) | 0.047** |
| <b>KS1 writing</b> | 0.138<br>(0.038; 0.239)  | 0.006*** | 0.118<br>(-0.091; 0.327)  | 0.247   | 0.103<br>(0.001; 0.204)  | 0.044** |
| <b>KS1 maths</b>   | 0.149<br>(0.051; 0.248)  | 0.003*** | 0.152<br>(-0.001; 0.305)  | 0.042** | 0.099<br>(-0.025; 0.223) | 0.113   |
| <b>KS2 reading</b> | 0.103<br>(-0.056; 0.262) | 0.198    | -0.119<br>(-0.395; 0.157) | 0.377   | 0.062<br>(-0.130; 0.254) | 0.519   |
| <b>KS2 maths</b>   | 0.075<br>(-0.060; 0.210) | 0.27     | 0.038<br>(-0.241; 0.318)  | 0.781   | 0.010<br>(-0.137; 0.157) | 0.893   |

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*).  
Standard errors clustered at school level.

### Sufficient overlap between treatment and control schools within strata

The robustness checks that control for London do not account for the non-random assignment of treatment status within strata. In particular, we might be concerned about the lack of ‘common sample’ (pairs of comparable schools that are assigned to the treatment and control groups).

Table Z8 summarises the breakdown of London/non-London and urban/non-urban by stratum.

**Table Z8: Split of London/urban/non-urban within strata and treatment and control**

| Strata | Control |                  |                      | Intervention |                  |                      | Compare Intervention/Control within group |
|--------|---------|------------------|----------------------|--------------|------------------|----------------------|---|
|        | London  | Non-London urban | Non-London non-urban | London       | Non-London urban | Non-London non-urban |   |
| 1      | 0       | 0                | 2                    | 1            | 0                | 1                    | Yes                                       |
| 2      | 0       | 0                | 1                    | 0            | 0                | 2                    | Yes                                       |
| 3      | 0       | 0                | 7                    | 0            | 3                | 3                    | Yes                                       |
| 4      | 0       | 0                | 9                    | 0            | 3                | 6                    | Yes                                       |
| 5      | 0       | 6                | 5                    | 10           | 1                | 1                    | Yes, limited                              |
| 6      | 0       | 1                | 10                   | 0            | 10               | 0                    | Yes, limited                              |
| 7      | 9       | 3                | 0                    | 12           | 0                | 0                    | Yes                                       |

Note: Stratum 1: Infant—below median KS1; 2: Infant—above median KS1; 3: Primary—missing data/small school; 4: Primary—below median KS2, below median EAL; 5: Primary—below median KS2, above median EAL; 6: Primary: above median KS2, below median EAL; 7: Primary: above median KS2, above median EAL.

**Table Z9: Split of Inner-London/Outer-London/urban within stratum 7 and treatment and control**

| Strata | Control      |              |                  | Intervention |              |                  | Compare intervention/control within group |
|--------|--------------|--------------|------------------|--------------|--------------|------------------|---|
|        | Inner London | Outer London | Non-London urban | Inner London | Outer London | Non-London urban |   |
| 7      | 2            | 7            | 3                | 12           | 0            | 0                | Yes, limited                              |

Note: Stratum 7: Primary—above median KS2, above median EAL.

From the balance of treatment and control across strata and London/urban/non-urban areas shown in Table 8 we decided that the following sub-samples would be informative:

treatment effect for strata 1, 2, 3, and 4 in non-London, non-urban areas (with the caveat that there might be some ordering in the non-urban LEA codes that still means there is selection into treatment in this group); and

treatment effect for stratum 7 in London (with the caveat that the treatment group are disproportionately in Inner London—see Table Z9).

The results of these specifications are given in Table Z10. For KS1, it is positive and significant (aside from KS1 maths) in the non-London non-urban strata (1, 2, 3, and 4) and positive but not significant in the London stratum (7). For KS2 reading, the estimate is negative in both samples although not significant. For KS2 maths, the estimate is positive although not significant in both samples.

**Table Z10: Results within strata with sufficient overlap**

| Outcome            | Original EEF report      |          | Strata 1,2,3,4 (non-London non-urban) |          | Strata 7 (London)         |         |
|--------------------|--------------------------|----------|---------------------------------------|----------|---------------------------|---------|
|                    | Effect size (95% CI)     | p-value  | Effect size (95% CI)                  | p-value  | Effect size (95% CI)      | p-value |
| <b>KS1 reading</b> | 0.104<br>(0.012; 0.196)  | 0.025**  | 0.151<br>(-0.023; 0.325)              | 0.077*   | 0.104<br>(-0.047; 0.254)  | 0.149   |
| <b>KS1 writing</b> | 0.138<br>(0.038; 0.239)  | 0.006*** | 0.284<br>(0.098; 0.469)               | 0.002*** | 0.128<br>(-0.077; 0.333)  | 0.191   |
| <b>KS1 maths</b>   | 0.149<br>(0.051; 0.248)  | 0.003*** | 0.192<br>(-0.053; 0.436)              | 0.109    | 0.169<br>(0.021; 0.316)   | 0.016** |
| <b>KS2 reading</b> | 0.103<br>(-0.056; 0.262) | 0.198    | -0.114<br>(-0.368; 0.140)             | 0.354    | -0.051<br>(-0.308; 0.206) | 0.679   |
| <b>KS2 maths</b>   | 0.075<br>(-0.060; 0.210) | 0.27     | 0.085<br>(-0.160; 0.330)              | 0.475    | 0.109<br>(-0.191; 0.408)  | 0.449   |

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*). Standard errors clustered at school level.

These results—which still have the set of controls that we used in the EEF report along with dummies for (inner/outer) London as appropriate—should account for variation in the prior attainment and other differences between pupils in treatment and control groups within a group of schools where there are some comparable treatment and control schools.

### Additional school-level controls for prior attainment

These specifications augment the estimation in the original EEF evaluation report by controlling for school-level prior attainment in addition to pupil-level prior attainment. This should account for any pre-existing differences in outcomes between schools allocated to the treatment and control groups. Multiple specifications for school-level prior attainment were tested. These are:

- continuous measures of school-level prior attainment in 2012, 2013, and 2014, with binary variables to account for missing data where applicable;
- continuous measure of school-level prior attainment in 2012, where there is no missing data;
- continuous measure of school-level prior attainment in 2014, with a binary variable to account for missing data where applicable;
- discrete measure of school-level prior attainment in 2012 (quartiles), where there is no missing data; and
- discrete measure of school-level prior attainment in 2014 (quartiles), with a binary variable to account for missing data where applicable.

The specifications reported also include a binary variable equal to one if the school is inside London and zero otherwise.

Results here are also in line with the original EEF evaluation report for KS1 results, with a similar magnitude and statistical significance. Results for KS2 are close to zero but not significantly different from those reported in the original evaluation report in any specification.

**Table Z11: Results accounting for school-level prior attainment**

| Adding London dummy and a continuous measure of school-level prior attainment |                          |          |  |          |                                   |          |                                   |          |
|---|--------------------------|----------|--|----------|-----------------------------------|----------|-----------------------------------|----------|
|   | Original EEF report      |          | 1. Continuous school-level (2012, 2013 and 2014) |          | 2. Continuous school-level (2012) |          | 3. Continuous school-level (2014) |          |
| Outcome   | Effect size (95% CI)     | p-value  | Effect size (95% CI)                             | p-value  | Effect size (95% CI)              | p-value  | Effect size (95% CI)              | p-value  |
| <b>KS1 reading</b>  | 0.104<br>(0.012; 0.196)  | 0.025**  | 0.141<br>(0.036; 0.245)                          | 0.001*** | 0.144<br>(0.044; 0.244)           | 0.004*** | 0.144<br>(0.045; 0.243)           | 0.004*** |
| <b>KS1 writing</b>  | 0.138<br>(0.038; 0.239)  | 0.006*** | 0.117<br>(0.008; 0.225)                          | 0.008*** | 0.130<br>(0.021; 0.240)           | 0.018**  | 0.130<br>(0.020; 0.240)           | 0.019**  |
| <b>KS1 maths</b>  | 0.149<br>(0.051; 0.248)  | 0.003*** | 0.159<br>(0.042; 0.275)                          | 0.000*** | 0.168<br>(0.060; 0.275)           | 0.002*** | 0.164<br>(0.055; 0.272)           | 0.003*** |
| <b>KS2 reading</b>  | 0.103<br>(-0.056; 0.262) | 0.198    | -0.030<br>(-0.209; 0.149)                        | 0.892    | -0.006<br>(-0.187; 0.175)         | 0.948    | -0.025<br>(-0.207; 0.157)         | 0.782    |
| <b>KS2 maths</b>  | 0.075<br>(-0.060; 0.210) | 0.27     | 0.005<br>(-0.152; 0.162)                         | 0.909    | 0.023<br>(-0.121; 0.167)          | 0.752    | -0.003<br>(-0.161; 0.154)         | 0.97     |

Stars for statistical significance: p <= 0.10 (\*), p <= 0.05 (\*\*), p <= 0.01 (\*\*\*).  
Standard errors clustered at school level.

**Table Z11 continued: Results accounting for school-level prior attainment**

| Adding London dummy and a discrete measure of school-level prior attainment |                          |          |                                 |          |                                 |          |
|---|--------------------------|----------|---------------------------------|----------|---------------------------------|----------|
| Outcome   | Original EEF report      |          | 4. Discrete school-level (2012) |          | 5. Discrete school-level (2014) |          |
|   | Effect size (95% CI)     | p-value  | Effect size (95% CI)            | p-value  | Effect size (95% CI)            | p-value  |
| <b>KS1 reading</b>  | 0.104<br>(0.012; 0.196)  | 0.025**  | 0.172<br>(0.072; 0.272)         | 0.001*** | 0.153<br>(0.056; 0.251)         | 0.002*** |
| <b>KS1 writing</b>  | 0.138<br>(0.038; 0.239)  | 0.006*** | 0.125<br>(0.016; 0.235)         | 0.023**  | 0.136<br>(0.029; 0.243)         | 0.012**  |
| <b>KS1 maths</b>  | 0.149<br>(0.051; 0.248)  | 0.003*** | 0.195<br>(0.083; 0.307)         | 0.001*** | 0.176<br>(0.070; 0.283)         | 0.001*** |
| <b>KS2 reading</b>  | 0.103<br>(-0.056; 0.262) | 0.198    | 0.028<br>(-0.116; 0.172)        | 0.701    | 0.001<br>(-0.184; 0.185)        | 0.995    |
| <b>KS2 maths</b>  | 0.075<br>(-0.060; 0.210) | 0.27     | 0.045<br>(-0.111; 0.201)        | 0.566    | 0.028<br>(-0.116; 0.173)        | 0.699    |

Stars for statistical significance: p <= 0.10 (\*), p <= 0.05 (\*\*), p <= 0.01 (\*\*\*).  
Standard errors clustered at school level.

## Robustness checks: school-level analysis using a difference-in-difference specification (KS2 only)

The results so far aim to capture the difference in attainment levels between treatment and control schools based on their location (and the other background characteristics that we included in the original EEF model) and rich controls for school-level prior attainment.

However, if it is not possible to account fully for pre-existing differences in school attainment, the treatment schools might have different levels of attainment over and above the average effect for London and non-London schools. A difference-in-difference specification might therefore be appropriate since it would allow us to use data from previous periods to estimate and therefore account for these differences in attainment levels (unrelated to the treatment). The aim of a difference-in-difference specification is to net out any difference in the level of attainment between treatment and control groups that existed before the trial and isolate the impact of the trial. However, a critical assumption for this approach to be valid is that the trends in attainment before the trial are the same ('parallel trends') between treatment and control groups. Otherwise, differences in pre-period trends can bias the estimated treatment effect.

Based on the data we have available to us, we can implement a school-level difference-in-difference (DiD) using the KS2 results from previous academic years to control for schools' differences. Unfortunately, school-level KS1 data is not made publicly available, but we suggest this as an additional robustness check in the final section, subject to data access.

The first step for this specification is to plot the trends for attainment prior to the treatment year to assess whether they are parallel.

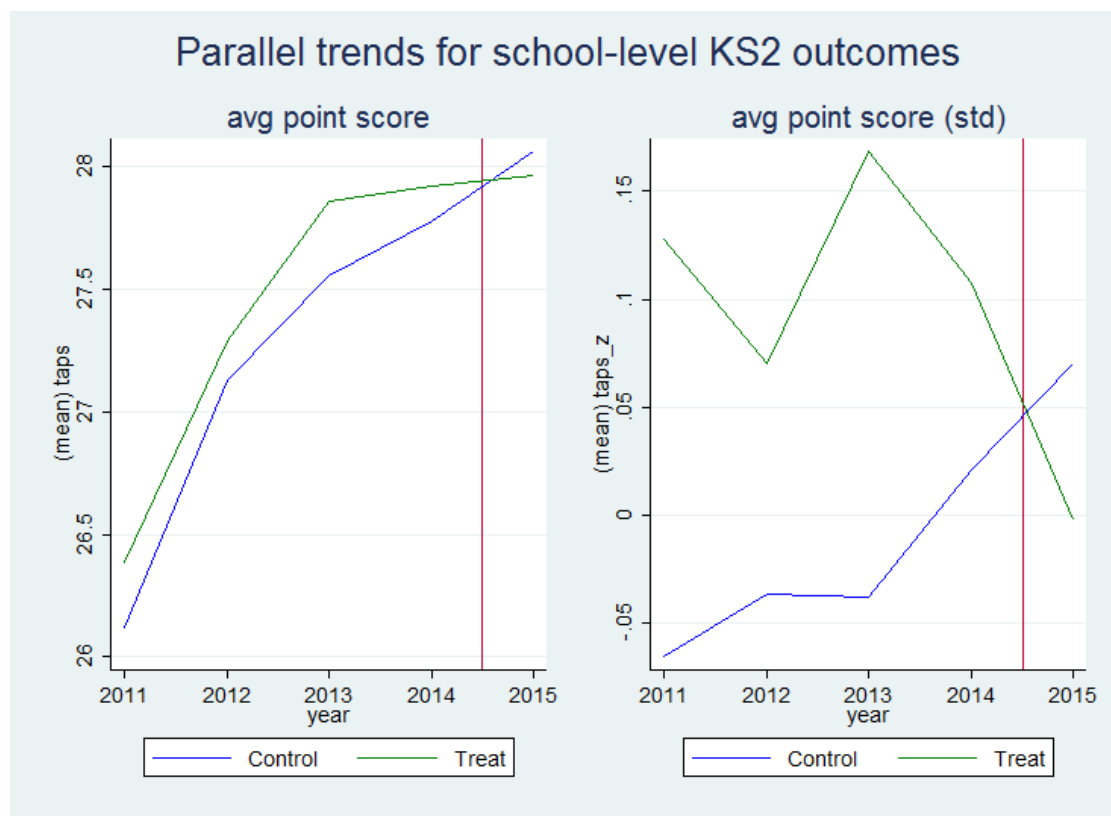
For Key Stage 2 we only have school-level data available for pre-trial cohorts. Our attainment measures are:

- school-level average point score (as constructed by DfE); and
- school-level average point score standardised within the trial sample to have a mean of zero and standard deviation of one.

The results in Figure 1 confirm that the treatment group has consistently higher attainment than the control group, even prior to the trial year (2015). In a pure randomised trial we would expect the pre-treatment trends to be (statistically) identical. The requirement of parallel trends is not met. For the average points score, the treatment group has a flatter increase in attainment between 2013 and 2014 than the control group. For the standardised average points score the difference in trends is more pronounced, with a decrease from 2013 for the treatment group. This means that the results from the difference-in-difference specification should be treated with caution as they are likely to be biased.



Figure Z1: Pre-trial trends for KS2 average points (full sample)



We estimate a standard difference-in-difference model that accounts for the difference in levels between treatment and control schools as well as any secular trend over time.

We run the following model:

$$y_{st} = \alpha + \tau_t + \beta Treatment_s + \delta Treatment * Post_{sT} + \pi Stratum_s + \gamma X_s + \varepsilon_{st}$$

Where  $y_{st}$  is an outcome;  $\tau_t$  is a set of dummies for each academic year;  $Treatment_s$  is an indicator for whether the school was assigned to the treatment group;  $Treatment * Post_{sT}$  interacts the treatment indicator with an indicator for whether the year is post-treatment (i.e. academic year 2014/2015 or later);  $Stratum_s$  is a set of randomisation strata dummies;  $X_s$  is a set of school-level controls (which vary between robustness checks); and  $\varepsilon_{st}$  is an error term, which we cluster by stratum.

The 'true' results come from correctly coding the post-treatment period  $Post_T$  as academic year 2014/2015 and beyond. But we can also derive a set of placebo results where we pretend that the treatment happened in earlier years (and so check whether we see 'effects' of the treatment even before it had happened). To do this, we drop the data from 2014/2015 and sequentially recode the  $Post_T$  indicator as being in 2013/2014, 2012/2013 onwards, and 2011/2012 onwards.

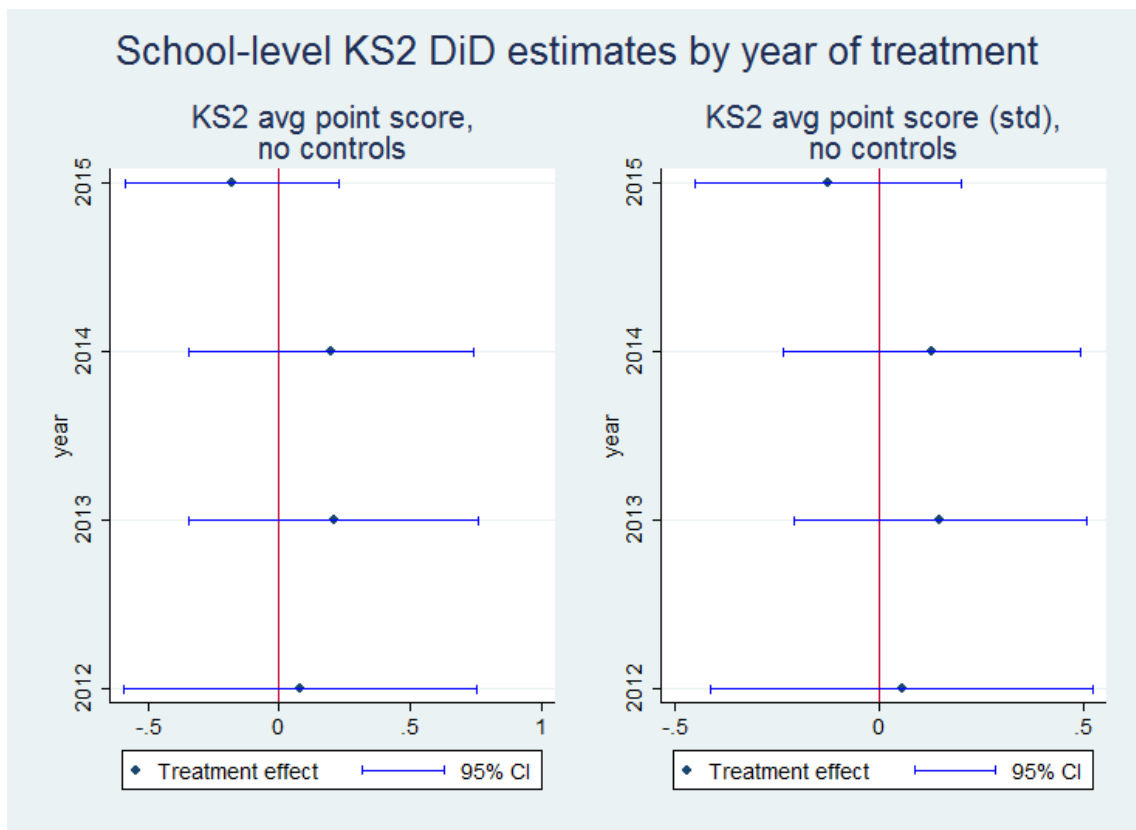
### Robustness checks: difference-in-difference with minimal controls

In this case,  $X_s$  contains binary variables for inner and outer London as well as for major urban areas only.

The results of both the main difference-in-difference and these placebo tests are summarised in Figure 2. In the top line (which corresponds to our actual treatment) we see negative but not statistically significant estimated effects.

The lower three lines on each graph give the results of the placebo checks. In each case, our placebo test does not find statistically significant effects.

**Figure Z2: Basic difference-in-difference results, Key Stage 2**



**Robustness checks: difference-in-difference with full controls**

The DiD results in the previous sub-section are quite far from the spirit of the specification we estimated for the EEF. In particular, they control for time-invariant differences between the treatment and control groups but do not account for differences in, for example, prior attainment, which might vary from cohort to cohort.

We can control for some of these time-varying, school-level characteristics. We have run two specifications:

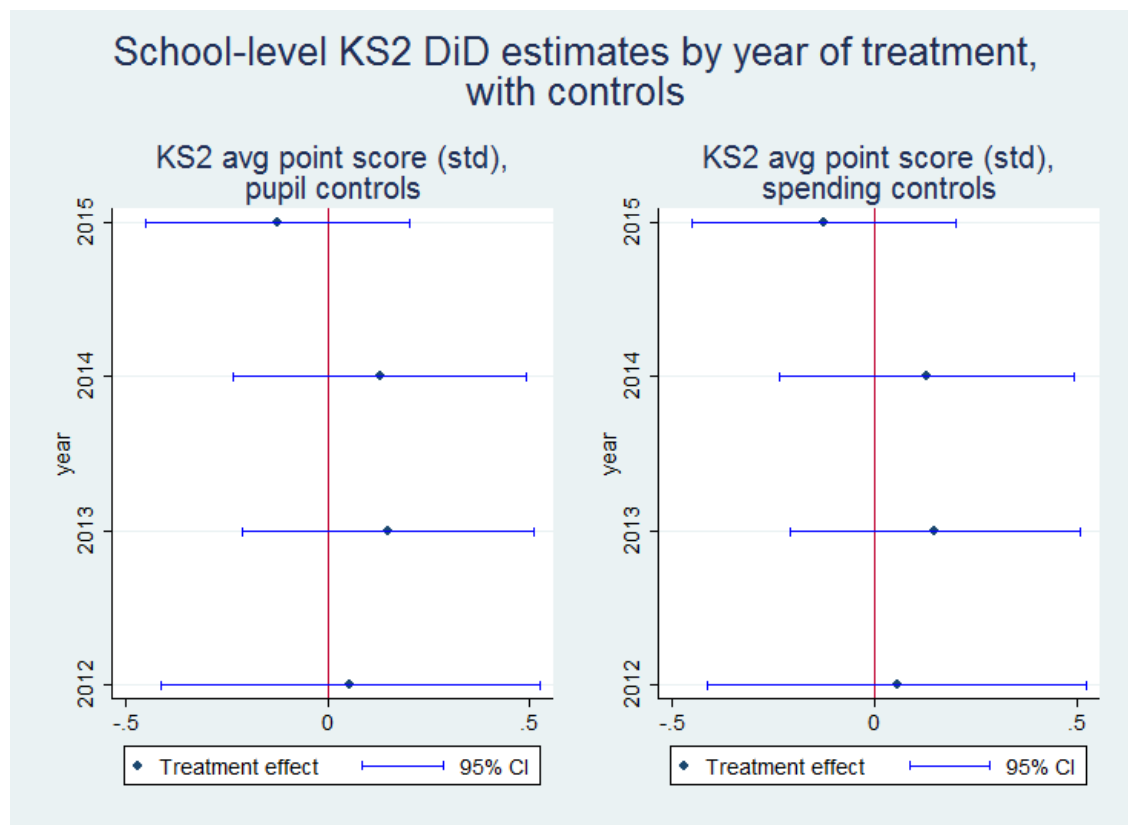
- controlling for (school-level) pupil characteristics: % on FSM, % with EAL, % with SEN, and prior attainment (FSP, standardised within sample)—this is analogous to the pupil-level controls in our original EEF report *except that we cannot control for ethnicity* (which is not available in the public-use school-level data); and
- controlling for school spending—total spending per pupil and the share of that spending on teaching staff, supply teachers, premises, and catering.

Figure 3 shows the results of these two specifications. Controlling for pupil characteristics does not majorly change the results for placebo years.

Since school spending information is not available before 2013, we can only estimate these specifications for ‘treatment’ in 2014 and 2015 (since we need at least one year of pre-treatment data). The addition of school-level characteristics as control variables does not affect the overall conclusion

that there is no statistically significant impact of the trial at KS2 according to the difference-in-difference specification.

**Figure Z3: Key Stage 2 DiD results with pupil and school spending controls**



### Robustness checks: DiD with time trends

The key assumption underlying difference-in-difference estimation is that the trends in the treatment and control groups would have been parallel if not for the intervention. The parallel trends graph (Figure 1) does not provide strong evidence for this, though the lines are quite noisy. This could reflect, for example, different trends in London schools or in the schools in different strata.

We can (partially) account for this econometrically by including linear time trends in our model. These allow different types of schools to be on different linear trajectories over time. We estimate these trajectories based on data from before 2015 (to avoid conflating the pre-existing trajectory with any effects from the treatment itself).

In practice, we include two types of trends:

- a linear time trend (yearly) for London schools; and
- linear time trends (yearly) for schools in the bottom 3 strata (missing data and infant schools); schools in stratum 4 (below median KS2 and EAL scores); schools in stratum 5 (below median KS2 scores, above median EAL); schools in stratum 6 (above median KS2, below median EAL); and schools in stratum 7 (above median KS2 and EAL).

We first estimate a modified version of our difference in difference, excluding 2015 data and the Treatment\*Post interaction term:

$$y_{st} = \alpha + \tau_t + \beta Treatment_s + \delta Treatment * Post_{sT} + \pi Stratum_s + \theta Trends_{st} + \gamma X_s + \varepsilon_{st}$$

Where  $Trends_{st}$  is a vector of the six trends described above and takes either value 0 (if the school is not in the relevant category) or value 1 in 2011, 2 in 2012, 3 in 2013, etc.

We then create six new estimated trend variables by multiplying the value of the trend with its coefficient and include these new trend variables—along with the Treatment\*Post interaction term—in our main specification.

At KS2, the trends make the estimated effects slightly more positive, but the point estimates remain negative and not statistically significant (Figure 4 and Figure 5).

**Figure Z4: KS2 DiD estimates for specifications with linear time trends**

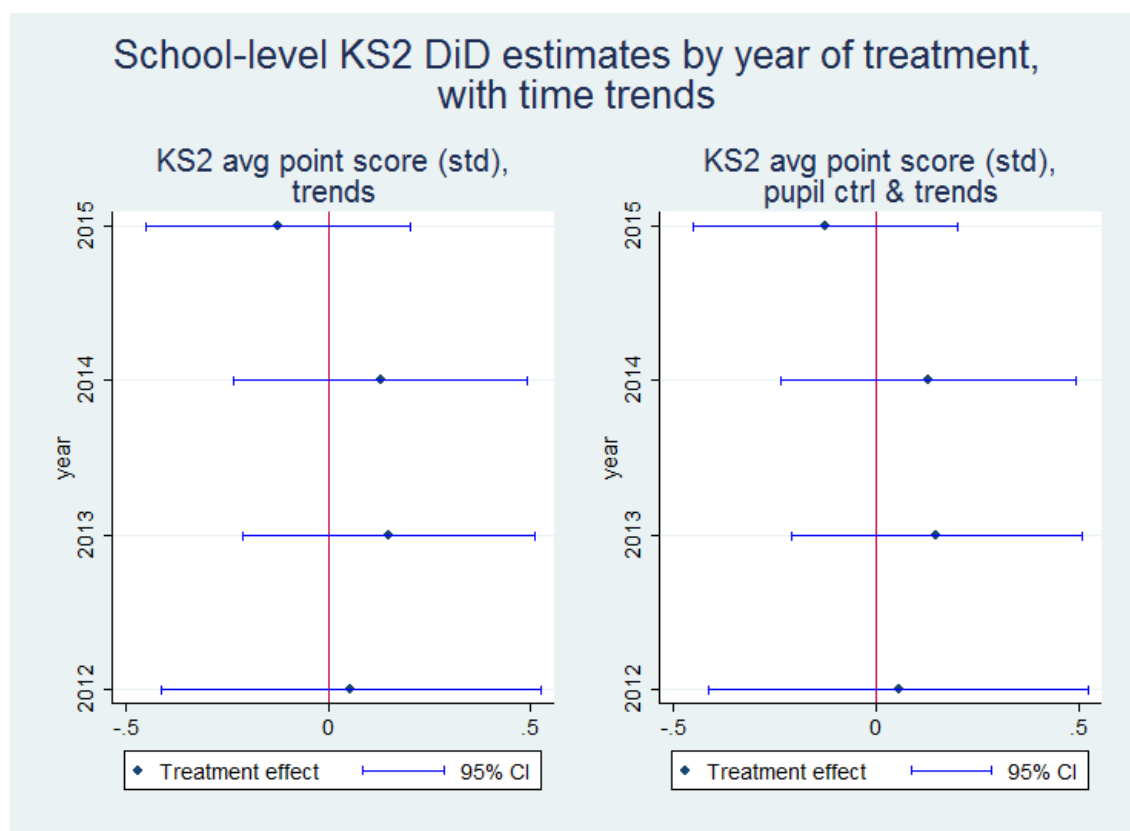
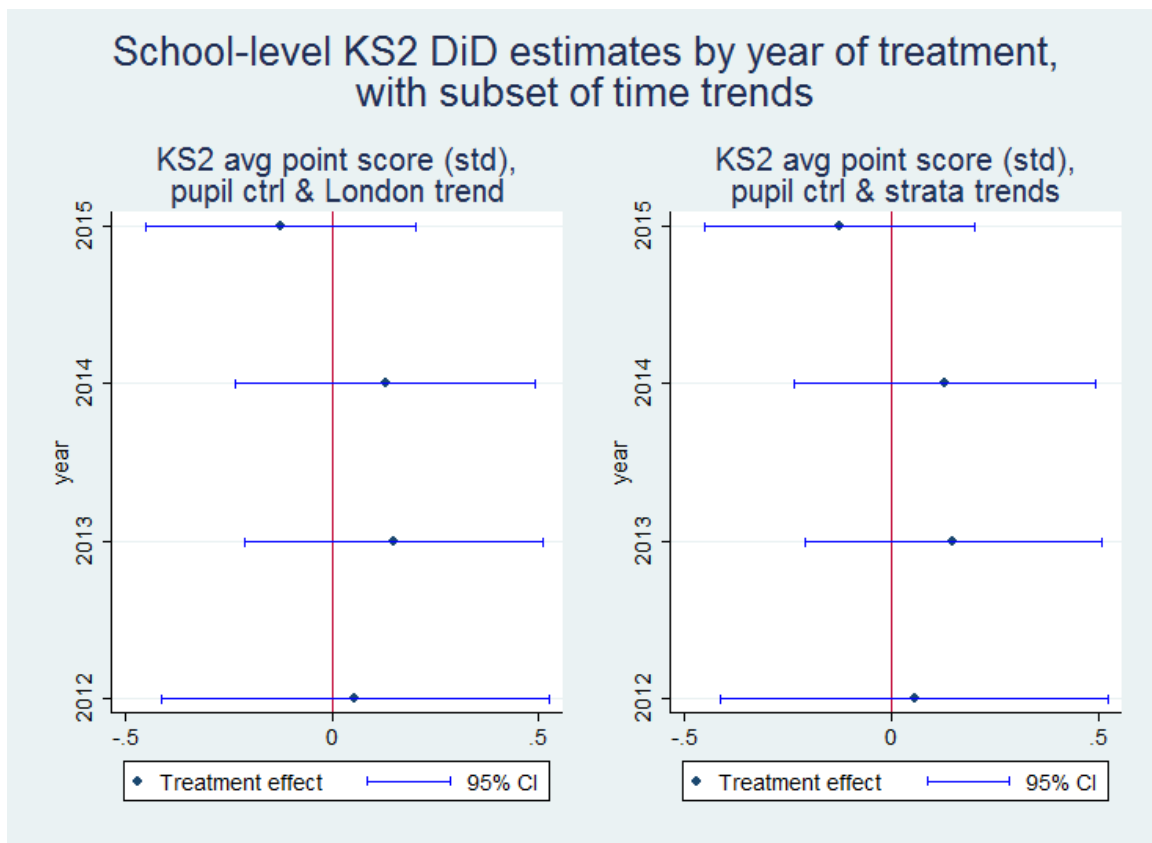


Figure Z5: KS2 DiD estimates for specification with linear time trends (subsets of trends)



### Robustness checks: DiD on common sample group

As discussed in the pupil-level robustness results, one remaining concern with these DiD results is the lack of common sample between treatment and control group schools. We can therefore repeat the DiD analysis on the two 'common sample' subgroups identified above: the schools in strata 1 to 4 in non-urban areas, and the London-based schools in stratum 7.

Figure 6 shows the results of a DiD analysis for the sample in schools in Strata 1 to 4, controlling for school-level pupil characteristics (EAL, FSM, SEN, and prior attainment) in order to be most comparable with our EEF results. We also show the results with and without the estimated pre-trends discussed above. Again, at KS2 we find effects that are very close to zero.

Figure 7 shows the same specifications on the stratum 7 London subsample. At KS2 the results are negative but not statistically significant.

Finally, Figure 8 repeats this analysis combining the two sets of common sample observations. This seems to combine the positive effect sizes of the non-urban results with the precision of the London results, which again is close to zero and not significantly different from zero.

Figure Z6: KS2 DiD estimates for common sample (strata 1–4, non-urban)

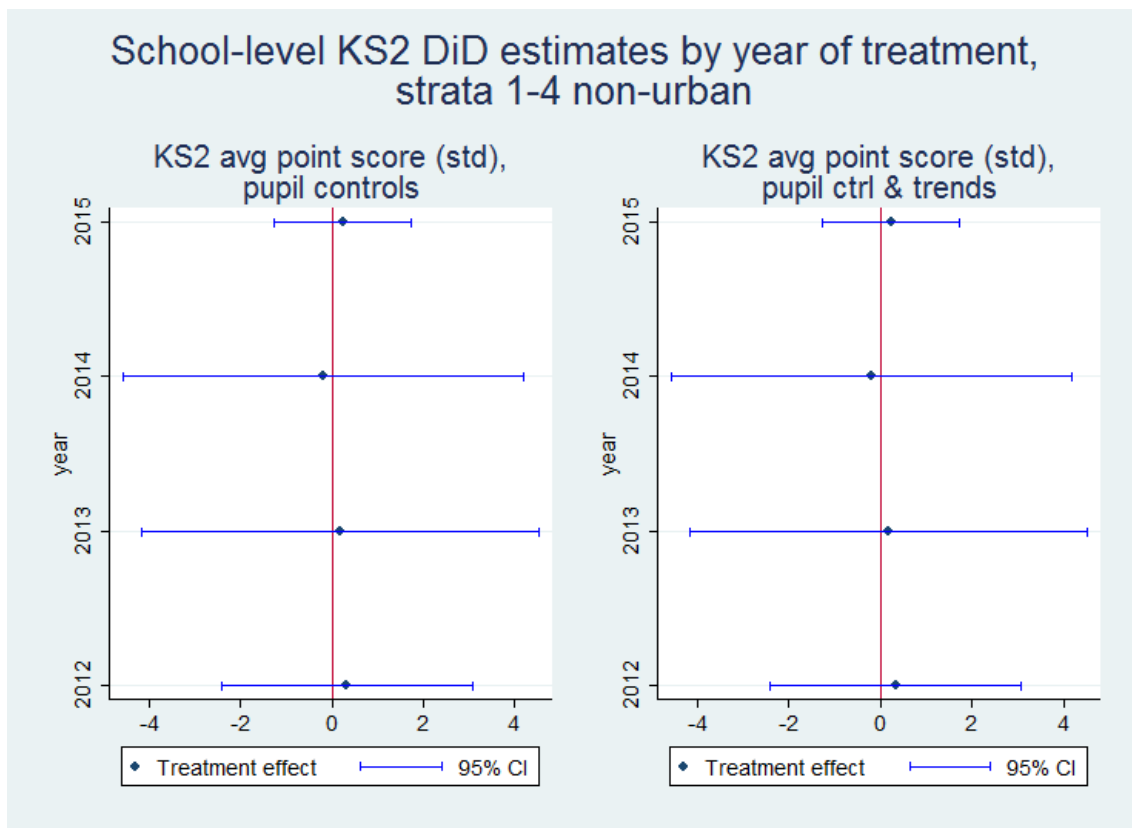


Figure Z7: KS2 DiD estimates for common sample (stratum 7, London)

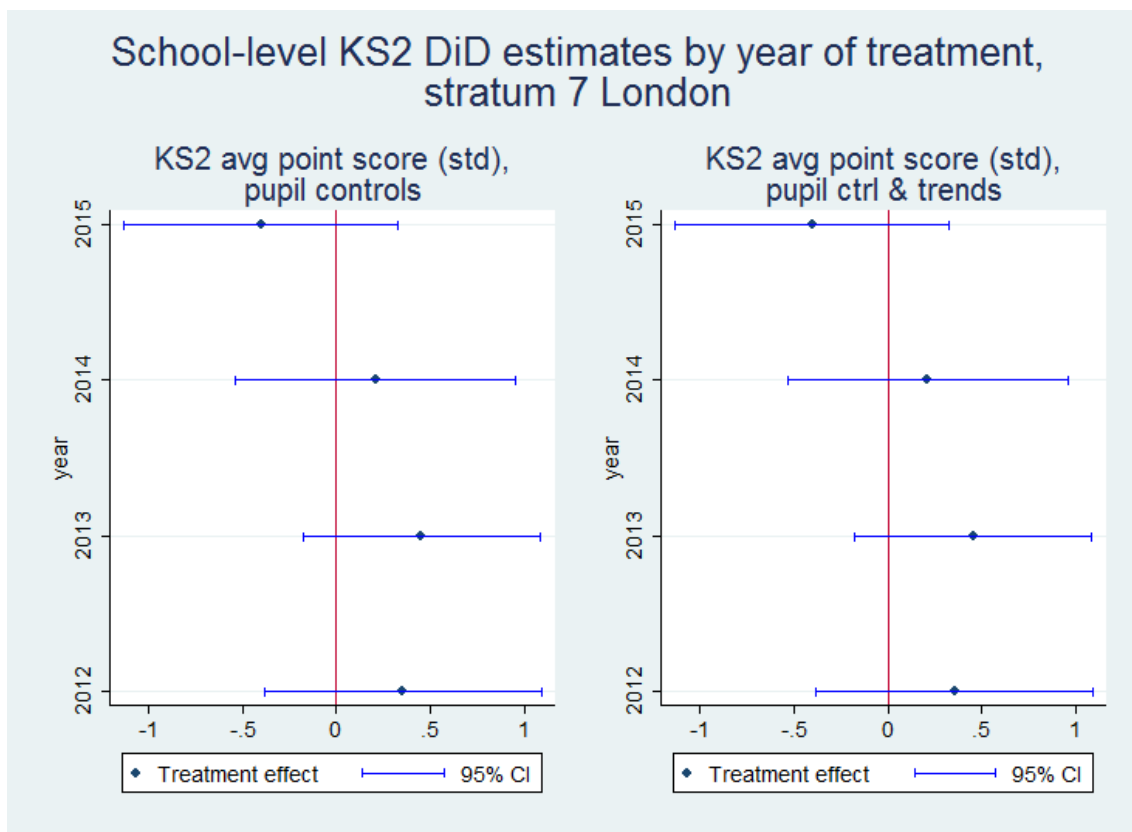
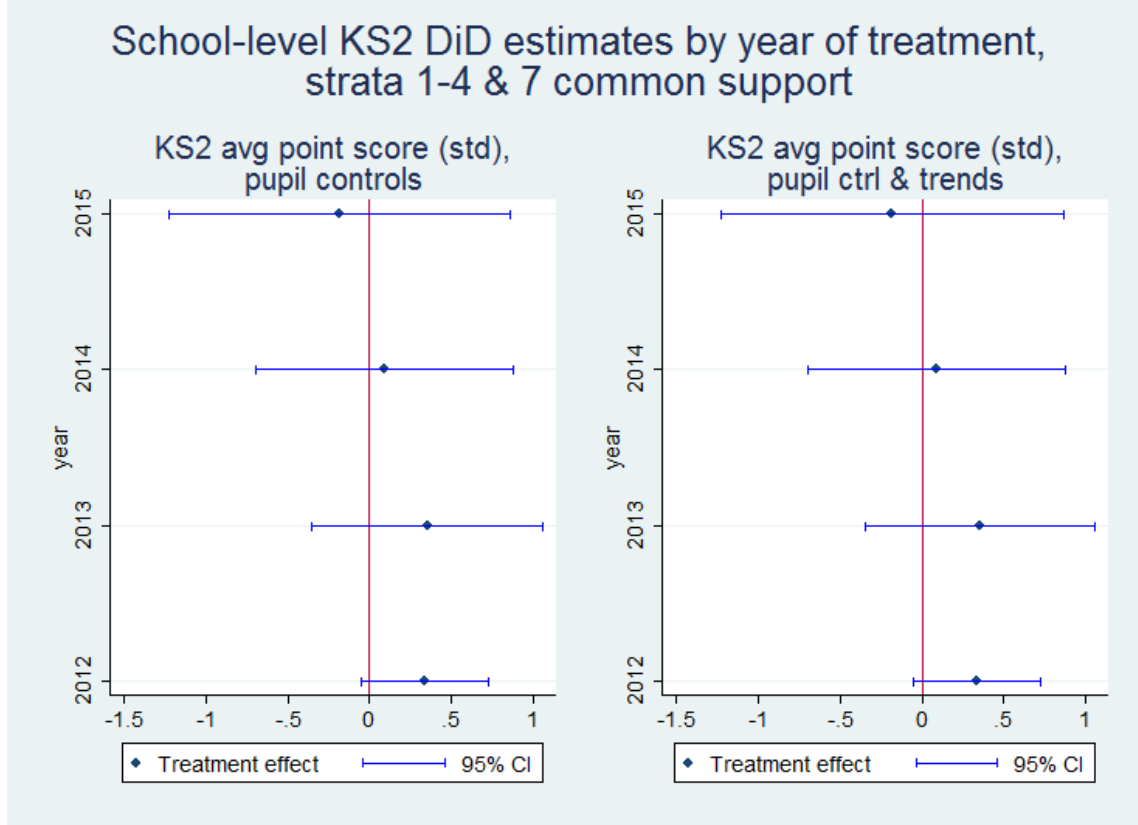


Figure Z8: KS2 DiD estimates for common sample (strata 1–4 non-urban, stratum 7 London)



## ORIGINAL REPORT AS PUBLISHED NOVEMBER 2016

### Executive summary

#### The project

The Magic Breakfast project provided 106 schools with support and resources to offer a free, universal, before-school breakfast club, including to all Year 2 and Year 6 pupils. The aim of the project was to improve attainment outcomes by increasing the number of children who ate a healthy breakfast. The schools in the project were schools in England with a relatively high proportion of disadvantaged pupils. The project ran between September 2014 and July 2015. Schools were provided with free food, support from a Magic Breakfast school change leader, and a £300 grant towards up-front costs. The intervention itself was delivered by school staff and volunteers.

The impact of the project was evaluated using a randomised controlled trial involving around 8,600 pupils. The process evaluation involved qualitative research with four case study schools. The project was jointly funded by the Department for Education and the Education Endowment Foundation and delivered by the charity Magic Breakfast.

#### Key conclusions

1. Year 2 children in breakfast club schools experienced around two months' additional progress compared to Year 2 children in the other schools in the trial. These positive results would be unlikely to occur by chance.
2. For Year 6 children in breakfast club schools, results for the main outcomes, reading and maths, were positive but could have occurred by chance. However, on other measures of writing and English they experienced around two months' progress compared to the other Year 6 children. These positive results would be unlikely to occur by chance.
3. The findings suggest that it is not just eating breakfast that delivers improvements, but attending a breakfast club. This could be due to the content of the breakfast itself, or to other social or educational benefits of the club.
4. Pupil behaviour, as measured by a teacher survey, improved in breakfast club schools. This is interesting because it shows that breakfast clubs may improve outcomes for children who do not even attend breakfast club, by improving classroom environments.
5. Activities thought to increase take-up of the breakfast provision included promoting it to parents and encouraging all children to attend while sensitively targeting pupils most likely to benefit. The project required additional staff time which some schools found difficult to provide without charging for breakfast.

#### How secure are the findings?

**Security rating awarded as part of the EEF peer review process**

The findings above have moderate to high security. The project was evaluated using a randomised controlled trial that compared the progress of pupils in the breakfast club schools to that of a control group of pupils receiving 'business as usual'. Randomisation was done by the independent evaluator. The trial was large and well-designed, and the number of pupils whose outcomes could not be measured at the end of the study, due to moving schools for example, was relatively low. The trial was an effectiveness trial, aiming to test the intervention under realistic conditions in a large number of schools.

However, around 40% of control schools established some form of breakfast club provision. While in some cases this was very limited, it is likely that some pupils in control schools were benefitting from



universal free before-school breakfast clubs which could result in an underestimation of the overall effect of the intervention.

### What are the findings?

The provision of a breakfast club led to an improvement in Key Stage 1 (KS1) outcomes of around two months' progress, roughly equal to the effect of providing universal free school meals in two pilot areas in 2011 (Brown *et al.*, 2012) that led to the roll-out of that programme in infant school. For KS2 assessments in reading and maths the impact was positive but slightly smaller, and may have been due to chance, so that it is not possible to say with confidence that it was due to the intervention. However, on measures of writing and English, KS2 students in breakfast club schools experienced the equivalent of around two months' progress compared to Year 6 children in the other schools in the trial. These positive results would be unlikely to occur by chance. The provision of a breakfast club led to larger improvements at KS1 and KS2 for pupils not eligible for free school meals (FSM) than for those eligible, although the effects were positive for both groups. This is despite FSM pupils being slightly more likely to eat breakfast at school as a result of the breakfast club. This either suggests that breakfast club attendance affects outcomes between groups differently, or that there was an indirect effect of the intervention on children who did not actually attend the breakfast club—perhaps because of an improved classroom environment—which was stronger for pupils not eligible for FSM.

Teacher perceptions of classroom behaviour and concentration indicate an improvement in the breakfast club schools relative to the other schools in the trial. These improvements mean that breakfast club provision can have benefits even for children who do not attend by improving their classroom learning environment. These spillovers between pupils could also play a role in explaining the stronger impact of the intervention for non-FSM pupils. Improved behaviour and concentration appears to have fed through to higher attainment, although to a greater extent at KS1 than KS2. Attendance at school also improved for children in breakfast club schools, resulting in about 26 fewer half-days of absence per year for a class of 30. There was no evidence of an impact on the body mass index of Year 6 students. (It was not possible to estimate the effect on Year 2 students' body mass index as this year group is not part of the National Child Measurement Programme.)

Key factors for successful implementation of the breakfast clubs were: communication with parents to encourage take-up, an established school breakfast routine, and a well-functioning delivery team supported by the wider school. The main challenges were compensating staff for additional hours of work and balancing the supply of, and demand for, food. In some schools, barriers to take-up included earlier start times for pupils, breakfast charges, and a lack of ongoing promotion from the school.

### How much does it cost?

The cost per pupil per year over three years is £11.86, averaged across all pupils in the breakfast club schools. The total cost was, on average, £4,462.11 per school. In addition, schools used 820 person-hours per year to deliver the intervention. On average, this included 87 teacher hours, 449 teaching assistant hours, 164 support staff hours, and 100 volunteer hours over the year.

**Table 1: Summary of impact on primary outcomes and impact by Free School Meal status**

| Group & outcome     | No. of schools | Effect size (95% confidence interval) | Estimated months' progress | Security rating | EEF cost rating |
|---------------------|----------------|---------------------------------------|----------------------------|-----------------|-----------------|
| Year 2: KS1 maths   | 102            | 0.149 (0.051;0.248)                   | + 2                        | ★★★★            | £ £ £ £ £       |
| Year 2: KS1 reading | 102            | 0.104 (0.012;0.196)                   | + 2                        | ★★★★            | £ £ £ £ £       |
| Year 2: KS1 writing | 102            | 0.138 (0.038;0.239)                   | + 2                        | ★★★★            | £ £ £ £ £       |
| Year 6: KS2 reading | 98             | 0.103 (-0.056;0.262)                  | + 2                        | ★★★★            | £ £ £ £ £       |
| Year 6: KS2 maths   | 98             | 0.075 (-0.060;0.210)                  | + 1                        | ★★★★            | £ £ £ £ £       |
| KS1 score: FSM      | 101            | 0.153 (0.068;0.237)                   | + 2                        | n/a             | £ £ £ £ £       |
| KS1 score: non-FSM  | 102            | 0.246 (0.152;0.341)                   | + 3                        | n/a             | £ £ £ £ £       |

|                    |    |                      |     |     |           |
|--------------------|----|----------------------|-----|-----|-----------|
| KS2 score: FSM     | 98 | 0.037 (-0.048;0.122) | + 1 | n/a | £ £ £ £ £ |
| KS2 score: non-FSM | 98 | 0.269 (0.150;0.387)  | + 4 | n/a | £ £ £ £ £ |

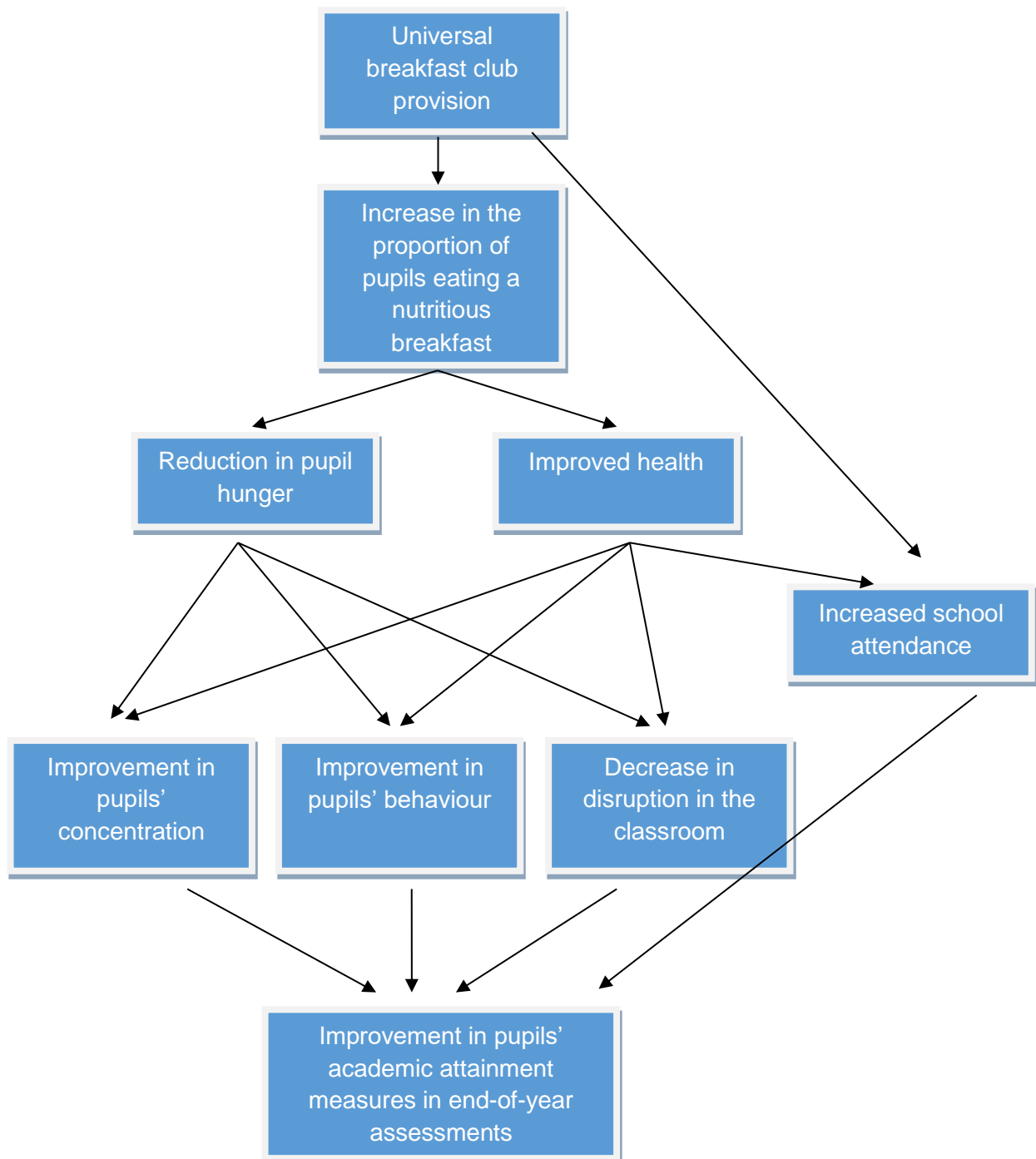
## Introduction

### Intervention

Magic Breakfast is a charity that supports the provision of breakfast clubs in disadvantaged schools in England, defined as those as having at least 35% of pupils eligible for free school meals or 50% of pupils having been eligible in the previous six years. This evaluation tested the impact of providing primary schools with resources and support to introduce a breakfast club which is before-school, universal, and free, over one academic year, from September 2014 to July 2015.

The rationale for this intervention is that addressing pupil hunger may lead to improvements in concentration and behaviour in the classroom, and ultimately improvements in pupil attainment. Teachers already involved with a breakfast club have also reported to Magic Breakfast that it improves the classroom environment for all pupils. This is because learning time increases as the interruptions due to hunger decrease, and there is less disruption from pupils arriving late. The mechanisms through which breakfast club provision might lead to improvements in pupil attainment are shown graphically in the Theory of Change model in Figure 1. The underlying logic is that the provision of a free universal breakfast club in relatively disadvantaged schools leads to an increase in the proportion of pupils that eat a nutritious breakfast; in turn, this will reduce pupil hunger (especially in the morning) and increase overall pupil health. An improvement in health and reduction in hunger is hypothesised to improve pupil concentration and behaviour, and increase the productivity of the learning environment. More indirectly it was suggested that the breakfast club may increase attendance at school (through improved pupil health and reduced sickness) and increase punctuality where the club is organised before the school day. Any increase in attendance or punctuality may have the potential to increase attainment as these pupils will have more hours of teaching and learning.

Figure 1: Theory of change model



The breakfast club in each school was supported by Magic Breakfast through the provision of free healthy food (as much as required), £300 in capital funding, and advice about establishing and successfully delivering a breakfast club (including the staffing arrangements for supervision of pupils during the breakfast club). Magic Breakfast's assistance included supporting schools to develop a sustainable breakfast club model that could continue after the intervention. The choice of breakfast club delivery for the intervention (universal, free, and before-school) was informed by theoretical reasoning and previous anecdotal evidence collected by Magic Breakfast. For example, previous evidence has suggested that universal rather than targeted provision is more likely to lead to higher take-up by disadvantaged pupils by changing norms within the school and reducing social stigma. Holding the breakfast club before school may also improve attendance and punctuality by creating an added incentive for children to arrive at school on time. In practice, however, schools in the intervention group implemented various forms of breakfast club; some introduced a small charge for some or all pupils, some capped the number of places available, and some introduced a breakfast club during school hours.

The implication of this deviation for the evaluation is that the effect of the intervention must be interpreted as the effect of providing schools with resources and support to introduce a universal, free, before-school breakfast club, rather than the effect of schools actually doing so. More detail on the fidelity to the programme (defined by the provision of a universal, free, before-school breakfast club) and take-up of the offer is given in the 'Outcomes and Analysis' section.

The control condition was 'business as usual', with a 'waitlist' design, where schools allocated to the control group were helped to establish their breakfast club the following academic year. As an incentive to minimise the attrition rate of control schools, Magic Breakfast supported breakfast clubs in these schools for two academic years rather than the one year of support given to intervention-group schools.

Around 40% of schools in the control group that responded to the follow-up headteacher survey adopted some form of breakfast club throughout the year. The majority of these schools (15 out of 17) did not charge for the breakfast provision, but it was targeted rather than universal in some cases (6 of 17). This level and type of provision is in line with the 'business-as-usual' design of the trial, but the relatively high presence of breakfast clubs in the control schools must be kept in mind when interpreting the evaluation findings. Around 90% of schools in the control group that responded to the follow-up survey also established or encouraged a breakfast club for Year 6 pupils in the week of testing for national exams: this may dilute any differences in attainment at KS2 between pupils in intervention and control schools.

The main motivation for schools joining the intervention, reported in the headteacher survey before randomisation and during case study visits, was a concern about pupils arriving at school hungry and the impact that this might have on students' wellbeing, behaviour, and experiences in the classroom. The introduction of Magic Breakfast provision was seen by staff as a way to tackle a number of pre-existing problems:

- to address problems of pupil hunger and, by virtue of this, improve children's wellbeing, concentration, and behaviour in class;
- to improve attendance and punctuality by creating an added incentive for children to arrive at school on time;
- to support parents by reducing the stress of morning routines and providing early morning childcare; and
- to establish and reinforce a welcoming and inclusive community ethos within the school.

## Background evidence

There has been significant interest from the Department for Education (DfE) in England in the relationship between nutrition and school attainment following the publication of the School Food Plan (2013) and the high-profile evaluation of the universal provision of free school meals in two pilot local authorities (Brown *et al.*, 2012). In response to the School Food Plan, the DfE committed to providing breakfast clubs in 183 schools where at least 35% of pupils are FSM-eligible with the aim of these clubs becoming financially self-sufficient after the intervention period. Magic Breakfast was awarded the contract to deliver free, healthy breakfasts and provide advice on establishing a sustainable breakfast club to meet this commitment, and work began around the same time as the intervention evaluated here.

School nutrition, in particular school breakfasts, are a focus for policy in Wales as well. The Primary School Free Breakfast Initiative (PSFBI) was piloted by the Welsh Assembly Government from September 2004 and subsequently made a duty for all maintained schools in 2013. Its aim is to encourage a healthy pattern of eating and improve concentration and standards of attainment for pupils. In Wales, local authorities must provide free 'healthy' breakfasts to all pupils in maintained schools that have requested them, with funding provided for food and staff costs.<sup>10</sup>

These policy initiatives are based on evidence that healthy breakfast consumption is linked to better outcomes for children. Much of this evidence is from the U.S. where breakfast provision in disadvantaged areas has a long history. For example, the School Breakfast Programme was established in 1966 (Shaw, 1988). Evidence has suggested that children eating breakfast regularly have better health and academic outcomes (Pollitt and Matthews, 1998; Rampersaud *et al.*, 2005), while 9- to 11-year-old children skipping breakfast are less able to distinguish similar images, show increased errors, and display slower memory recall (Pollitt *et al.*, 1998).

Relevant to this study, some research finds that breakfast *at school* helps children perform better. For example, Mahoney *et al.* (2005) conducted a three-week intervention in American schools and found improved cognitive performance in 9- to 11-year-olds. These benefits were also evident among 6- to 8-year-olds, where they also differed by type of food consumed. An evaluation of Minnesota's three-year universal school breakfast pilot found that students exposed to the intervention tended to improve their percentile rank on standardised tests written at age 11, relative to their rank three years earlier (Minnesota Department of Children, Families, and Learning, 1998). Massachusetts' expansion of its School Breakfast Program led to improved test scores and lower rates of absence and tardiness among low-income elementary school children aged 8 to 11 (Meyers *et al.*, 1989).

Many of these studies typically suffer from methodological issues, however, such as inadequate or absent control groups (Grantham-McGregor, 2005). There are some exceptions that use a randomised controlled trial design. For example, Murphy *et al.* (2011) used an experimental design to evaluate the effects of free school breakfasts in Wales, finding positive effects for pupil diet (more 'healthy' breakfast items consumed) but no effect on memory or behaviour. In this trial, breakfast consumption moved from home to school, and there was no effect on food consumption later in the day. Shemilt *et al.* (2004) also use a randomised controlled trial to test the impact of an early Department for Health breakfast club initiative in England, finding lower absences and higher levels of pupil concentration among schools randomly assigned to breakfast club provision, although also some evidence of worse conduct. The evaluation focused primarily on student health and behaviours, such as emotions and relationships—

<sup>10</sup> A 'healthy' breakfast is defined in the Healthy Eating Regulations, regulation 4, Schedule 1. Food categories are: milk-based drinks or yoghurts, cereals—not coated or flavoured, either alone or in combination with sugar or chocolate or cocoa powder, fruit and vegetables, breads, and toppings (thinly spread).

measured through the Strengths and Difficulties Questionnaire (Goodman, 1997)—and their concentration—assessed through the Trail Making Test (AITB, 1944). Information on absences and punctuality was only available at the school level, and the authors did not analyse any attainment data. The evaluation was also challenged by a relatively short time between baseline and first follow-up (three months) and substantial contamination between study arms before the second follow-up at one year. In a developing country context and using a within-school randomised study design, Powell *et al.* (1998) found that provision of a school breakfast for pupils in grades two to five (age six to nine) in Jamaica improved nutritional status, school attendance, and achievement in arithmetic (mainly for younger pupils in grades two and three).

The rationale for conducting this new evaluation is to provide evidence from a randomised controlled trial on the impact of breakfast club provision on a range of pupil outcomes, including attainment, focusing on relatively disadvantaged schools in England where pupil hunger was thought to depress pupil learning. The original rationale of the trial was also to test the effectiveness of alternative methods of provision to inform the expansion of breakfast clubs (see the Trial Design section for more detail on why only one model was eventually tested). The trial was an effectiveness trial as it was implemented to a relatively large number of schools in a ‘real world’ setting. This means that the effect of the intervention is likely to represent the effect if implemented by other schools in a similar context. The outcomes are chosen to show the direct effect of breakfast club provision on academic attainment and the mechanisms through which any effect occurs. The findings from this evaluation are particularly relevant for England and the U.K., but of interest internationally given the limited existing evidence from randomised controlled trials.

## Evaluation objectives

The primary objective of the impact evaluation was to measure the impact of breakfast club provision on pupils’ academic attainment, by comparing pupil outcomes in the intervention group with a ‘business as usual’ control group. The research questions for the impact and process evaluations, as defined in the evaluation protocol, are:

- What is the impact of free universal breakfast club delivery before the school day on teacher-assessed levels of KS1 attainment in English (reading, writing, and speaking and listening) and maths?
- What is the impact of breakfast club delivery on test scores and teacher-assessed levels of KS2 attainment in English and maths?
- What mechanisms are likely to explain any improvement in academic attainment?
  - Breakfast consumption.
  - Classroom behaviour and concentration.
  - Attendance at school.
  - Health (proxied by measures of underweight, normal weight, and overweight).
- Does the impact of breakfast club delivery vary across groups of pupils and different types of schools?
  - Pupils eligible and not eligible for free school meals.
  - Pupils with low prior attainment.
  - Pupils that have and do not have breakfast.
  - Boys and girls.
- What is the cost-effectiveness of free universal breakfast club delivery before the school day?
- How attractive is the free breakfast provision model to schools?
- What does planning, design and set-up involve?
- What approaches of delivery have schools developed? What are the challenges and success factors?
- How well has delivery worked in practice?
- What do schools, children and families perceive the impact of free breakfasts to be?

- What are the schools' plans or expectations for future provision of school breakfasts when support from Magic Breakfast ends?
- How do all the above factors vary for different types of schools in different contexts?

The motivation for each research question is to test the assumptions in the theory of change illustrated in Figure 1, and therefore inform which of the mechanisms relating to the provision of a universal and free breakfast club are likely to lead to improved academic attainment. The rationale for analysing each of the subgroups stated is that each group could reasonably be expected to respond differently to the intervention. FSM pupils may be more affected by the intervention than those not eligible if their breakfast consumption is more likely to increase or their breakfast composition more likely to become more nutritious during the intervention. Pupils with low prior attainment may be more affected by the intervention than those with expected or high prior attainment if their performance and concentration in the classroom improves more during the intervention. This would be expected if these pupils are more easily distracted by poor behaviour from their peers, or their own feelings of hunger. Pupils that do not have breakfast before the intervention may be more affected by breakfast club provision than those that do, as their breakfast consumption, and therefore concentration and behaviour in classroom, has the most potential to improve. Finally, boys and girls may be affected by the intervention differently as social norms for attending the breakfast club and concentration and behaviour in class may be different.

The evaluation protocol can be read here:

[https://educationendowmentfoundation.org.uk/public/files/Projects/EEF\\_Project\\_Protocol\\_MagicBreakfast.pdf](https://educationendowmentfoundation.org.uk/public/files/Projects/EEF_Project_Protocol_MagicBreakfast.pdf)

### **Ethical review**

Ethical approval for this study was granted by the UCL Research Ethics Committee (project 5749/001). Opt-out consent for data collection and sharing was sought from all parents in intervention and control schools.

### **Project team**

The intervention was developed and supported by the charity Magic Breakfast, led by Alex Cunningham with Fatima Khan and Maureen Healy. The data collection for the evaluation was assisted by many members of staff at Magic Breakfast, and was greatly assisted by those with responsibility for contact with each school.

Ellen Greaves led the impact evaluation, working with Christine Farquharson and Claire Crawford, and with support from Laura Westwood. The process evaluation was carried out by a team from the National Children's Bureau (NCB) Research Centre: Emma Wallace and Clarissa White oversaw the design and delivery of the process evaluation and edited the report. Grace Trevelyan and Amy Edwards carried out the case studies, analysis and reporting; Jo Lea was involved in the early stages of the research and helped to design the topic guides used to explore breakfasts with young people.

### **Trial registration**

This trial has been registered on the international standard randomised controlled trial number (ISRCTN) registry at <http://www.isrctn.com/ISRCTN14385822>.

## Methods

### Trial design

The trial used a randomised controlled design, with randomisation at the school (or cluster) level. Randomisation at the school level rather than pupil level was chosen as spillovers between pupils were likely. For example, access to breakfast could improve the concentration of one pupil, which then affects his/her classmates through less disruption in the classroom.

Schools were randomly allocated to one of two groups; the intervention group, who received support and resources to establish a universal free breakfast club before school in the academic year 2014/2015, and the control group, who received support and resources for the two following academic years (2015/2016 and 2016/2017). This 'waitlist' design was chosen to minimise attrition from the trial and maximise the survey response rate.

It was originally intended that there would be three intervention groups offering three separate models of school breakfast provision for relatively disadvantaged schools:

1. a free universal breakfast offer within the existing school morning—breakfast is offered free of charge to all pupils at some point between registration and the end of the morning break;
2. a free universal offer before the existing school day—breakfast is offered free of charge to all pupils before registration, with the exact time decided by each school; and
3. a breakfast club before the school day, with a charge, but offering free breakfast to FSM pupils—breakfast is offered free to FSM pupils (with a charge for other pupils) before registration, with the exact time and charge to be decided by each school.

For each breakfast model the schools would order their choice of cereals, wheat biscuits, porridge, bagels, and juice from Magic Breakfast, which would provide as much of these foods as requested free of cost (with the exception of juice, the amount of which was capped). Schools would also be able to supplement the breakfast foods offered to their pupils at their own cost.

The intervention changed because the number of schools eligible for the evaluation (that is, schools with limited or no existing breakfast provision—defined as no provision at all, or provision usually attended by 6% of pupils or less) was lower than expected. Of the 1,765 primary, infant, and junior schools that were approached that met the criteria for pupil disadvantage (at least 35% of pupils currently eligible for free school meals), only 374 (21%) met this existing provision eligibility criterion (compared to the initial expectations that around half of schools eligible in terms of pupil disadvantage would also meet this criterion). Because of the lower-than-expected number of eligible schools, rather than continue with the three variations listed above, the intervention included only the most popular and prevalent model of school breakfast provision: a free universal offer before the school day starts (model 2 in the list above). This means that all pupils in the intervention schools are offered a free breakfast (in contrast to model 3) and that the breakfast club takes place before registration (in contrast to model 1).

The choice of model was informed by previous anecdotal evidence from the project team and theoretical reasoning that this would be most appropriate for reaching disadvantaged pupils. For example, it was thought that take-up by disadvantaged pupils might be increased most under a universal provision model that could change norms within the school and reduce social stigma. In support of this reasoning, the evaluation of the free school meal pilot commissioned by the Department for Education found that extending the eligibility criteria for free school meals in secondary schools (rather than making it universal) had no impact on the attainment of eligible pupils, and that take-up was low amongst the eligible group (Brown *et al.*, 2012). Universal provision may also have a positive impact on pupil attainment if there are complementarities between pupils (positive 'spillovers' between pupils in the classroom).



## Participant selection

All pupils in selected schools were eligible for the trial as the breakfast club offer was universal and free. Schools were eligible for the trial if at least 35% of pupils were currently FSM-eligible and they had no existing breakfast club provision or existing provision usually attended by 6% of pupils or less. This meant that 374 of the 1,765 schools that were initially approached were eligible. In addition, schools must have agreed to the conditions of the project and evaluation by signing the Memorandum of Understanding shown in Appendix C.

Schools were recruited by the Magic Breakfast project team. This process involved: identifying schools that met the pupil disadvantage criterion (through the latest available census information in the School Performance Tables); contacting these schools to see whether the breakfast club criterion was also met; raising awareness of Magic Breakfast among schools in general (for example through national media advertisements and information circulated by existing partner schools); targeted approaches through phone calls, tailored emails, and letters; and finally, recruitment meetings with potentially interested schools. The purpose of this recruitment meeting was to provide the opportunity to discuss in more detail how the partnership between the school and Magic Breakfast would work, and discuss the challenges faced by the school, such as the perceived impact of hunger.

Parent/guardian opt-out consent for participation in the research analysis and linking between pupil survey and administrative data in the National Pupil Database (NPD) was sought after randomisation to intervention or control group.

Consent forms are available in Appendix C.

## Data collection

The data collected for this evaluation included administrative data (for pupils of parents/guardians that granted opt-out consent) and online and paper surveys of teachers, headteachers, and pupils. These are described in turn below, and survey questions are given in full in Appendix D. Survey response and the level of missing data is shown in Figure 2.

- National Pupil Database: The results of both teacher- and externally-marked national assessments for all pupils in state-funded schools in England are held in the NPD, which is maintained by the Department for Education. These outcome measures were collected from the NPD to reduce the data collection burden on schools and to ensure reliability and representativeness. The NPD provides the primary and secondary outcome measures related to academic attainment and data relating to pupil absence and punctuality.
- National Child Measurement Programme (NCMP): Information on child health was collected from the NCMP, which records the height and weight of all Year 6 pupils (aged 10–11) in state-funded schools in England. This data was provided at the school level for schools in the intervention and control groups with Year 6 pupils, and shows the percentage of Year 6 pupils with a 'healthy weight' and the average Body Mass Index (BMI) of the year group for each school. The data was supplied under a Data Sharing Agreement which did not allow the data to be further released in a way in which an individual school could be identified.
- Baseline pupil survey: The baseline pupil survey was completed by pupils in Year 2 and Year 6 in intervention and control schools, around September–October 2014. Pupils were asked, among other things, 'Did you have breakfast today?', 'Did you have breakfast at school today?', and 'Did you feel hungry when you started your first lesson, or not?'. They were also asked to describe or draw the foods they had eaten that morning. The surveys were completed on paper and sent to Magic Breakfast to be inputted manually before being securely uploaded to DfE for linkage to the NPD.
- Follow-up pupil survey: The questions and format of the baseline pupil survey were repeated in the follow-up pupil survey, completed by Year 2 and Year 6 pupils in June–July 2015.

- **Baseline teacher survey:** The baseline teacher survey was completed by teachers of Year 2 and Year 6 classes in intervention and control schools around September–October 2014. The survey measured teachers' perceptions of the typical level of pupil concentration and behaviour in their class at the time of response with questions such as: 'Thinking about the first lesson your class had today, what percentage of children in your class do you think had a poor level of concentration?' and 'Please give the percentage of your class that are usually ready to learn at the start of the first lesson of the day'. The surveys were completed online using Google Forms.
- **Follow-up teacher survey:** The questions and format of the baseline teacher survey were repeated in the follow-up teacher survey completed by teachers of Year 2 and Year 6 teachers in June–July 2015.
- **Baseline headteacher survey:** The baseline headteacher survey was completed by headteachers in intervention and control schools as a condition of being told the allocation to intervention or control group for the trial. The baseline survey was typically completed in June to August 2014. The baseline survey asked about the motivations for signing up to the Magic Breakfast intervention, the number of Year 2 and Year 6 classes at the school, and the contact email addresses for teachers of these classes. The surveys were completed online using Google Forms.
- **Follow-up headteacher survey:** The follow-up headteacher survey was also completed online using Google Forms, in June–July 2015. For headteachers of control schools, the questions included information about whether a breakfast club had been established, and the format of the club if so. For headteachers of intervention schools, the questions also included information about the format and coverage of the breakfast club that was implemented, and detailed questions on the cost of running the breakfast club and the additional staff time it required. Headteachers in intervention schools were also asked whether the breakfast club would continue after the trial and associated support from Magic Breakfast ended. Headteachers in intervention and control schools were also asked whether students in Year 6 were encouraged to take up a breakfast club offer in the week of testing for KS2, and whether this was the first year this had been done.
- **Food orders data:** Schools placed orders with Magic Breakfast, which delivered the requested food free of charge each week. Magic Breakfast provided us with the order sheets for intervention schools during two representative weeks in the spring term (weeks of 2nd and 9th March) as well as information on the termly porridge orders for the spring term. Magic Breakfast also provided information on the regular retail prices of each item which we have used to estimate the cost of food for a breakfast club in a school that is not receiving support from a charity like Magic Breakfast.

## Outcome measures

The primary outcome measures are national assessments in English and maths taken by all pupils in state-funded schools in England. These assessments are held at the end of each Key Stage in primary schools, when children are Year 2 (aged 6–7, KS1) and Year 6 (aged 10–11, KS2). These assessments are done exclusively by the teacher at KS1 (known as 'teacher assessments') and by both teacher assessments and externally-marked tests at KS2. We use the teacher assessments in reading, writing, and maths at KS1, and the externally-marked tests in English and maths at KS2, as our primary outcomes (see below for a more detailed discussion of why we chose the externally-marked tests as primary outcomes for KS2). In particular, we use the point scores obtained in KS1 and the fine point scores awarded by external markers on the KS2 tests. We standardise each academic outcome within the sample to have a mean of zero and a standard deviation of one. This means that the coefficients derived from the analysis are expressed in terms of standard deviations and are comparable across the academic, behaviour, and concentration outcomes within this evaluation. It should be noted that these

standard deviations are related to the sample presented here and are not comparable with standard deviations derived from other samples.

We also identified a set of secondary academic outcomes. For these we use additional teacher-assessed outcomes at KS1 (speaking and listening) and teacher-assessed outcomes at KS2: reading, writing, English, maths, and science.

For all teacher-assessed outcomes, teachers assess students against levels of achievement, which we convert to numeric points using the DfE's 'KS1 and KS2 test and examination point scores used in the 2015 school performance tables' guidance, then standardise as discussed above (DfE, 2015a).

For our primary outcomes we used externally-marked tests rather than teacher-assessed measures, where available. This is because teacher assessments have the potential to be affected by teachers' knowledge of the intervention or to conflate improvements in mechanisms (such as pupil behaviour) with improvements in attainment. They are also less finely-scored than the externally-marked tests. For these reasons, they are considered less robust outcome measures than externally-marked tests. However, the teacher-assessed measures we use are moderated and used for national performance data and therefore still provide sufficiently reliable outcome measures. For some subjects our data allow us to compare the impact of the intervention on both teacher assessments and test scores (for example, for KS2 maths and reading). In these cases results are similar which suggests that any bias introduced by teacher assessment is minimal.

In addition to supplementary attainment measures, other secondary outcomes measures were chosen to show the mechanisms through which any impact on attainment occurs. First, breakfast consumption and student hunger were measured through the pupil surveys at baseline and follow-up. Second, classroom teacher surveys at baseline and follow-up measured teachers' perceptions of student behaviour and concentration. Finally, the presence of a breakfast club could improve attendance and punctuality at school which could in turn have an impact on academic attainment. Attendance and punctuality could improve because of the additional incentive to arrive at school on time, or because pupils' health improves. These mechanisms are measured through two sources of data. The NPD contains information on pupil-level absences, and whether these were authorised (for example because of illness) or unauthorised. Information on child health, proxied by healthy weight, was collected from the NCMP.

## Sample size

The target for recruitment of schools was 50 schools per treatment arm. This was decided on the basis of sample size calculations, shown in Table 3. As described above, the number of treatment arms was reduced from three to one when recruitment to the trial was lower than expected. A control group of 50 schools provided the 'business as usual' counterfactual.

Table 3 shows that with a realistic level of intra-class correlation (around 0.1), it would be possible to detect an effect size of around 0.15 standard deviations when there are 50 schools in each group (with power 80% and a significance level of 5%).<sup>11</sup> This is similar to the effect size found for the universal provision of free school meals on KS1 and KS2 test scores found in previous research; standardised scores increased by around 0.15 standard deviations, varying slightly across subgroups and areas (Brown *et al.*, 2012).

These sample size calculations were replicated under various assumptions of the correlation in outcomes between pupils at the same school (the intra-cluster correlation of outcomes), and various

<sup>11</sup> The value of 0.1 is roughly the ICC observed in data gained from the IFS evaluation of The LIT Programme for the sub-sample of schools with at least 35% of FSM pupils. The EEF guidance on ICCs for KS1 average point score is 0.110 and for KS2 overall fine grade score is 0.139.

assumptions about the proportion of the variance in outcomes that is unexplained by observable characteristics (as this was not known for the relevant schools at the time). Model 1 reports the detectable effect size when the variance of the outcome unexplained by attributes of the pupils and schools (including prior attainment) is 60% (which is similar to that found in the LIT Programme evaluation in Crawford and Skipp, 2014). This evaluation was used as a benchmark as it was available internally at IFS before guidance from the EEF on reasonable assumptions for power calculations was made available. Model 2 reports a less optimistic scenario (70% unexplained), while Model 3 is more optimistic (50% unexplained). The final minimum detectable effect sizes are presented in Table 3. All minimum detectable effect sizes were computed using Microsoft Excel.

**Table 3: Sample size calculations**

| Intra class correlation                         |       |       |       |       |       |       |
|---|-------|-------|-------|-------|-------|-------|
|   | 0     | 0.05  | 0.1   | 0.15  | 0.2   | 0.25  |
| <b>Minimum Detectable Effect Size (Model 1)</b> | 0.069 | 0.118 | 0.152 | 0.180 | 0.204 | 0.225 |
| <b>Minimum Detectable Effect Size (Model 2)</b> | 0.074 | 0.127 | 0.164 | 0.194 | 0.220 | 0.243 |
| <b>Minimum Detectable Effect Size (Model 3)</b> | 0.063 | 0.108 | 0.139 | 0.164 | 0.186 | 0.205 |

*Note: These calculations represent the effect size that will be possible to detect using a two-sided hypothesis test with significance level of 5%, and with power against an alternative hypothesis of 80%. We have assumed 40 pupils per cohort (based on publicly available school-level data for schools with more than 35% of pupils currently eligible for free school meals) for 50 intervention and 50 control schools.*

## Randomisation

Randomisation was done using a random number generator using Stata 13 within strata (or sampling blocks) to achieve a good balance of baseline observable characteristics. Strata were defined according to the following characteristics drawn from the publicly-available school census and performance tables, which either differed widely within the recruited sample (percentage of pupils with English as an additional language, ‘EAL pupils’) or which provide important context to any potential treatment effects that might be observed (average prior attainment). The strata characteristics were:

- type of school—infant school or junior/primary school;
- prior school-level average attainment—above median in the sample, below sample median (where attainment is measured at KS1 for infant schools and KS2 for junior and primary schools); and
- percentage of EAL pupils —above sample median, below sample median.

The protocol pre-specified that either the percentage of EAL pupils or the percentage of FSM pupils would be used. The percentage with EAL was chosen as there was more variation in this school characteristic in the sample of schools recruited.

The number of schools in each stratum is:

- infant: below median KS1, 4;
- infant: above median KS1, 3<sup>12</sup>;
- junior/primary, missing data/small schools, 13;
- junior/primary, below median KS2 and below median EAL, 18;

<sup>12</sup> Publicly-available data collected pre-randomisation based on 2012 information indicated that seven of the schools in our sample were infant schools. However, one school also admitted Year 6 pupils in 2014/15, so we use a figure of six infant schools elsewhere in the report.

- junior/primary, below median KS2 and above median EAL, 23;
- junior/primary: above median KS2 and below median EAL, 21; and
- junior/primary, above median KS2 and above median EAL, 24.

Randomisation was done by IFS in August 2014 after all schools had signed the Memorandum of Understanding and all headteachers had completed a baseline survey, but before the baseline teacher and pupil surveys were completed.

The Stata code for this randomisation is given in Appendix E.

## Analysis

Primary and secondary analyses are on an ‘intention-to-treat’ basis, which means that all schools allocated to the intervention and control groups are used to estimate the effect of the intervention independent of their level of engagement with the intervention, control group activity, or continued participation. We have also presented sensitivity analyses that impose additional restrictions on the schools included in the analysis. One analysis uses a sample restriction that includes only schools where no pupils are known to be observing Ramadan (the Muslim month of fasting) as it became apparent that Ramadan overlapped with the school year and therefore affected some students’ breakfast consumption (see Tables 7 and 8).<sup>13</sup> Another ‘on-treatment’ analysis uses a restriction which excludes schools that deviated from the intended treatment, either by charging for the breakfast club or by restricting attendance (see Table 7).

Effect sizes are calculated as the difference in the standardised outcome between intervention and control groups, on average, conditional on pupil and school characteristics that may differ across the two groups. The outcome is standardised on the sample to have a mean of zero and standard deviation of one. The effect size is computed through ordinary least squares (OLS) regression for all outcomes except those relating to pupil hunger. The three hunger outcomes (‘ate breakfast’, ‘ate breakfast at school’, and ‘hungry at the start of the morning’) are indicator variables, so the analysis is conducted with logistic regression. Average marginal effects are then computed and presented as the effect size. For primary and secondary outcomes, which are measured at the pupil level, OLS is run at the pupil level, controlling for randomisation strata, the relevant baseline measure, pupil demographics, and school characteristics. In detail, these are:

- randomisation strata (entered as a series of binary variables);
- relevant baseline measure—
  - for KS1, binary pupil-level variables for whether development measured at the Foundation Stage Profile was at the expected level or missing (relative to not at the expected level),
  - for KS2, binary pupil-level variables for categories of average attainment on KS1 tests—11–14.99 points (did not achieve grade level), 15–16.99 points (at grade level), 17+ points (above grade level), or missing (relative to fewer than 11 points—did not achieve any level 2),
  - for pupils’ breakfast consumption and hunger—pupil-level breakfast consumption and hunger measured in the baseline pupil survey,
  - for teachers’ perceptions of pupil behaviour and concentration—classroom-level pupil behaviour and concentration measured in the baseline teacher survey,
  - for attendance—a continuous pupil-level measure of sessions lost in the previous academic year (2013/2014), and
  - for late arrival—a continuous pupil-level measure of late arrivals to sessions in the previous academic year (2013/2014);

<sup>13</sup> This exclusion is based on schools in which at least one pupil spontaneously discusses ‘Ramadan’ or ‘fasting’ in their survey. This is unlikely to capture all fasting and will necessarily also exclude some non-fasting students.

- pupil demographics—binary variables for female, ever eligible for FSM, any special educational need, English as an Additional Language, and major ethnic group (Black, Asian, Mixed, and Other, with White as the reference group); and
- school characteristics—binary variables for latest measure of Ofsted effectiveness prior to the intervention ('good', 'requires improvement', 'inadequate', and 'missing', with 'outstanding' as the reference group), rural-urban category ('semi-urban' and 'major urban', with 'rural' as a reference category), continuous variables for percentile of the school Lower Super Output Area's rank on the Index of Multiple Deprivation, and the total number of students in the school.

Clustering of pupils within schools is taken into account using robust standard errors (clustered sandwich estimator) that allow observations within schools to be correlated. This approach was used rather than multi-level modelling (or random effects) to account for the clustering of pupils within schools because the model does not impose the assumption that the school-level effect and pupil- and school-level covariates in the model are independent. The method of randomisation (stratification) is accounted for by including the variables used to stratify as regressors in the OLS regression.

Health outcome (BMI) data is available at the school level rather than the pupil level. The relevant baseline measure (BMI data for the Year 6 pupils in the year before the trial) is also at the school level under the assumption that the health of pupils in adjacent cohorts within schools is positively correlated. We account for relevant school characteristics as in the pupil-level regressions specified above.

The inclusion of measures of the outcomes of interest taken before the intervention started means we are effectively comparing the change in outcomes over time among pupils and schools in the treatment group against the change in outcomes among pupils and schools in the control group. We include these and other baseline characteristics in our analysis because it maximises the chance of detecting a statistically significant impact on each outcome of interest (as outlined in the section on sample size calculations above). As we shall see below, there are also small differences in characteristics between the treatment and control group at baseline which might otherwise bias our estimates if we did not account for them.

The small number of observations with missing demographic covariates in the NPD was not included in the final sample. This is 62 pupils at KS1 and 65 pupils at KS2. Appendix F discusses this sample restriction.

Subgroup analysis is conducted according to the evaluation protocol. Of particular interest is the impact for pupils that have ever been eligible for free school meals, a group important for the EEF and Magic Breakfast. Other subgroups of interest are boys versus girls, students with low prior attainment versus those who achieved the expected level of attainment or better in their last assessment, and students who did not report having breakfast in the baseline surveys versus those who did.

Additional analysis not pre-specified in the evaluation protocol relates to the results for the sub-set of schools where no students mentioned observing Ramadan in the follow-up pupil survey, and the sub-set of schools in the intervention group whose breakfast offer was universal and free (two of the components of the intended model as described in the Trial Design section).

## Implementation and process evaluation

Four case studies were carried out to explore how schools were delivering their breakfast provision. As far as was feasible, the four schools were selected to ensure a mix in terms of school size, delivery approach, area type, ethnic profile, and number of children in receipt of free school meals. A breakdown of the achieved sample can be found in Appendix G.

Visits lasting one or two days were carried out in each of the four schools. During these visits researchers observed the delivery of breakfast provision and interviewed:

- staff managing the provision—headteachers, School Breakfast Co-ordinators and Magic Breakfast school change leaders (ten members of staff altogether);
- staff delivering the provision—teaching assistants, pastoral staff, and parents (nine members of staff);
- teachers (five members of staff);
- children who took up the breakfast provision (30 children);
- children who did not take up the breakfast provision (eight children); and
- parents of children who took up the breakfast provision (eight parents).

The interviews and focus groups were based on topic guides that outlined the main topics that were to be addressed depending on the role of the participant and the nature of their involvement in the delivery of Magic Breakfast. Interviews with staff lasted between 40 and 90 minutes, while interviews with parents ranged from 20 to 40 minutes. Interviews with children were tailored according to their needs and varied in length from 25 to 45 minutes. Pupils were asked to create their ideal breakfast plate using paper breakfast plates and stickers of popular breakfast foods. They were also invited to design their ideal breakfast menu using printed menus.

The interviews and focus groups were recorded and transcribed and then the data was systematically and comprehensively analysed using 'Framework' (Ritchie and Spencer, 1994). This approach involves producing a thematic framework consisting of a series of worksheets, each of which addressed a particular theme from the research. It enabled participants' views and experiences to be compared and contrasted within a particular school and across different schools and delivery models.

The qualitative findings presented in this report have been illustrated through the use of quotes. Adopting a qualitative approach has enabled detailed exploration of how four schools delivered their breakfast provision and provided useful learning for other schools. However, the purposive nature of the sample design and the small number of case studies means that we cannot provide any statistical data about the prevalence of these views. The report deliberately avoids giving numerical findings relating to the qualitative evidence, or trying to generalise the findings to other schools that may have chosen different delivery approaches.

## Costs

The cost is calculated as the cost per pupil per year averaged over three years, in line with EEF guidance. The relevant figure is the cost to a school of delivering the same intervention outside the conditions of the trial. For example, schools in the intervention group received free food throughout the year and a £300 grant from Magic Breakfast towards the capital set-up costs which would not normally be available to schools establishing a breakfast club, so these costs have been included in the figures presented.

The costs of universal breakfast club provision for schools in the absence of the trial are: capital set-up costs (including food storage, preparation, and games), ongoing food costs, and ongoing staff costs (for supervision and food preparation/serving). Capital set-up costs were collected from a survey of headteachers at the end of the intervention in June 2015. This survey also collected ongoing staff costs. This varied dramatically across schools, depending on whether volunteers or existing members of staff were used, and whether and how these existing members of staff were remunerated. In line with EEF guidance, we have reported these staffing costs in units of time rather than as financial costs. Ongoing food costs were reported by Magic Breakfast, which in the trial provided most of the breakfast food and drink at no cost to the school. Outside the trial the schools would incur a cost for this food, unless partnering with a charity such as Magic Breakfast.

## Timeline

### Table 4: Timeline

| Date                     | Activity  |
|--------------------------|---|
| May to June 2014         | Magic Breakfast recruited schools   |
| June–August 2014         | Baseline headteacher survey delivered through email   |
| Early August 2014        | Randomisation<br>Informed schools about randomisation outcome   |
| August–September 2014    | Magic Breakfast co-ordinated breakfast club implementation  |
| September 2014           | Ensured baseline pupil survey of breakfast and lunch consumption is delivered to schools and completed<br>Delivered baseline classroom survey to Year 2 and Year 6 teachers through email   |
| September 2014–July 2015 | Magic Breakfast maintained support for schools<br>Magic Breakfast maintained record of costs of food and delivery associated with each breakfast club   |
| February 2015            | Developed tools—topic guides, survey, focus group activities  |
| March 2015               | Case study schools approached   |
| April–June 2015          | Fieldwork in four schools   |
| June 2015                | Delivered final headteacher survey which collected information about <ul style="list-style-type: none"> <li>- costs of additional food items provided by the school</li> <li>- equipment (including cost if possible) used by the schools but not provided by Magic Breakfast</li> <li>- other activities associated with breakfast clubs (e.g. games)</li> </ul> Ensured final pupil survey of breakfast and lunch consumption is delivered to schools and completed<br>Delivered final classroom survey to Year 2 and Year 6 teachers via email<br>Delivered headteacher survey through email |
| July–October 2015        | Analysed and reported process evaluation findings   |
| August–September 2015    | Magic Breakfast co-ordinated breakfast club provision for control schools   |
| October 2015             | Applied for access to NPD for KS1 and KS2 results (that took place in May 2015), and absence data for September 2013 to July 2015<br>Applied for access to National Child Measurement Programme BMI data  |
| January 2016–June 2016   | Analysed impact   |
| June 2016                | Draft final report sent to the EEF  |

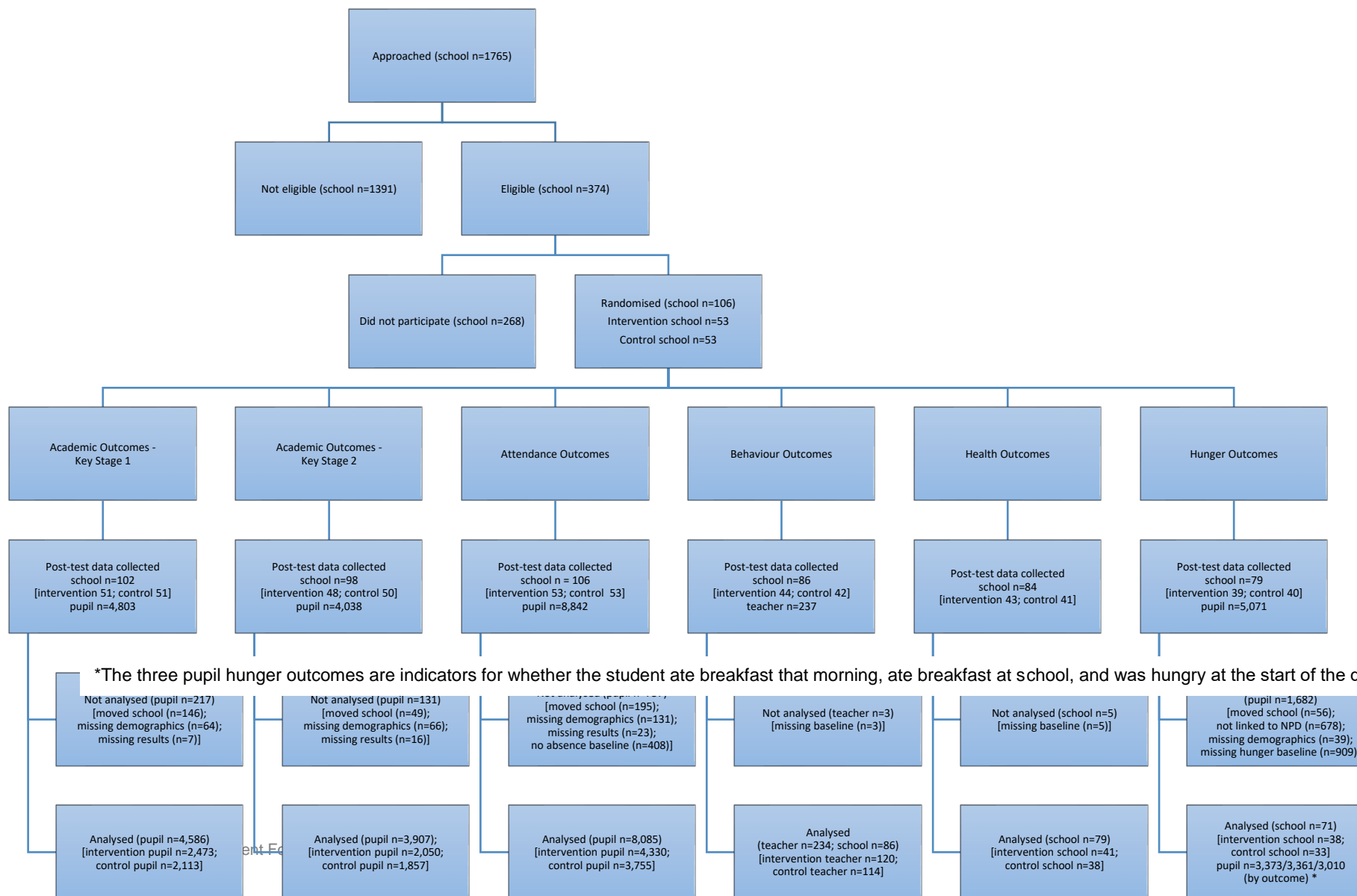


|                     |                              |
|---------------------|------------------------------|
| <b>October 2016</b> | <b>Final report released</b> |
|---------------------|------------------------------|

## Impact evaluation

The flow of participants through the evaluation is shown in Figure 2 below.

Figure 2: Participant flow diagram



**Table 5: Minimum detectable effect size at different stages**

| Stage                                 | N [schools]<br>(n = intervention;<br>n = control) | Correlation<br>between pre-<br>test (+other<br>covariates) &<br>post-test | ICC  | Blocking/<br>stratification or<br>pair matching | Power | Alpha | Minimum<br>detectable<br>effect size<br>(MDES) |
|---------------------------------------|---|---|------|---|-------|-------|--|
| Protocol                              | 100 (50; 50)                                      | 0.60  | 0.10 | Stratification                                  | 80%   | 0.05  | 0.15   |
| Randomisation<br>and analysis,<br>KS1 | 102 (51; 51)                                      | 0.61  | 0.04 | Stratification<br>(106 schools)                 | 80%   | 0.05  | 0.11   |
| Randomisation<br>and analysis,<br>KS2 | 98 (50; 48)                                       | 0.65  | 0.11 | Stratification<br>(106 schools)                 | 80%   | 0.05  | 0.16   |

The minimum detectable effect size at different stages of the evaluation is presented in Table 5. For KS1, there are slightly more schools than initially assumed, and more pupils per school-cohort than originally assumed (46 versus 40). This means that the sample size and hence the power of statistical tests is greater than expected, all else equal, which reduces the minimum detectable effect size. For KS2 outcomes, the intra-cluster correlation (ICC) is slightly higher than initially assumed, which reduces the power of statistical tests, all else equal. For KS1 the ICC is much lower than initially assumed which contributes to the lower minimum detectable effect size for these outcomes (0.11 at KS1 compared to 0.16 at KS2).

### Pupil characteristics

Magic Breakfast approached 1,765 schools to participate in this trial, but the majority were not eligible to participate due to existing breakfast club provision. Of the 374 who were eligible, 106 schools across 59 local authorities signed the Memorandum of Understanding to participate in the intervention. These schools were then randomised into the intervention and control groups, as described above, with 53 schools in each.

Table 6 shows the characteristics for schools in the intervention and control group for the Magic Breakfast intervention, and the characteristics for pupils at these schools. Statistically significant differences between the intervention and control groups are denoted by stars on the relevant characteristics.

Magic Breakfast targets schools where over 35% of pupils are currently FSM-eligible, or where more than half have been eligible at any point in the previous six years. The schools that were eligible to participate in this trial therefore have a more disadvantaged pupil intake than the average school. For example, the NPD data indicates that the median school in this trial has 43% of its Year 2 and Year 6 students currently receiving FSM, whereas nationally 15% of pupils are claiming FSM (DfE, 2015b). There is a wide range of EAL prevalence: while one school in the trial has no EAL students in Year 2 or Year 6, 95% of these cohorts in another school have a first language other than English.

Most of the participating schools are located in major urban areas, with just 18% of the schools in a rural or mostly rural local authority district (this is shown split by treatment status in Table 6). The schools that were randomly allocated to receive the intervention are particularly urban—over three-quarters of them are located in a major urban centre. The schools are mostly located in highly deprived areas, with 82% of schools in the 30% most deprived neighbourhoods (based on the 2010 Index of Multiple Deprivation scores). There is a mixture of school types and religious affiliations. Around 60% of

participating schools are community schools, with the rest an even split between academies and voluntary schools. Four-fifths of the schools have no religious affiliation. The majority of the schools that were inspected by Ofsted prior to the start of the intervention were rated as 'good', and just two schools were rated 'inadequate' at their most recent inspection.

All headteachers of schools in the trial completed a baseline survey prior to randomisation. The responses to this survey are reasonably balanced. For example, the proportion of schools where 'some' or 'most' of the pupils at the school have 'bad behaviour' is roughly 30% in the intervention and control groups. Improving pupil health and wellbeing is the main reason for signing up to the trial in both intervention and control groups (49% and 47% respectively). Reducing pupil hunger is the main motivation for 34% of intervention schools and 26% of control schools, but this difference is not statistically significant. There is also some difference in the proportion of schools that offered some ad-hoc breakfast provision prior to signing up (42% of intervention schools and 47% of control schools), but again this difference is not statistically significant.

With the exception of a school's urban-rural location and the percentage of pupils with healthy weight, Table 6 indicates that there are no significant differences between intervention and control schools according to school-level characteristics. Intervention schools are much more likely to be in urban areas than control schools (around 76% compared to 36%). Pupils in intervention schools are less likely to be a healthy weight before the trial than pupils in control schools (57% compared to 61%).

There are also few significant differences between the two groups of schools in pupil-level characteristics. There are differences between pupils in the intervention and control schools, on average, but these are typically small (in the range of two to three percentage points) and not statistically significant. For example, the intervention group contains more Black and fewer Asian students than the control group. The intervention group schools also have a slightly lower proportion of FSM students, students with special educational needs, and non-native English speakers.

Pupils in intervention schools have significantly fewer authorised absences in the academic year prior to the intervention, on average. Although this difference is small in percentage point terms, it is large in percentage terms given the low baseline number.

In addition to data on demographics and attainment from the NPD, information from pupils on their consumption of breakfast and their level of hunger was collected through paper surveys. Pupils in the intervention and control schools were equally likely to consume breakfast prior to the start of the intervention, with around 91% of pupils in both intervention and control schools eating breakfast overall, and around 10% of pupils in both intervention and control schools eating breakfast at school. A higher proportion of pupils in the intervention schools reported being hungry at the start of the school day, however, which was a significant difference to the control schools.

**Table 6: Baseline comparison**

| Variable                                  | Intervention group         |               | Control group |               |            |
|---|----------------------------|---------------|---------------|---------------|------------|
|   | School-level (categorical) | n/N (missing) | Percentage    | n/N (missing) | Percentage |
| <b>Religiously affiliated<sup>a</sup></b> |                            | 11/53 (0)     | 20.8          | 11/53 (0)     | 20.8       |
| <b>Community school</b>                   |                            | 30/53 (0)     | 56.6          | 32/53 (0)     | 60.4       |
| <b>Academy</b>                            |                            | 12/53 (0)     | 22.6          | 10/53 (0)     | 18.9       |
| <b>Voluntary or Foundation school</b>     |                            | 11/53 (0)     | 20.8          | 11/53 (0)     | 20.8       |
| <b>Ofsted: Outstanding<sup>b</sup></b>    |                            | 4/53 (0)      | 7.5           | 4/53 (0)      | 7.5        |
| <b>Ofsted: Good<sup>b</sup></b>           |                            | 30/53 (0)     | 56.6          | 33/53 (0)     | 62.3       |

|   |                      |                   |                      |                   |
|---|----------------------|-------------------|----------------------|-------------------|
| <b>Ofsted: Satisfactory<sup>b</sup></b>                                       | 12/53 (0)            | 22.6              | 10/53 (0)            | 18.9              |
| <b>Ofsted: Inadequate<sup>b</sup></b>   | 0/53 (0)             | 0.0               | 2/53 (0)             | 3.8               |
| <b>School in urban area<sup>c***</sup></b>                                    | 40/53 (0)            | 75.5              | 19/53 (0)            | 35.8              |
| <b>School in rural area<sup>c***</sup></b>                                    | 3/53 (0)             | 5.7               | 16/53 (0)            | 30.2              |
| <b>'Some' or 'most' pupils badly behaved</b>                                  | 17/53 (0)            | 32.1              | 16/53 (0)            | 30.2              |
| <b>Main motivation for joining trial: improve pupil health and well-being</b> | 26/53 (0)            | 49.1              | 25/53 (0)            | 47.2              |
| <b>Main motivation for joining trial: reduce pupil hunger</b>                 | 18/53 (0)            | 34.0              | 14/53 (0)            | 26.4              |
| <b>School offers some ad hoc provision before trial</b>                       | 22/53 (0)            | 41.5              | 25/53 (0)            | 47.2              |
| <b>School-level (continuous)</b>  | <b>n (missing)</b>   | <b>Mean</b>       | <b>n (missing)</b>   | <b>Mean</b>       |
| <b>Number of pupils</b>   | 53 (0)               | 333.9             | 53 (0)               | 291.5             |
| <b>% Free School Meals</b>  | 53 (0)               | 43.8              | 53 (0)               | 44.1              |
| <b>% English as Additional Language</b>                                       | 50 (3)               | 35.5              | 47 (6)               | 31.3              |
| <b>KS2 Average Point Score (overall)<sup>d</sup></b>                          | 46 (7)               | 27.5              | 48 (5)               | 27.4              |
| <b>KS2 Overall Value-Added Measure<sup>d</sup></b>                            | 46 (7)               | 100.2             | 48 (5)               | 100.4             |
| <b>School LSOA's IMD 2010 rank<sup>e</sup></b>                                | 53 (0)               | 15.5              | 53 (0)               | 19.6              |
| <b>School LSOA's IMD 2010 score<sup>e</sup></b>                               | 53 (0)               | 42.5              | 53 (0)               | 41.3              |
| <b>% of pupils with healthy weight<sup>**</sup></b>                           | 44 (9)               | 57.0              | 41 (12)              | 61.2              |
| <b>Pupil-level (categorical)</b>  | <b>n/N (missing)</b> | <b>Percentage</b> | <b>n/N (missing)</b> | <b>Percentage</b> |
| <b>Female<sup>f</sup></b>   | 2283/4609 (0)        | 49.5              | 1969/4038 (0)        | 48.8              |
| <b>Ethnicity: White<sup>f</sup></b>   | 2505/4546 (63)       | 55.1              | 2228/4000 (38)       | 55.7              |
| <b>Ethnicity: Black<sup>f</sup></b>   | 961/4546 (63)        | 21.1              | 653/4000 (38)        | 16.3              |
| <b>Ethnicity: Asian<sup>f</sup></b>   | 478/4546 (63)        | 10.5              | 640/4000 (38)        | 16.0              |
| <b>Ethnicity: Mixed<sup>f</sup></b>   | 342/4546 (63)        | 7.5               | 252/4000 (38)        | 6.3               |
| <b>FSM: Currently Eligible<sup>f</sup></b>                                    | 1553/4574 (35)       | 34.0              | 1430/4009 (29)       | 35.7              |
| <b>FSM: Ever Eligible<sup>f</sup></b>   | 2478/4574 (35)       | 54.2              | 2300/4009 (29)       | 57.4              |

|  |                    |             |                    |             |
|--|--------------------|-------------|--------------------|-------------|
| <b>SEN: Any recorded<sup>f</sup></b>                                   | 983/4574 (35)      | 21.5        | 954/4009 (29)      | 23.8        |
| <b>English as Additional Language<sup>f</sup></b>                      | 1786/4568 (41)     | 39.1        | 1645/4003 (35)     | 41.1        |
| <b>Ate breakfast today</b>   | 3211/3526 (1611)   | 91.1        | 2326/2556 (1887)   | 91.0        |
| <b>Ate breakfast at school</b>   | 346/3506 (1631)    | 9.9         | 269/2536 (1907)    | 10.6        |
| <b>Hungry at start of day*</b>   | 1230/3212 (1925)   | 38.3        | 786/2356 (2087)    | 33.4        |
| <b>Will eat lunch today</b>  | 3449/3538 (1599)   | 97.5        | 2483/2564 (1879)   | 96.8        |
| <b>Good level of development at FSP<sup>g</sup></b>                    | 1059/2376 (139)    | 44.6        | 838/2012 (130)     | 41.7        |
| <b>Pupil-level (continuous)</b>  | <b>n (missing)</b> | <b>Mean</b> | <b>n (missing)</b> | <b>Mean</b> |
| <b>Total half-sessions absent in 2013/14<sup>h</sup></b>               | 4376 (233)         | 13.0        | 3792 (246)         | 13.8        |
| <b>Total authorised half-sessions absent in 2013/14<sup>h*</sup></b>   | 4376 (233)         | 9.5         | 3792 (246)         | 10.3        |
| <b>Total unauthorised half-sessions absence in 2013/14<sup>h</sup></b> | 4376 (233)         | 3.6         | 3792 (246)         | 3.5         |
| <b>Total half-sessions late in 2013/14<sup>h</sup></b>                 | 4609 (233)         | 0.3         | 4038 (246)         | 0.4         |
| <b>Reading points at KS1<sup>i</sup></b>                               | 1914 (180)         | 14.5        | 1698 (197)         | 14.4        |
| <b>Writing points at KS1<sup>i</sup></b>                               | 1914 (180)         | 13.3        | 1698 (197)         | 13.2        |
| <b>Maths points at KS1<sup>i</sup></b>                                 | 1913 (181)         | 14.8        | 1697 (198)         | 14.6        |

**Notes:**

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*)

Pupil-level variables account for the clustering of pupils within schools.

<sup>a</sup> Religious affiliation includes both Church of England and Catholic schools.

<sup>b</sup> Ofsted ratings are the most recent overall effectiveness ratings available as at 2014.

<sup>c</sup> Urban-rural classification is from the six-point 2011 DEFRA classifications of U.K. local authority districts. 'Rural' refers to areas in the bottom three tiers; 'urban' refers to those in Tier 6, 'Major Urban' settlements.

<sup>d</sup> KS2 average point score and overall value-added measures are measured for the previous cohort since test results for the cohort of interest are not available for the year before the intervention.

<sup>e</sup> Index of Multiple Deprivation rank is from the 2010 Department for Communities and Local Government's overall LSOA-level IMD. Lower ranks and higher scores imply greater deprivation.

<sup>f</sup> Innate demographic characteristics (ethnicity, FSM, EAL, etc.) are measured on the cohorts of interest.

<sup>g</sup> A good level of development at Foundation Stage Profile is checked for balance only among Year 2 pupils as the relevant measure of prior attainment for this group.

<sup>h</sup> Absences are measured for the cohorts of interest but prior to the start of the intervention.

<sup>i</sup> KS1 point scores are measured only for pupils in Year 6, who wrote the KS1 tests three years prior to the intervention.

## Sample sizes and missing data

To provide evidence on the primary outcome of interest and mechanisms through which any impact occurs, this report considers a wide range of outcomes from diverse sources. Where possible, the project team has linked these data sources together to allow for a richer set of controls in the analysis. However, to maximise analytical power for the primary outcome, and to minimise potential bias from non-response, we allow the sample to differ by outcome. For the primary academic and pupil-reported

outcomes we also conduct sensitivity analysis, exploring the effect of imposing different restrictions on the sample.

The sample of interest is students in intervention and control schools in Year 2 and Year 6 in the academic year 2014/2015. These pupils were assessed at the end of KS1 and KS2, respectively, in May 2015.

The precise number of pupils initially randomised into the trial is unknown. This is because administrative data, which provides the best information on pupil counts, is typically only collected at three points in the year (in the termly School Census). The best estimate of the initial number of pupils randomised that we have access to is from the NPD, which provides information about each pupil who wrote their national assessments in one of the schools in our sample in the 2014/15 academic year. However, this figure could differ from the count of pupils initially randomised in two ways: (a) it will exclude students who attended one of the schools in our sample at the start of the academic year, but moved to a different school part-way through the year; and (b) it will include students who joined one of the schools in our sample part-way through the year. Previous research suggests that about six per cent of Year 2 students enter a new school part-way through the school year, while the corresponding figure at Year 6 is about three per cent (authors' calculations based on Rodda et al., 2013 and national state-funded pupil counts).

Academic outcomes (KS1 and KS2 test scores and teacher assessments) come from the NPD, provided by the DfE. This is a total sample of 8,842 students (4,803 in KS1 and 4,038 in KS2, with one student missing information on their year group).

DfE matched this sample of pupils to surveys of breakfast consumption at the start and end of the academic year. Not all pupils are matched to their breakfast survey(s), due to non-response by some schools, pupil absence on the day of the survey, and insufficient or incorrect information to enable the link. In total, 6,834 pupil records were successfully matched to at least one survey response (at baseline or follow-up). There were 2,013 NPD records that did not match to any survey, and 968 students who completed a survey but whose responses were not matched to the NPD. The latter is equivalent to 12% of the 7,802 pupils who completed either survey.

In addition, there are a number of schools with missing data. Although 106 schools in total enrolled in the trial, as indicated in Figure 2 KS1 data is available for only 102 of them. This is because there are four junior schools in the sample, with no pupils in Year 2. At KS2 we have data for 98 schools. This reflects the six infant schools in the sample (with no Year 6 pupils). However, there are an additional two schools for which no KS2 data is available. One of these schools has a special programme at Year 6 that could explain why this data is not available. The reason for the unavailability of the other school's KS2 data is unknown.

The effective sample size for academic outcomes is 4,586 pupils in KS1 and 3,907 pupils in KS2 (8,493 in total). This sample includes all students assessed in an intervention or control school in the NPD (those that did not move schools mid-year) with non-missing demographic characteristics (gender, FSM, SEN, and EAL status, and major ethnic group) and a full set of academic outcomes (tests and teacher assessments). There are fewer KS2 than KS1 pupils, partly because six of the trial schools were infant schools with no Year 6 pupils, while only four schools were junior schools with no students in Year 2.

There were some missing values in the school-level data. Of the 106 schools, 11 did not have pre-intervention Ofsted inspection ratings. We created an additional category for missing rating rather than dropping these schools' observations when running regressions with school-level controls.

For the secondary outcomes that illuminate the mechanisms through which breakfast club provision may have influenced attainment, separate samples are used:



- *Breakfast consumption and reported hunger*: The sample is restricted to the pupils that responded to the baseline and follow-up survey and with both surveys linked to the NPD. This leads to a sample of 3,379 (2,019 in Year 2 and 1,359 in Year 6). Sensitivity analysis, which does not impose the condition that a baseline survey was linked to the NPD, gives a sample of 4,290 students. In addition, some students did not respond to all questions, so there is slight variation in the sample size by outcome. This is particularly pronounced for the question on pupil hunger, where non-response means the sample size drops to 3,010.
- For *absence and punctuality*, the sample is based on the main academic analysis sample but restricted to those with absence observed in the previous academic year (2013/2014). This leads to a sample of 8,085 (4,512 in Year 2 and 3,573 in Year 6).
- For *classroom behaviour and concentration*, the sample is teachers who responded to both the baseline and follow-up surveys. This includes 234 teachers.
- For *health*, proxied by body mass index and the percentage of the year group at a healthy weight, school-level data (for Year 6 pupils) was provided by the National Child Measurement Programme (NCMP). The sample is schools where the school-level average was observed in the year of the intervention (2014/2015) and the previous academic year. (The data includes 84 schools in 2015 and 85 schools in 2014; 79 of these schools have data for both years.)

Separate samples were used for these secondary outcomes to ensure relevant baseline information was available in each regression analysis while maintaining the larger sample size (without these sample conditions) for the primary outcome.

## Outcomes and analysis

This evaluation assesses the effect of the intervention on academic outcomes (the primary outcomes of interest) and on mechanisms through which academic attainment may be influenced. These are: breakfast consumption, pupil absence and punctuality, pupil behaviour, pupil concentration, and pupil health (as proxied by Body Mass Index). Breakfast consumption was hypothesised to increase (and pupil hunger decrease) in the intervention schools more than the control schools, although the use of breakfast clubs in some of the control schools may have compromised this in that, as a result, pupil attainment might also have risen in some control schools, diluting the estimated treatment effect of the intervention. Pupil absence and late arrivals would be expected to decrease more in the intervention schools, partly due to the additional incentive to arrive at school on time and ease for parents, and perhaps due to improvements in pupil health (meaning fewer days of sickness). Pupil behaviour and concentration in class would be expected to improve more in intervention schools if the provision of breakfast allows pupils to focus more in the classroom. This could be because hunger limits pupil concentration, because of higher nutritional value of the food provided at the breakfast club, or because the social nature of the breakfast club induces learning once in the classroom. We discuss each primary and secondary outcome in turn. Each specification includes pupil and school covariates to account for the small imbalances in these factors across the intervention and control groups, and to improve precision of the estimate (as discussed in the Analysis section of the Introduction).

The estimated difference between pupil outcomes in the intervention and control group is an intention-to-treat analysis. This means that all control schools are included, regardless of whether they established a breakfast club during the trial, and all intervention schools are included, regardless of whether they implemented the breakfast club in the agreed form.

Survey data provided by 79% of intervention schools (42 schools) revealed that there was limited fidelity to the agreed intervention, defined as establishing a universal, free, before-school breakfast club at the school. Just a third of the intervention group schools that responded to the follow-up headteacher survey reported that they implemented a universal, free, before-school breakfast club as agreed before the trial. Around one quarter of schools in the intervention group (that responded to the follow-up headteacher survey) had breakfast club provision during the school day (including during the

registration period), instead of, or in addition to, before the school day. Of schools in the intervention group that provided breakfast before the school day as intended, around 57% made access universal and free. Overall, 39% of responding schools in the intervention group had a charge for some or all pupils that attended the breakfast club, and this was less prevalent (11%) in schools where the breakfast club was during the school day.

This was echoed in the process evaluation, where one of the case study schools organised its breakfast club provision as part of the school day rather than before and two of the schools charged for breakfast club attendance. As discussed earlier, schools deviated from the universal, free, before-school provision in part due to funding and access considerations specific to each school. Schools' deviation from the agreed method of breakfast club provision might suggest that some types of breakfast club are more practical in particular school contexts. For example, schools in the case studies weighed access and funding considerations differently, and parents' demand for childcare influenced the timing of the breakfast club.

Around 40% of schools in the control group that responded to the follow-up survey adopted some form of breakfast club throughout the year. The majority of these schools (15 of 17) did not charge for the breakfast provision, but it was targeted rather than universal in some cases (6 of 17). This level and type of provision is in line with the 'business as usual' design of the trial, but the relatively high presence of breakfast clubs in the control schools must be kept in mind when interpreting the evaluation findings. Around 90% of schools in the control group that responded to the follow-up survey also established or encouraged a breakfast club for Year 6 pupils in the week of testing for national exams which may dilute any differences in attainment between Year 6 pupils in intervention and control schools.

Pupil take-up of the breakfast club offer was lower than expected before the intervention. This may be a response to schools' decisions to deviate from the universal and free breakfast club offer, an indication of additional barriers to take-up, or a reflection that pupil hunger was not as large a problem as initially thought. Of intervention schools that responded to the follow-up headteacher survey, the majority (58%) of schools that provided breakfast before the school day fed on average 21 to 50 pupils per day, and around one third (36%) fed on average more than 50 pupils per day. Two-thirds of the nine intervention schools that offered breakfast during the school day on average provided breakfast to over 50 pupils per day.<sup>14</sup>

Analysis code is given in Appendix H.

## Academic outcomes

Table 7 reports the primary outcomes: the effect of breakfast club provision on pupils' attainment at KS1 and KS2.

Breakfast club provision has positive, statistically significant impacts on KS1 scores in maths, reading, and writing. Progress in maths and writing was moderately enhanced by the intervention; scores in intervention schools were on average 0.14–0.15 standard deviations higher, equivalent to about two months' progress. The impact for KS1 reading was a slightly smaller, though still significant, at 0.10 standard deviations.

The magnitude is similar to the estimated effect on KS1 attainment of providing universal free school meals in two relatively disadvantaged pilot areas (Brown *et al.*, 2012). Taken together, these results suggest that breakfast and lunch consumption may have additive effects as the improvements in

<sup>14</sup> We provide absolute numbers of attendees rather than proportions because the range of proportions is very large and would have an upper bound of 100% for the top category, limiting the usefulness of the data.

attainment from breakfast club provision in trial occurred even after the introduction of universal free school meals for all pupils in Year 2.

At KS2, the effects of intervention on primary outcomes are smaller and not statistically significant. KS2 reading scores were affected to a similar degree as reading scores at KS1, but a larger degree of uncertainty means that the KS2 results are not statistically different from zero. Maths results at KS2 were positively affected but by a small amount (0.07 standard deviations) that is not significantly different from zero.

This is in contrast to the evaluation of universal free school meals provision, where universal provision was found to improve academic attainment *more* for KS2 than KS1 pupils (Brown *et al.*, 2012). Possible reasons include:

- the fact that 91% of control schools (who responded to our follow-up survey) offered large-scale breakfast provision to Year 6 pupils during the week of KS2 tests;
- higher breakfast consumption prior to the intervention for Year 6 pupils (as reported in the baseline pupil survey);
- Year 6 pupils being more likely to fast during Ramadan, which occurred during the school year; and
- the fact that the KS1 primary outcomes were teacher-assessed and the KS2 primary outcomes were not (see the Outcome Measures section for more detail on this).

In relation to the first point above, both the process evaluation and the headteacher surveys indicate that Year 6 pupils were actively encouraged to take-up the breakfast club offer during the KS2 testing period. About 95% of intervention and 91% of control schools who responded to the survey provided breakfast for KS2 students for this short period of the year. If the effects of breakfast provision are mostly immediate (affecting students' ability to focus on the test that day, for example) rather than cumulative (allowing students to learn more each day during the year), the high levels of provision in KS2 control schools could also attenuate the impact of the intervention on academic outcomes assessed by these tests.

In relation to the final point above, it might be possible that the intervention improves behaviour and causes teachers to perceive their class more positively, which affects the teacher-assessed outcomes at KS1 (but not the KS2 tests). However, Table 8 indicates that the intervention has similar effects on test and teacher assessment scores for reading and maths at KS2, which are assessed in both ways. This suggests that teachers' judgments are not substantially affected by the presence of a breakfast club and are a reliable measure of attainment in this context.

**Table 7: Primary outcomes—academic attainment**

| Outcome                | Raw means: standardised variables |                         |                |                         | Effect size                              |                         |          |
|------------------------|-----------------------------------|-------------------------|----------------|-------------------------|--|-------------------------|----------|
|                        | Intervention group                |                         | Control group  |                         | n in model<br>(intervention;<br>control) | Effect size<br>(95% CI) | p-value  |
|                        | n<br>(missing)                    | Mean<br>(95% CI)        | n<br>(missing) | Mean<br>(95% CI)        |  |                         |          |
| <b>KS1<br/>maths</b>   | 2,473<br>(42)                     | 0.204<br>(0.167; 0.242) | 2113<br>(29)   | 0.053<br>(0.011; 0.095) | 102<br>(51; 51)                          | 0.149<br>(0.051; 0.248) | 0.003*** |
| <b>KS1<br/>reading</b> | 2,473<br>(42)                     | 0.206<br>(0.169; 0.242) | 2113<br>(29)   | 0.075<br>(0.033; 0.116) | 102<br>(51; 51)                          | 0.104<br>(0.012; 0.196) | 0.025**  |
| <b>KS1<br/>writing</b> | 2,473<br>(42)                     | 0.228<br>(0.191; 0.266) | 2113<br>(29)   | 0.097<br>(0.056; 0.138) | 102<br>(51; 51)                          | 0.138<br>(0.038; 0.239) | 0.006*** |

|                    |               |                          |              |                           |                |                          |       |
|--------------------|---------------|--------------------------|--------------|---------------------------|----------------|--------------------------|-------|
| <b>KS2 reading</b> | 2,050<br>(44) | 0.047<br>(0.006; 0.089)  | 1857<br>(38) | -0.015<br>(-0.061; 0.031) | 98<br>(48; 50) | 0.103<br>(-0.056; 0.262) | 0.198 |
| <b>KS2 maths</b>   | 2,050<br>(44) | 0.008<br>(-0.035; 0.051) | 1857<br>(38) | 0.027<br>(-0.018; 0.072)  | 98<br>(48; 50) | 0.075<br>(-0.060; 0.210) | 0.270 |

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*)

Standard errors clustered at school level.

Outcomes are the standardised fine point scores for the relevant tests. KS1 outcomes reported in the NPD are teacher-assessed, informed in part by the pupil's test results.

Controls for KS1 outcomes: randomisation strata; Foundation Stage attainment; demographics (sex, ever FSM, ethnic group, SEN, EAL); pre-intervention school characteristics (Ofsted rating, IMD rank, urban-rural, class size).

Controls for KS2 outcomes are the same, but control for prior attainment at KS1 rather than for Foundation Stage attainment. This is done through binary variables for categories of average attainment on KS1 tests: fewer than 11 points (did not achieve any level 2); 11–14.99 points (did not achieve grade level); 15–16.99 points (at grade level); 17+ points (above grade level); missing.

Although a significant impact on primary outcomes was only observed for KS1, the secondary outcomes reported in Table 8 suggest that the intervention had a moderate positive impact on KS2 secondary outcomes of writing and English equivalent to about two months' progress. The effect on science, maths, and reading in KS2 was small and not significant, but positive.

**Table 8: Secondary outcomes—academic attainment**

| Outcome                       | Raw means: standardised variables |                          |                |                           | Effect size                                |                          |          |
|-------------------------------|-----------------------------------|--------------------------|----------------|---------------------------|--|--------------------------|----------|
|                               | Intervention group                |                          | Control group  |                           | n in model<br>(intervention<br>n; control) | Effect size<br>(95% CI)  | p-value  |
|                               | n<br>(missing)                    | Mean<br>(95% CI)         | n<br>(missing) | Mean<br>(95% CI)          |  |                          |          |
| <b>KS1 speak &amp; listen</b> | 2,473<br>(42)                     | 0.162<br>(0.124; 0.200)  | 2,113<br>(29)  | -0.012<br>(-0.055; 0.030) | 102 (51; 51)                               | 0.112<br>(-0.023; 0.248) | 0.099*   |
| <b>KS2 reading</b>            | 2,050<br>(44)                     | 0.049<br>(0.008; 0.090)  | 1,857<br>(38)  | -0.005<br>(-0.051; 0.041) | 98 (48; 50)                                | 0.087<br>(-0.024; 0.198) | 0.120    |
| <b>KS2 writing</b>            | 2,050<br>(44)                     | 0.065<br>(0.024; 0.107)  | 1,857<br>(38)  | -0.022<br>(-0.067; 0.024) | 98 (48; 50)                                | 0.151<br>(0.042; 0.259)  | 0.006*** |
| <b>KS2 English</b>            | 2,050<br>(44)                     | 0.048<br>(0.007; 0.089)  | 1,857<br>(38)  | -0.001<br>(-0.047; 0.044) | 98 (48; 50)                                | 0.127<br>(0.033; 0.221)  | 0.008*** |
| <b>KS2 maths</b>              | 2,050<br>(44)                     | 0.010<br>(-0.032; 0.053) | 1,857<br>(38)  | 0.030<br>(-0.016; 0.075)  | 98 (48; 50)                                | 0.045<br>(-0.062; 0.151) | 0.404    |
| <b>KS2 science</b>            | 2,050<br>(44)                     | 0.042<br>(-0.000; 0.084) | 1,857<br>(38)  | 0.001<br>(-0.044; 0.046)  | 98 (48; 50)                                | 0.070<br>(-0.065; 0.205) | 0.301    |

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*)

Standard errors clustered at school level.

Teacher-assessed levels were converted to a point score using DfE guidance and then standardised.

Controls for KS1 outcomes: randomisation strata; Foundation Stage attainment; demographics (sex, ever FSM, ethnic group, SEN, EAL); pre-intervention school characteristics (Ofsted rating, IMD rank, urban-rural, class size).

Controls for KS2 outcomes are the same, but control for bins of the child's average score on the KS1 reading, writing, and maths tests rather than for Foundation Stage attainment.

Table 9 displays the results of an exploratory sensitivity analysis, investigating the extent to which choices about the sample influenced the effect size detected. Four different samples were used:

- the main analysis sample and a full set of controls (analysis sample);
- the main analysis sample controlling only for prior attainment and randomisation strata (prior attainment);
- the subset of the main analysis sample which linked to a pupil survey (survey link); and
- the sample restricted to schools whose breakfast club provision was universal and free. (Before-school delivery was not necessary for schools to be included in this sample even though the model tested by this evaluation was the free, universal, *and* before-school model. This is because relatively few schools met all three criteria, and the first two were judged to be the most important components of fidelity.)

The three main KS1 results are quite robust to variations: similar results are obtained using the first three samples.

KS2 results were somewhat more sensitive to the different samples. Both controlling only for prior attainment and randomisation strata and restricting the sample to pupils who completed a survey caused the effect size to fall. The much smaller number of Year 6 surveys returned indicates that the degree of self-selection into completing the survey is greater at KS2, so it is possible that the intervention was less effective for survey-completers because they were already more motivated to be in class and complete assigned work. This is supported by the higher mean scores in the survey-linked sample relative to the analysis sample. Excluding schools where at least one student mentioned fasting or Ramadan in the follow-up survey (which was the case for Year 6 pupils only) had relatively little effect on the analysis.

For both year groups, restricting the sample to schools whose breakfast club provision was universal and free—the intended model of delivery—dramatically reduces the effect size, and in the case of KS1 removes its significance. This could indicate that schools that varied their provision from the agreed template were able to find a model that better suited their student body and therefore was more effective. However, because these deviations from the agreed model were not random, it is impossible to say whether it is the alternative model per se or some other feature of these deviating schools (such as more inventive leadership, parents more willing to contribute time or money to a breakfast club, or easier integration of educational activities into the alternate breakfast provision models) that is responsible for the difference in results.

**Table 9: Sensitivity analysis—academic outcomes**

| Outcome                | Sample           | Raw means: standardised variables |                         |                |                         | Effect size                              |                          |          |
|------------------------|------------------|-----------------------------------|-------------------------|----------------|-------------------------|--|--------------------------|----------|
|                        |                  | Intervention group                |                         | Control group  |                         | n in model<br>(intervention;<br>control) | Effect size<br>(95% CI)  | p-value  |
|                        |                  | n<br>(miss)                       | Mean<br>(95% CI)        | n<br>(miss)    | Mean<br>(95% CI)        |  |                          |          |
| <b>KS1<br/>maths</b>   | Analysis sample  | 2,473<br>(42)                     | 0.204<br>(0.167; 0.242) | 2,113<br>(29)  | 0.053<br>(0.011; 0.095) | 102 (51; 51)                             | 0.149<br>(0.051; 0.248)  | 0.003*** |
|                        | Prior attainment | 2,473<br>(42)                     | 0.204<br>(0.167; 0.242) | 2,113<br>(29)  | 0.053<br>(0.011; 0.095) | 102 (51; 51)                             | 0.134<br>(0.045; 0.223)  | 0.003*** |
|                        | Survey link      | 2,084<br>(431)                    | 0.238<br>(0.198; 0.277) | 1,643<br>(499) | 0.093<br>(0.047; 0.139) | 89 (46; 43)                              | 0.153<br>(0.052; 0.253)  | 0.003*** |
|                        | Universal & Free | 861<br>(1,654)                    | 0.136<br>(0.073; 0.200) | 2,113<br>(29)  | 0.053<br>(0.011; 0.095) | 68 (17; 51)                              | 0.024<br>(-0.090; 0.138) | 0.669    |
| <b>KS1<br/>reading</b> | Analysis sample  | 2,473<br>(42)                     | 0.206                   | 2,113<br>(29)  | 0.075                   | 102 (51; 51)                             | 0.104                    | 0.025**  |

|                    |                  |                |                            |                |                           |              |                           |          |
|--------------------|------------------|----------------|----------------------------|----------------|---------------------------|--------------|---------------------------|----------|
|                    |                  |                | (0.169; 0.242)             |                | (0.033; 0.116)            |              | (0.012; 0.196)            |          |
|                    | Prior attainment | 2,473<br>(42)  | 0.206<br>(0.169; 0.242)    | 2,113<br>(29)  | 0.075<br>(0.033; 0.116)   | 102 (51; 51) | 0.113<br>(0.030; 0.196)   | 0.007*** |
|                    | Survey link      | 2,084<br>(431) | 0.237<br>(0.198; 0.275)    | 1,643<br>(499) | 0.112<br>(0.066; 0.157)   | 89 (46; 43)  | 0.117<br>(0.022; 0.213)   | 0.014**  |
|                    | Universal & Free | 861<br>(1,654) | 0.150<br>(0.088; 0.213)    | 2,113<br>(29)  | 0.075<br>(0.033; 0.116)   | 68 (17; 51)  | -0.008<br>(-0.114; 0.099) | 0.888    |
| <b>KS1 writing</b> | Analysis sample  | 2,473<br>(42)  | 0.228<br>(0.191; 0.266)    | 2,113<br>(29)  | 0.097<br>(0.056; 0.138)   | 102 (51; 51) | 0.138<br>(0.038; 0.239)   | 0.006*** |
|                    | Prior attainment | 2,473<br>(42)  | 0.228<br>(0.191; 0.266)    | 2,113<br>(29)  | 0.097<br>(0.056; 0.138)   | 102 (51; 51) | 0.124<br>(0.037; 0.210)   | 0.005*** |
|                    | Survey link      | 2,084<br>(431) | 0.260<br>(0.220; 0.300)    | 1,643<br>(499) | 0.142<br>(0.096; 0.188)   | 89 (46; 43)  | 0.127<br>(0.021; 0.233)   | 0.017**  |
|                    | Universal & Free | 861<br>(1,654) | 0.198<br>(0.134; 0.261)    | 2,113<br>(29)  | 0.097<br>(0.056; 0.138)   | 68 (17; 51)  | 0.058<br>(-0.049; 0.166)  | 0.278    |
| <b>KS2 reading</b> | Analysis sample  | 2,050<br>(44)  | 0.047<br>(0.006; 0.089)    | 1,857<br>(38)  | -0.015<br>(-0.061; 0.031) | 98 (48; 50)  | 0.103<br>(-0.056; 0.262)  | 0.198    |
|                    | Prior attainment | 2,050<br>(44)  | 0.047<br>(0.006; 0.089)    | 1,857<br>(38)  | -0.015<br>(-0.061; 0.031) | 98 (48; 50)  | 0.066<br>(-0.072; 0.204)  | 0.340    |
|                    | Survey link      | 1,540<br>(554) | 0.114<br>(0.069; 0.159)    | 1,306<br>(589) | 0.110<br>(0.063; 0.157)   | 84 (41; 43)  | 0.008<br>(-0.138; 0.154)  | 0.913    |
|                    | No Ramadan       | 1,612<br>(482) | 0.076<br>(0.030; 0.123)    | 1,117<br>(778) | 0.017<br>(-0.037; 0.071)  | 75 (40; 35)  | 0.119<br>(-0.050; 0.288)  | 0.160    |
|                    | Universal & Free | 773<br>(1,321) | -0.027<br>(-0.100; 0.045)  | 1,857<br>(38)  | -0.015<br>(-0.061; 0.031) | 66 (16; 50)  | 0.085<br>(-0.104; 0.275)  | 0.368    |
| <b>KS2 maths</b>   | Analysis sample  | 2,050<br>(44)  | 0.008<br>(-0.035; 0.051)   | 1,857<br>(38)  | 0.027<br>(-0.018; 0.072)  | 98 (48; 50)  | 0.075<br>(-0.060; 0.210)  | 0.270    |
|                    | Prior attainment | 2,050<br>(44)  | 0.008<br>(-0.035; 0.051)   | 1,857<br>(38)  | 0.027<br>(-0.018; 0.072)  | 98 (48; 50)  | 0.008<br>(-0.119; 0.135)  | 0.896    |
|                    | Survey link      | 1,540<br>(554) | 0.059<br>(0.011; 0.106)    | 1,306<br>(589) | 0.111<br>(0.062; 0.159)   | 84 (41; 43)  | 0.005<br>(-0.128; 0.137)  | 0.946    |
|                    | No Ramadan       | 1,612<br>(482) | 0.001<br>(-0.047; 0.050)   | 1,117<br>(778) | 0.036<br>(-0.018; 0.090)  | 75 (40; 35)  | 0.075<br>(-0.078; 0.228)  | 0.330    |
|                    | Universal & Free | 773<br>(1,321) | -0.097<br>(-0.168; -0.026) | 1,857<br>(38)  | 0.027<br>(-0.018; 0.072)  | 66 (16; 50)  | 0.035<br>(-0.121; 0.192)  | 0.652    |

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*).

Standard errors clustered at school level. Outcomes are standardised.

Analysis sample: Children with a full set of outcome variables and non-missing demographics.

Prior attainment: Identical to the analysis sample, but controlling only for prior attainment (not demographics or school characteristics).

Survey link: Children in the analysis sample whose NPD record was successfully linked to a pupil survey.

*No Ramadan: Children in the analysis sample in schools where no child mentioned fasting or Ramadan in the pupil survey (Year 6 only).*  
*Universal & Free: Children in the analysis sample whose headteachers reported offering a universal, free breakfast club as part of the intervention.*  
*Controls: randomisation strata; prior attainment; demographics (sex, ever FSM, ethnic group, SEN, EAL); pre-intervention school characteristics (Ofsted rating, IMD rank, urban-rural, class size).*

Overall, this suggests the breakfast club intervention did have an impact on academic attainment, particularly for Year 2 students. These impacts on Year 2 students were strongest in maths but were broadly similar across subject groups, and were generally robust to different choices of sample. The impacts on KS2 tests were more modest, but the intervention's impact on KS2 writing and English was comparable to the effect at KS1. The results for the test-assessed KS2 results were somewhat more sensitive to the choice of sample.

### **Breakfast consumption and hunger**

The theory of change for this intervention suggests that an increase in breakfast consumption and reduction in pupil hunger within the school are necessary conditions for the introduction of a breakfast club to have any positive impact on pupil attainment. This section therefore explores the self-reported breakfast consumption of pupils in intervention and control schools at the start and end of the academic year. It is also interesting to compare consumption and hunger in Year 2 and Year 6 pupils, given the difference in results at KS1 and KS2 seen in the primary analysis.

The baseline level of breakfast consumption was high (91% of pupils who completed the survey reported having breakfast that day), and slightly (but significantly) smaller for children eligible for FSM (90% of whom ate breakfast, compared to 92% of non-FSM students).<sup>15</sup> Given the high baseline breakfast consumption, any increase in consumption due to the intervention is therefore meaningful.

Table 10 shows that, on average, the intervention increased the proportion of pupils eating breakfast *at school* by around 15 percentage points. Breakfast consumption *overall* increased only marginally, suggesting that breakfast consumption moved from home to school for most affected pupils. This is consistent with evidence for Wales (Murphy *et al.*, 2011). Any direct effect from breakfast consumption on pupil attainment is therefore likely to be due to changing the content and context of breakfast, rather than whether or not any breakfast is consumed. For example, while the share of pupils eating breakfast who reported eating food from at least one of the four healthy food groups<sup>16</sup> was significantly higher in control schools at baseline (71% in control schools compared to 68% in intervention schools), in the follow-up survey more pupils in intervention schools reported eating a breakfast containing at least one of the healthy food categories (69% in control schools versus 71% in intervention schools), though this difference was not significant. These changes seem to be driven by improvements in the quality of school breakfasts: in intervention schools, the proportion of students eating breakfast at school who reported eating at least one healthy food increased from 70% at baseline to 86% at follow-up, while the share of healthy breakfasts among intervention-school pupils eating at home remained the same. It should be noted, however, that these results are dependent on pupils' self-reported breakfast intake and cannot reliably be used to assess pupil nutrition.

It is important to note that pupil-reported breakfast consumption may have some measurement error. Self-reported breakfast consumption may not be reliable as some children may feel under pressure to disguise their true consumption, perhaps due to fear of investigation from social services or teachers'

<sup>15</sup> Note that these figures and the figures below describing the type of foods consumed at breakfast were produced from the pupil survey but are not presented in a summary table elsewhere.

<sup>16</sup> The four 'healthy' breakfast categories defined in the Healthy Eating Regulations are milk-based drinks or yoghurts, cereals (not coated or flavoured, either alone or in combination with sugar or chocolate or cocoa powder), fruit and vegetables, and breads with toppings (thinly spread).

views. There may also be particular measurement error among Year 6 students, as the follow-up pupil surveys went out during Ramadan, the Muslim month of fasting. Twenty-three students indicated on their surveys that they had not eaten breakfast for religious reasons, but it is impossible to know conclusively how many students this affects. It is likely to have been higher since pupils were not asked about fasting directly. The timing was common to the intervention and control schools, so measures of the effect of the intervention—which are based on comparison between the two groups—will be more reliable than before/after comparisons for pupils within the same school. However, the lack of balance in ethnicity between intervention and control schools suggests that Ramadan may have depressed breakfast consumption more in the control group than the intervention group, assuming that Asian students are more likely to observe Ramadan, so even these results should be treated with caution. (10.5% of students in the intervention group are of Asian ethnicity, compared to 16.0% of students in the control group.) We have included the results of a specification in which schools where at least one student mentioned Ramadan or fasting were excluded, which attempts to provide the effect of the intervention for schools not affected by the timing of Ramadan; however, this is a crude way to control for the impact of fasting since it relies on students to spontaneously mention their fast and then excludes all students in that school from the analysis (even those who were not fasting).

In addition to the effects of Ramadan, pupils' self-reported hunger may also have some measurement error due to using a question that is potentially ambiguous. Students were asked: 'Did you feel hungry when you started your first lesson, or not?'. Although we have interpreted responses of 'yes' as students being hungry, it is possible that some students responded 'yes' to mean that they were not hungry.

Exploring the differences between Year 2 and Year 6 pupils, the estimated impact of the intervention on eating breakfast at school was much larger for Year 6 pupils (22 percentage points compared to 11 percentage points). This suggests that the differences in academic impacts at KS1 and KS2 are not driven by lower take-up among Year 6 students, although there was also a significant increase in the proportion of Year 6 students that reported being hungry in the intervention group (of five percentage points). This effect was even larger when schools where students mentioned Ramadan were excluded.

The fact that not all pupils ate breakfast following the introduction of the breakfast club suggests that at least some pupils in the intervention group felt that the breakfast club offer was not suitable for them or their family, although this result could be affected by religious or access concerns (for example, where breakfast clubs were not universal and free). The headteacher survey suggests a lower level of take-up in most schools than expected, with fewer than 40% of schools in the intervention group usually serving more than 50 pupils per day.

These shifting breakfast consumption patterns indicate that the intervention is effective in encouraging students who were already eating breakfast to attend the breakfast club, potentially reducing stigma and making it easier to encourage other students who do not eat breakfast at home to attend. Inducing students to eat breakfast at school could also benefit pupils if, for example, breakfast club foods are healthier and more nutritious than the breakfast foods available at home, or if the additional time at school and the activities undertaken during this time facilitates learning later in the day. For example, the case study evidence suggests that staff perceived the social environment of the breakfast club as positive for children's relationships and independence.



**Table 10: Secondary outcomes—breakfast consumption and hunger**

| Outcome (subgroup)                                    | Raw means          |                      |               |                      | Effect size                        |                       |          |
|---|--------------------|----------------------|---------------|----------------------|------------------------------------|-----------------------|----------|
|   | Intervention group |                      | Control group |                      | n in model (intervention; control) | Effect size (95% CI)  | p-value  |
| n (missing)   | Mean (95% CI)      | n (missing)          | Mean (95% CI) |                      |                                    |                       |          |
| <b>Ate breakfast</b>                                  | 1,842 (2,357)      | 0.902 (0.888; 0.915) | 1,531 (1,839) | 0.894 (0.878; 0.909) | 71 (38; 33)                        | 0.032 (0.001; 0.064)  | 0.045**  |
| <b>Ate breakfast (Year 2)</b>                         | 1,167 (1,116)      | 0.921 (0.906; 0.937) | 848 (931)     | 0.923 (0.905; 0.941) | 64 (34; 30)                        | 0.017 (-0.024; 0.058) | 0.419    |
| <b>Ate breakfast (Year 6)</b>                         | 675 (1,241)        | 0.868 (0.843; 0.894) | 683 (908)     | 0.857 (0.830; 0.883) | 53 (24; 29)                        | 0.030 (-0.013; 0.072) | 0.169    |
| <b>Ate breakfast (Year 6 - no Ramadan)</b>            | 509 (1,046)        | 0.882 (0.854; 0.910) | 368 (751)     | 0.864 (0.829; 0.899) | 41 (20; 21)                        | 0.052 (-0.009; 0.112) | 0.095*   |
| <b>Breakfast at school</b>                            | 1,835 (2,364)      | 0.236 (0.217; 0.255) | 1,488 (1,882) | 0.079 (0.065; 0.092) | 70 (38; 32)                        | 0.146 (0.076; 0.216)  | 0.000*** |
| <b>Breakfast at school (Year 2)</b>                   | 1,161 (1,122)      | 0.240 (0.216; 0.265) | 821 (958)     | 0.099 (0.078; 0.119) | 63 (34; 29)                        | 0.106 (0.041; 0.172)  | 0.001*** |
| <b>Breakfast at school (Year 6)</b>                   | 674 (1,242)        | 0.228 (0.197; 0.260) | 667 (924)     | 0.054 (0.037; 0.071) | 52 (24; 28)                        | 0.220 (0.104; 0.337)  | 0.000*** |
| <b>Breakfast at school (Year 6 - no Ramadan)</b>      | 508 (1,047)        | 0.181 (0.148; 0.215) | 352 (767)     | 0.048 (0.026; 0.071) | 40 (20; 20)                        | 0.204 (0.102; 0.307)  | 0.000*** |
| <b>Hungry at start of class</b>                       | 1,602 (2,597)      | 0.280 (0.258; 0.302) | 1,408 (1,962) | 0.249 (0.226; 0.271) | 71 (38; 33)                        | 0.032 (-0.019; 0.083) | 0.219    |
| <b>Hungry at start of class (Year 2)</b>              | 991 (1,292)        | 0.357 (0.327; 0.387) | 762 (1,017)   | 0.366 (0.332; 0.400) | 64 (34; 30)                        | 0.007 (-0.057; 0.072) | 0.823    |
| <b>Hungry at start of class (Year 6)</b>              | 611 (1,305)        | 0.154 (0.125; 0.183) | 631 (960)     | 0.113 (0.088; 0.137) | 52 (24; 28)                        | 0.046 (-0.005; 0.097) | 0.080*   |
| <b>Hungry at start of class (Year 6 - no Ramadan)</b> | 457 (1,098)        | 0.155 (0.122; 0.189) | 333 (786)     | 0.111 (0.077; 0.145) | 40 (20; 20)                        | 0.056 (0.003; 0.109)  | 0.039**  |

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*).

Reported effect sizes are average marginal effects following logistic regression.

Standard errors clustered at school level.

Controls: randomisation strata; average of baseline responses to the three (yes/no) outcome questions<sup>17</sup>; demographics (sex, ever FSM, ethnic group, SEN, EAL); pre-intervention school characteristics (Ofsted rating, IMD rank, urban-rural, class size).

<sup>17</sup> The indicators for 'eating breakfast' and 'eating breakfast at school' were reversed, so that higher values on this index denote worse outcomes (more hunger and/or less breakfast consumption). These baseline measures were

The 'no Ramadan' specification excludes pupils in schools where at least one child mentioned Ramadan or fasting in their survey response.

One potential concern with the pupil survey data is that the definition of the sample (requiring both baseline and follow-up survey responses) could affect the estimated results. Appendix I explores how sensitive the results on hunger and breakfast consumption are to different definitions of the sample. Overall, results are similar across the three specifications, although the effect of the intervention on eating breakfast at school is larger when restricting the analysis to schools with a universal, free breakfast offer.

### Absence and punctuality

Breakfast club provision may be expected to reduce school absences and late arrivals through increasing the incentive to arrive on time (for the breakfast club), reducing the amount parents are responsible for at home (no longer preparing breakfast), and/or improving pupils' health (and thereby reducing absences due to illness or medical appointments). The evidence shown in Table 11 suggests that the intervention led to small reductions in late arrivals and absences, equivalent to around 0.15 fewer late arrivals per pupil per academic year (4.5 per academic year in a class of 30) and 0.88 fewer half-days of absence in the academic year per pupil per academic year (or 26.4 per academic year in a class of 30). For absences, the size of the impact is larger if we consider only authorised absences (which includes absence due to illness), with the intervention causing a significant reduction of 1.37 half-days of absence per pupil per academic year (or 41.1 per academic year in a class of 30). However, *unauthorised* absences were higher in the intervention group, although this result was not statistically significant.

The effect of the intervention on students' presence and punctuality was stronger for Year 2 students, with an effect size for late arrivals around three times as large as that for Year 6 students.

These findings suggest that the effect of the intervention may have been partly mediated by reduced authorised absences, potentially due to fewer days off for illness. Improved punctuality appears to have played less of a role, although the four schools in the process evaluation did perceive a decrease in the proportion of pupils arriving late to class. Another randomised controlled trial on an early breakfast club programme in England also found small improvements in school attendance (Shemilt *et al.*, 2004).

**Table 11: Secondary outcomes—absence and late arrivals**

| Outcome (subgroup)            | Raw means          |                         |               |                         | Estimated difference               |   | p-value |
|-------------------------------|--------------------|-------------------------|---------------|-------------------------|------------------------------------|---|---------|
|                               | Intervention group |                         | Control group |                         | n in model (intervention; control) | Estimated difference (number of half-day sessions) (95% CI) |         |
| n (miss)                      | Mean (95% CI)      | n (miss)                | Mean (95% CI) |                         |                                    |   |         |
| <b>Total absence</b>          | 4,330 (279)        | 12.819 (12.411; 13.227) | 3,755 (282)   | 13.493 (13.033; 13.953) | 106 (53; 53)                       | -0.884 (-1.940; 0.172)                                      | 0.097*  |
| <b>Total absence (Year 2)</b> | 2,435 (80)         | 13.913 (13.350; 14.475) | 2,077 (65)    | 14.390 (13.785; 14.995) | 102 (51; 51)                       | -1.129 (-2.242; -0.015)                                     | 0.044** |

combined into an index in order to create a more holistic baseline measure, particularly in light of concerns about pupil misreporting on individual questions outlined in this section.

|                               |                |                            |                |                            |                 |                            |          |
|-------------------------------|----------------|----------------------------|----------------|----------------------------|-----------------|----------------------------|----------|
| <b>Total absence (Year 6)</b> | 1,895<br>(199) | 11.413 (10.829;<br>11.996) | 1,678<br>(217) | 12.383<br>(11.680; 13.087) | 98<br>(48; 50)  | -0.892<br>(-2.330; 0.546)  | 0.218    |
| <b>Authorised absence</b>     | 4,330<br>(279) | 9.030<br>(8.735; 9.326)    | 3,755<br>(282) | 10.199<br>(9.849; 10.549)  | 106<br>(53; 53) | -1.367<br>(-2.280; -0.454) | 0.003*** |
| <b>Unauthorised absence</b>   | 4,330<br>(279) | 3.788<br>(3.539; 4.038)    | 3,755<br>(282) | 3.294<br>(3.044; 3.544)    | 106<br>(53; 53) | 0.454<br>(-0.235; 1.143)   | 0.191    |
| <b>Late arrivals</b>          | 4,330<br>(279) | 0.236<br>(0.200; 0.272)    | 3,755<br>(282) | 0.352<br>(0.293; 0.410)    | 106<br>(53; 53) | -0.147<br>(-0.337; 0.043)  | 0.126    |
| <b>Late arrivals (Year 2)</b> | 2,435<br>(80)  | 0.287<br>(0.231; 0.344)    | 2,077<br>(65)  | 0.403<br>(0.327; 0.479)    | 102<br>(51; 51) | -0.222<br>(-0.495; 0.051)  | 0.107    |
| <b>Late arrivals (Year 6)</b> | 1,895<br>(199) | 0.170<br>(0.129; 0.211)    | 1,678<br>(217) | 0.288<br>(0.196; 0.379)    | 98<br>(48; 50)  | -0.073<br>(-0.187; 0.040)  | 0.199    |

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*)

Standard errors clustered at school level.

Total absences: The total number of half-day sessions missed during the 2014/15 academic year.

Authorised reasons for absence include illness, medical appointments, religious observances, study leave, traveller absence, agreed family holiday, and pupil exclusion.

Unauthorised reasons for absence include unauthorised family holiday and late arrival.

Late arrivals measure just those unauthorised absences which were caused by the pupil's tardy (but eventual) arrival.

Controls: randomisation strata; the student's prior absence record (from 2013/2014); demographics (sex, ever FSM, ethnic group, SEN, EAL); pre-intervention school characteristics (Ofsted rating, IMD rank, urban-rural, class size).

Note: The effect sizes for total absences for Year 2 and Year 6 are both stronger than the overall effect size. Generally, the overall effect size would be in between the effect sizes for the two groups when run separately. However, that is not always the case because the model accounts for school and pupil covariates that might have a different effect on the outcome for the two subgroups.

## Behaviour and concentration

In addition to increasing the amount of time pupils spend in the classroom, the intervention might have improved the quality of that learning time. These effects could operate both at the individual level (if a student is better able to concentrate because they are not hungry, for example) and at the classroom level (if a more pleasant learning environment helps students indirectly by reducing distractions and disruptions).

These mechanisms were assessed through an online survey of classroom teachers, asking about students' disruptiveness and focus. Using factor analysis, these questions were grouped into two indices measuring reported behaviour and concentration. This data-driven approach attempts to identify and analyse the common factors driving the response to each question, incorporating responses for all relevant questions into a single index. However, these indices and the estimates of the intervention's impact on them are necessarily an approximation of the estimate that would result from using each question as an outcome separately.

The behaviour index incorporates teachers' perception of their classroom's learning environment, lost time due to disruption, disruptive noise, and lost time waiting for students to be quiet. The concentration index is based on teachers' assessment of the share of their class that was ready to learn first thing in the morning and after lunch, and the share of students that had good and poor concentration in the mornings. The teacher survey containing these questions is given in Appendix D.

As Table 12 shows, the breakfast club intervention had a large positive effect on teacher-perceived student behaviour and concentration in the classroom. Classes in intervention schools scored 0.48

standard deviations higher on the behaviour index. The results for pupil concentration were even larger, with classes in intervention schools scoring 0.64 standard deviations higher. For a class with an average score on the concentration index at baseline, this is roughly equivalent to moving from the 50th to the 74th percentile. An equivalent move in the each of the underlying measures of concentration would be (separately):

- increasing the share of the class with good concentration in the mornings from 61–80% to 81–90%;
- no change in the share of the class with poor concentration in the mornings (as the 50th and 74th percentile in the distribution of this outcome are equal);
- increasing the share of the class ready to learn at the start of the day from 61–80% to 81–90%; or
- increasing the share of the class ready to learn after lunch from 61–80% to 81–90%.

These findings are consistent with findings from the school case studies, where respondents felt that children were more settled, less disruptive, and better able to concentrate as a result of eating breakfast at school, and also with evidence from an earlier randomised controlled trial in England (Shemilt *et al.*, 2004).

Table 12 also breaks down these results by year group. Effect sizes are higher for Year 2 students for both behaviour and concentration. Most strikingly, the intervention increased Year 2 students' scores on the concentration index by almost 0.7 standard deviations. The particularly strong effects for Year 2 students on these drivers of effective learning could partially explain the much stronger effects on Key Stage 1 attainment relative to Key Stage 2, although it is surprising that the positive reports from Year 6 teachers did not feed through to more positive test results or teacher assessments. One explanation is that teachers' survey responses may have been influenced by involvement with the intervention, so that only *perceived* behaviour and concentration increased for both age groups. It is also possible that actual behaviour and concentration increased for both groups, but this did not translate through into improved results on the KS2 primary outcome measures.

**Table 12: Secondary outcomes: classroom behaviour and concentration**

| Outcome (subgroup)                        | Raw means: standardised variables |                       |               |                         | Effect size                        |                       |          |
|---|-----------------------------------|-----------------------|---------------|-------------------------|------------------------------------|-----------------------|----------|
|   | Intervention group                |                       | Control group |                         | n in model (intervention; control) | Effect size (95% CI)  | p-value  |
|   | n (missing)                       | Mean (95% CI)         | n (missing)   | Mean (95% CI)           |                                    |                       |          |
| <b>Classroom behaviour index</b>          | 120 (51)                          | 0.216 (0.031; 0.400)  | 114 (26)      | -0.231 (-0.404; -0.059) | 86 (44; 42)                        | 0.478 (0.166; 0.790)  | 0.002*** |
| <b>Classroom behaviour index (Year 2)</b> | 66 (24)                           | 0.269 (0.045; 0.492)  | 55 (13)       | -0.375 (-0.596; -0.153) | 74 (39; 35)                        | 0.494 (0.151; 0.837)  | 0.004*** |
| <b>Classroom behaviour index (Year 6)</b> | 54 (27)                           | 0.151 (-0.163; 0.465) | 59 (13)       | -0.097 (-0.360; 0.166)  | 76 (36; 40)                        | 0.440 (-0.028; 0.907) | 0.061*   |
| <b>Classroom concentration index</b>      | 120 (51)                          | 0.307 (0.150; 0.464)  | 114 (26)      | -0.345 (-0.536; -0.155) | 86 (44; 42)                        | 0.641 (0.329; 0.954)  | 0.000*** |

|   |            |                          |            |                            |                |                         |          |
|---|------------|--------------------------|------------|----------------------------|----------------|-------------------------|----------|
| <b>Classroom concentration index (Year 2)</b> | 66<br>(24) | 0.349<br>(0.155; 0.542)  | 55<br>(13) | -0.431<br>(-0.676; -0.185) | 74<br>(39; 35) | 0.699<br>(0.276; 1.122) | 0.001*** |
| <b>Classroom concentration index (Year 6)</b> | 54<br>(27) | 0.256<br>(-0.007; 0.520) | 59<br>(13) | -0.265<br>(-0.560; 0.029)  | 76<br>(36; 40) | 0.565 (0.122;<br>1.009) | 0.011**  |

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*)

Standard errors clustered at school level.

Outcome variables are standardised averages of the questions from the teacher survey, constructed using maximum likelihood factor analysis. The behaviour index depends on four questions about waiting for quiet, pleasant learning environment, lost time, and disruptive noise. The concentration index is based on four questions about good and poor concentration in the morning and readiness to learn at the start of the day and after lunch. The behaviour and concentration indices are standardised to have a mean of zero and a standard deviation of one. A negative mean therefore indicates that the average classroom in the group performed worse on that index than the average classroom in the total sample (which would have a score of zero).

The ordinal reliability alpha scores (Gadernann, Guhn, and Zumbo, 2012) are 0.8914 for the behaviour index and 0.8884 for the concentration index.

Controls: randomisation strata; the teacher's average response to behaviour and concentration questions at baseline; pre-intervention school characteristics (Ofsted rating, IMD rank, urban-rural, class size).

### Body Mass Index

Finally, Table 13 outlines the impact of the Magic Breakfast intervention on Year 6 pupils' health, as measured by their Body Mass Index. This information was made available at the school level by the National Child Measurement Programme (NCMP) and is unfortunately not available for Year 2 pupils; only Year 6 and Reception pupils are measured each year in the national programme. This measure of health is imperfect, as BMI may not adjust quickly to changes in breakfast consumption, but it is the best available source of information. One further caveat is that measurement happens throughout the school year, and so in some cases may have occurred before the intervention began, but the timing should be balanced across intervention and control schools as a result of randomisation.

Students in intervention schools were less likely to be recorded as 'healthy weight' and had higher BMIs, on average, although these results are not statistically significant. The direction of the effects suggests that the Magic Breakfast intervention did not improve the health of pupils in the intervention schools as measured by the proportion of pupils who were of 'healthy weight'. These results could imply that schools implementing a breakfast club should be mindful that students could over-consume, although the results here are not conclusive enough to give a strong recommendation.

**Table 13: Secondary outcomes: Body Mass Index and healthy weight**

| Outcome                 | Raw means          |                            |                  |                            | Effect size                              |                           |         |
|-------------------------|--------------------|----------------------------|------------------|----------------------------|--|---------------------------|---------|
|                         | Intervention group |                            | Control group    |                            | n in model<br>(intervention;<br>control) | Effect size<br>(95% CI)   | p-value |
| n<br>(missing)          | Mean<br>(95% CI)   | n<br>(missing)             | Mean<br>(95% CI) |                            |  |                           |         |
| <b>BMI z-score</b>      | 41<br>(12)         | 0.728<br>(0.645; 0.812)    | 38<br>(15)       | 0.588<br>(0.507; 0.669)    | 79<br>(41; 38)                           | 0.112<br>(-0.055; 0.279)  | 0.179   |
| <b>% healthy weight</b> | 41<br>(12)         | 57.351<br>(54.354; 60.348) | 38<br>(15)       | 60.560<br>(57.722; 63.398) | 79<br>(41; 38)                           | -2.689<br>(-8.721; 3.343) | 0.373   |

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*)

Analytical weights are applied to account for the different number of pupils included in each school-level average.

Standard errors are robust to heteroscedasticity.

Body Mass Index z-score: the z-score of the average BMI of Year 6 students in the school, standardised based on the British 1990 growth reference population.

% healthy weight: the share of the Year 6 pupils in the school who are at a healthy weight (defined as between the 2nd and 85th centiles of the 1990 Growth Reference Tables).

Controls: randomisation strata; the outcome variable measured on the 2013/2014 Year 6 cohort; pre-intervention school characteristics (Ofsted rating, IMD rank, urban-rural, class size).

## Summary

Overall, the Magic Breakfast intervention had a positive effect, especially for Year 2 students and for teachers' assessment of their classroom learning environment. Secondary attainment outcomes also improved for Year 6 students. In addition to large improvements in teacher-reported pupil concentration and behaviour, many academic outcomes also moderately improved. Breakfast club provision also improved student attendance but seems to have done little to reduce late arrivals. Students in intervention schools are more likely to eat breakfast at school as a result of the intervention, although not to eat breakfast overall. This suggests that the improved outcomes are due to the content of the breakfast (the volume or type of food consumed) or the context of the breakfast (the social environment of the breakfast club, more time to relax before lessons start, the nature of other activities taking place alongside breakfast).

There were large improvements in the perceived levels of concentration and behaviour in the classroom. This means that the classroom environment improved for teachers, and presumably pupils, which may have been beneficial for learning. This is a potential mechanism by which even pupils who did not change their breakfast consumption patterns may have benefitted from the intervention.

The results showed a larger effect size for KS1 pupils on primary outcomes. This is in contrast to the evaluation of the universal free school meals pilot (Brown *et al.*, 2012) where the effects were largest for Year 6 pupils. Possible reasons for this include:

- the fact that around 90% of control schools who responded to our follow-up survey offered large scale breakfast provision to Year 6 pupils during the week of KS2 tests;
- higher breakfast consumption prior to the intervention for Year 6 pupils (as reported in the baseline pupil survey);
- Year 6 pupils being more likely to fast during Ramadan, which occurred during the school year; and
- the fact that the KS1 primary outcomes were teacher assessed and the KS2 primary outcomes were not (although the comparable data we have suggests that teacher assessment did not introduce any substantial bias).

Ramadan is more likely to affect older pupils as in most traditions fasting is compulsory for those who have reached puberty, although some children may decide to begin at a younger age. If fasting influences pupil attainment, the impact of Ramadan should therefore be stronger in KS2 than KS1. The results remain similar even when schools with at least one pupil mentioning fasting or Ramadan are excluded from the sample, however. These schools had higher reported pupil hunger, but this didn't affect the estimated difference between intervention schools and control schools, on average.

The most noticeable difference between KS1 and KS2 pupils are the conditions in control schools, where over 90% of Year 6 pupils were encouraged to take up a breakfast club offer in the week of testing. This would not affect the teachers' reports of concentration and behaviour throughout the year, but may have affected performance on the tests. If the free breakfasts offered in control schools during the testing week have a positive impact on control group pupils' attainment, this could attenuate the effect size of the intervention for academic outcomes for KS2 pupils. The sample size for the older age group is also smaller due to the recruitment of a number of infant schools, which reduces the power of the intervention to detect a significant effect at KS2.

## Subgroup analysis

This section shows the effect of the Magic Breakfast intervention for subgroups of pupils, as specified in the evaluation protocol. Of particular interest is the impact for pupils that have ever been eligible for

free school meals, a group important for both the EEF and Magic Breakfast. Other subgroups of interest are boys versus girls, students with low prior attainment versus those who achieved expected or better performance in their last assessment, and students who did not report having breakfast in the baseline surveys compared to those who did.

Table 14 presents the results of these subgroup analyses for two summary outcome measures, average points on the KS1 assessments (reading, writing, and maths) and the KS2 tests (reading and maths). Since the outcomes are standardised to have mean zero and standard deviation one, a negative score indicates that the intervention had a negative effect for the children in that subgroup, relative to children in that subgroup whose schools were in the control group. Conversely, positive scores indicate a positive effect of the intervention for that subgroup. The final column of Table 14 gives the p-value of a test for whether the effect of the intervention (relative to the relevant group of children in control schools) differs significantly by subgroup.

The overall effect for all KS1 pupils and all KS2 pupils on these summary outcome measures is shown in the first two rows of Table 14. These first two rows show that, for each Key Stage group, the summary outcomes measures are not substantively different to the effect sizes for each of the primary outcome variables individually.

The subgroup analysis is conducted by repeating the analysis presented in the first two rows on specific subgroups. The pupils in that subgroup from intervention schools are compared to pupils in that subgroup in control schools. For example, in the 'KS1: FSM' row, pupils eligible for FSM in KS1 in intervention schools are compared with pupils eligible for FSM in KS1 in control schools. Repeating the analysis on separate samples also allows the effect of prior attainment, pupil demographics, and school characteristics included in the model to differ between the two subgroups. For example, the relationship between prior attainment and current attainment may be stronger for a particular subgroup. These factors mean that the overall effect size for KS1 average points (shown in row 1) is not necessarily an average of the two effect sizes for a given pair of subgroups, such as FSM and non-FSM. Repeating the analysis on separate samples for subgroup analysis gives an accurate picture of how attainment is affected by the intervention on each particular group. Comparing the results from subgroup analysis with the overall effect is informative but not the most appropriate way to test for differences in the effect size across groups.

### **Pupil disadvantage**

The FSM subgroup analysis indicates that the intervention was significantly more effective for non-FSM students than FSM students, both at KS1 and KS2, although the effects were positive for both groups.

The effect sizes for both the KS1 FSM subgroup and the KS1 non-FSM subgroup are both larger than the overall effect size for KS1. At first glance, this appears unintuitive: we might expect the overall KS1 effect size to be an average of the two subgroup effect sizes. There are several reasons why this is not the case. First, since the FSM subgroup is larger than the non-FSM group, it dominates the total sample used to calculate the overall effect size. Second, as outlined above, the subgroup analysis changes the comparison group from *all* pupils in control schools to pupils in control schools with the same FSM status. The overall analysis includes the effect for both FSM and non-FSM students (relative to FSM and non-FSM students in control schools), while the subgroup analysis compares the effect within these subgroups only. Finally, the subgroup analysis allows the influence of prior attainment, pupil demographics, and school characteristics on the outcomes to differ between the two groups.

From this subgroup analysis we conclude that the effect for non-FSM students is stronger than for FSM students at KS1 and KS2, although the provision of a breakfast club had a positive effect for both groups. There are several possible explanations for this: students from non-FSM families could be more likely to take up the breakfast club offer; students from non-FSM families could benefit more from positive spillovers in the classroom; or Year 6 students from marginal, non-FSM families could benefit

from the free breakfast offer while many of the potential benefits to FSM students have already been realised through the free lunchtime meals to which they are already entitled. (This explanation would not apply at Year 2 since all Year 2 students now receive free schools meals, regardless of whether they meet FSM eligibility criteria.)

Table 15 provides evidence against the first hypothesis, as the intervention increased the probability of eating breakfast at school more for FSM students (by 17 percentage points compared to around ten percentage points for non-FSM students). However, overall breakfast consumption increased for non-FSM students, but not for FSM students, which suggests that any gains for FSM students would be derived from the composition and context of the breakfast club compared to eating breakfast outside school, or indirect effects through other pupils. Table 13 provides some support for the hypothesis that non-FSM students may also suffer from hunger, as the proportions of FSM and non-FSM pupils that report being hungry at the start of their first lesson are roughly similar (and relatively high, at 27% and 29% respectively).

### Gender

Table 14 indicates that the intervention did not have significantly different effects on academic outcomes by gender.

### Prior attainment

At Key Stage 2, the intervention was less effective for students with low prior attainment. This could suggest that pupils who were at or above expected grade level were better-placed to take advantage of the improved classroom learning environment.

### Prior breakfast consumption

For KS1 students the intervention was also significantly more effective for pupils who had eaten breakfast at the baseline, supporting the hypothesis that much of the effect of the intervention came through the specific context of eating a nutritious breakfast among peers and staff at school, rather than consumption of breakfast itself.

**Table 14: Primary outcomes—subgroup analysis**

| Outcome: subgroup        | Raw means: standardised variables |                        |               |                         | n in model (intervention; control) | Effect size           |           | Model Difference |
|--------------------------|-----------------------------------|------------------------|---------------|-------------------------|------------------------------------|-----------------------|-----------|------------------|
|                          | Intervention group                |                        | Control group |                         |                                    | Effect size (95% CI)  | p-value   | p-value          |
|                          | n (miss)                          | Mean (95% CI)          | N (miss)      | Mean (95% CI)           |                                    |                       |           |                  |
| <b>KS1: whole sample</b> | 2,473 (42)                        | 0.088 (0.050; 0.127)   | 2,113 (29)    | -0.064 (-0.107; -0.021) | 102 (51; 51)                       | 0.144 (0.045; 0.242)  | 0.004* ** | N/A              |
| <b>KS2: whole sample</b> | 2,050 (44)                        | 0.030 (-0.013; 0.072)  | 1,857 (38)    | 0.008 (-0.037; 0.053)   | 98 (48; 50)                        | 0.095 (-0.052; 0.242) | 0.198     |                  |
| <b>KS1: FSM</b>          | 1,242 (15)                        | -0.039 (-0.093; 0.014) | 1,116 (6)     | -0.209 (-0.267; 0.151)  | 101 (51; 50)                       | 0.153 (0.068; 0.237)  | 0.000* ** | 0.083*           |
| <b>KS1: non-FSM</b>      | 1,231 (10)                        | 0.217 (0.163; 0.271)   | 997 (7)       | 0.098 (0.036; 0.160)    | 102 (51; 51)                       | 0.246 (0.152; 0.341)  | 0.000* ** |                  |



|                                   |               |                            |               |                            |                 |                           |              |          |
|-----------------------------------|---------------|----------------------------|---------------|----------------------------|-----------------|---------------------------|--------------|----------|
| <b>KS2: FSM</b>                   | 1,206<br>(15) | -0.052<br>(-0.105; 0.002)  | 1,158<br>(19) | -0.027<br>(-0.081; 0.026)  | 98<br>(48; 50)  | 0.037<br>(-0.048; 0.122)  | 0.394        | 0.001*** |
| <b>KS2: non-FSM</b>               | 844<br>(11)   | 0.146<br>(0.079; 0.213)    | 699<br>(6)    | 0.067<br>(-0.014; 0.148)   | 98<br>(48; 50)  | 0.269<br>(0.150; 0.387)   | 0.000*<br>** |          |
| <b>KS1: male</b>                  | 1,270<br>(26) | -0.045<br>(-0.102; 0.011)  | 1,103<br>(18) | -0.172<br>(-0.233; -0.111) | 102<br>(51; 51) | 0.145<br>(0.053; 0.237)   | 0.002*<br>** | 0.199    |
| <b>KS1: female</b>                | 1,203<br>(16) | 0.229<br>(0.179; 0.280)    | 1,010<br>(11) | 0.054<br>(-0.006; 0.114)   | 102<br>(51; 51) | 0.233<br>(0.147; 0.320)   | 0.000*<br>** |          |
| <b>KS2: male</b>                  | 1,012<br>(18) | 0.037<br>(-0.026; 0.100)   | 927 (21)      | -0.030<br>(-0.096; 0.037)  | 98<br>(48; 50)  | 0.191<br>(0.088; 0.294)   | 0.000*<br>** | 0.117    |
| <b>KS2: female</b>                | 1,038<br>(26) | 0.022<br>(-0.034; 0.078)   | 930 (17)      | 0.046<br>(-0.015; 0.107)   | 97<br>(48; 49)  | 0.063<br>(-0.031; 0.157)  | 0.189        |          |
| <b>KS1: low prior attainment</b>  | 1,298<br>(19) | -0.345<br>(-0.394; -0.296) | 1,162<br>(12) | -0.462<br>(-0.516; -0.408) | 101<br>(51; 50) | 0.121<br>(0.040; 0.201)   | 0.003*<br>** | 0.791    |
| <b>KS1: high prior attainment</b> | 1,051<br>(8)  | 0.694<br>(0.655; 0.733)    | 834<br>(4)    | 0.602<br>(0.558; 0.647)    | 102<br>(51; 51) | 0.103<br>(0.028; 0.177)   | 0.007*<br>** |          |
| <b>KS2: low prior attainment</b>  | 633<br>(10)   | -0.645<br>(-0.719; -0.572) | 620 (11)      | -0.486<br>(-0.562; -0.411) | 97<br>(48; 49)  | -0.026<br>(-0.148; 0.097) | 0.681        | 0.072*   |
| <b>KS2: high prior attainment</b> | 1,417<br>(34) | 0.331<br>(0.288; 0.374)    | 1,237<br>(27) | 0.256<br>(0.205; 0.307)    | 98<br>(48; 50)  | 0.186<br>(0.107; 0.265)   | 0.000*<br>** |          |
| <b>KS1: no prior breakfast</b>    | 160<br>(4)    | -0.209<br>(-0.344; -0.075) | 104<br>(0)    | -0.097<br>(-0.299; 0.105)  | 74<br>(42; 32)  | -0.072<br>(-0.346; 0.201) | 0.601        | 0.078*   |
| <b>KS1: prior breakfast</b>       | 1,719<br>(10) | 0.172<br>(0.128; 0.215)    | 1,207 (7)     | 0.041<br>(-0.013; 0.095)   | 84<br>(46; 38)  | 0.148<br>(0.073; 0.223)   | 0.000*<br>** |          |
| <b>KS2: no prior breakfast</b>    | 125<br>(1)    | -0.158<br>(-0.316; -0.000) | 108<br>(1)    | -0.193<br>(-0.388; 0.002)  | 67<br>(38; 29)  | -0.155<br>(-0.425; 0.115) | 0.258        | 0.362    |
| <b>KS2: prior breakfast</b>       | 1,292<br>(14) | 0.118<br>(0.068; 0.168)    | 934 (10)      | 0.152 (0.098;<br>0.206)    | 78<br>(41; 37)  | -0.024<br>(-0.108; 0.061) | 0.578        |          |

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*)

Standard errors clustered at school level.

Outcome variables are standardised averages of the relevant scores (reading, writing, and maths teacher assessments for KS1; reading and maths test scores for KS2).

For each set of subgroups, the main model has been re-estimated separately for pupils in and out of the subgroup. The final column indicates whether the effect size is significantly different between the two groups.

The low attainment subgroup comprises KS1 students who did not attain a good level of development at Foundation Stage, and KS2 students who received a Level 1 (below grade level) in any KS1 subject (reading, writing, maths, speaking and listening, and science).

Controls for KS1 outcomes: randomisation strata; Foundation Stage attainment; demographics (sex, ever FSM, ethnic group, SEN, EAL); pre-intervention school characteristics (Ofsted rating, IMD rank, urban-rural, class size).

Controls for KS2 outcomes are the same, but control for bins of the child's average score on the KS1 reading, writing, and maths tests rather than for Foundation Stage attainment.

**Subgroup analysis for consumption and hunger outcomes**

As shown in Table 15, the intervention had negligible impact on breakfast consumption both for students who had eaten breakfast prior to the start of the breakfast club, and those who had not. Given the small number of children who did not eat breakfast at baseline, this reinforces the finding that the majority of breakfast club take-up came from students switching from a home to a school breakfast (or possibly having something to eat at school in addition to their breakfast at home). The intervention was also significantly more likely to increase school breakfast consumption for those who had not eaten breakfast at baseline and for FSM students. However, FSM students were significantly *less* likely to eat any breakfast as a result of the intervention than their non-FSM peers.

It is interesting that, although the intervention appears to be having an impact by increasing consumption of school breakfast, and although the impact on eating school breakfasts is significantly higher for FSM students, the impact on attainment is significantly lower for FSM students.

**Table 15: Breakfast consumption and hunger—subgroup analysis**

| Outcome: subgroup                        | Raw means          |                      |               |                      | Effect size                        |                        |          | Model Difference |
|--|--------------------|----------------------|---------------|----------------------|------------------------------------|------------------------|----------|------------------|
|  | Intervention group |                      | Control group |                      | n in model (intervention; control) | Effect size (95% CI)   | p-value  | p-value          |
| n (miss)                                 | Mean (95% CI)      | N (miss)             | Mean (95% CI) |                      |                                    |                        |          |                  |
| <b>Ate breakfast: FSM</b>                | 957 (1,039)        | 0.877 (0.856; 0.898) | 911 (841)     | 0.892 (0.872; 0.913) | 71 (38; 33)                        | -0.007 (-0.049; 0.034) | 0.739    | 0.001***         |
| <b>Ate breakfast: non-FSM</b>            | 885 (778)          | 0.929 (0.912; 0.946) | 620 (597)     | 0.895 (0.871; 0.919) | 70 (37; 33)                        | 0.099 (0.049; 0.149)   | 0.000*** |                  |
| <b>School breakfast: FSM</b>             | 952 (1,044)        | 0.251 (0.223; 0.279) | 882 (870)     | 0.075 (0.057; 0.092) | 70 (38; 32)                        | 0.170 (0.095; 0.246)   | 0.000*** | 0.036**          |
| <b>School breakfast: non-FSM</b>         | 883 (780)          | 0.220 (0.192; 0.247) | 606 (611)     | 0.084 (0.062; 0.106) | 69 (37; 32)                        | 0.103 (0.027; 0.179)   | 0.008*** |                  |
| <b>Hungry AM: FSM</b>                    | 836 (1,160)        | 0.273 (0.242; 0.303) | 828 (924)     | 0.238 (0.209; 0.267) | 71 (38; 33)                        | 0.043 (-0.020; 0.106)  | 0.180    | 0.513            |
| <b>Hungry AM: non-FSM</b>                | 766 (897)          | 0.287 (0.255; 0.319) | 580 (637)     | 0.264 (0.228; 0.300) | 70 (37; 33)                        | 0.016 (-0.056; 0.087)  | 0.665    |                  |
| <b>Ate breakfast: no prior breakfast</b> | 160 (155)          | 0.700 (0.628; 0.772) | 130 (100)     | 0.669 (0.587; 0.751) | 62 (34; 28)                        | 0.030 (-0.077; 0.137)  | 0.580    | 0.991            |
| <b>Ate breakfast: prior breakfast</b>    | 1,678 (1533)       | 0.921 (0.908; 0.934) | 1,395 (931)   | 0.914 (0.899; 0.929) | 71 (38; 33)                        | 0.029 (-0.000; 0.059)  | 0.051*   |                  |

|  |                 |                         |                  |                         |             |                          |          |          |
|--|-----------------|-------------------------|------------------|-------------------------|-------------|--------------------------|----------|----------|
| <b>School breakfast: no prior b'fast</b> | 150<br>(165)    | 0.387<br>(0.308; 0.466) | 121<br>(109)     | 0.058<br>(0.016; 0.100) | 60 (32; 28) | 0.331<br>(0.191; 0.472)  | 0.000*** | 0.000*** |
| <b>School breakfast: prior breakfast</b> | 1,671<br>(1540) | 0.224<br>(0.204; 0.244) | 1,352<br>(974)   | 0.081<br>(0.067; 0.096) | 70 (38; 32) | 0.133<br>(0.067; 0.200)  | 0.000*** |          |
| <b>Hungry AM: no prior breakfast</b>     | 132<br>(183)    | 0.356<br>(0.273; 0.439) | 118<br>(112)     | 0.305<br>(0.221; 0.389) | 61 (33; 28) | 0.102<br>(-0.036; 0.240) | 0.146    | 0.275    |
| <b>Hungry AM: prior breakfast</b>        | 1,467<br>(1744) | 0.273<br>(0.251; 0.296) | 1,284<br>(1,042) | 0.244<br>(0.220; 0.267) | 71 (38; 33) | 0.029<br>(-0.022; 0.079) | 0.264    |          |

Stars for statistical significance:  $p \leq 0.10$  (\*),  $p \leq 0.05$  (\*\*),  $p \leq 0.01$  (\*\*\*).

Reported effect sizes are average marginal effects following logistic regression.

Standard errors clustered at school level.

For each set of subgroups, the main model has been re-estimated separately for pupils in and out of the subgroup. The final column indicates whether the effect size is significantly different between the two groups.

The 'no prior breakfast' subgroup comprises students who reported not having eaten breakfast on their baseline survey.

Controls: randomisation strata; average of baseline responses to the three outcome questions; demographics (sex, ever FSM, ethnic group, SEN, EAL); pre-intervention school characteristics (Ofsted rating, IMD rank, urban-rural, class size).

## Cost

The schools in this intervention received support from Magic Breakfast to deliver a breakfast club. Magic Breakfast provided as much food as required free of charge to the school, support with the logistics of setting up a breakfast club, and a £300 grant to defray up-front costs. Schools were responsible for any additional costs, including capital costs over £300 (for example, purchasing a new freezer), sundries for breakfast such as milk, spreads, and juice above a certain limit, and any supervision or other staff time costs. In line with EEF guidance, the costs presented here are the full non-labour costs of the intervention, including the regular retail cost of food and up-front costs. A school deciding to establish a breakfast club could reduce its per-pupil costs below those estimated here, however, by joining a free or subsidised breakfast club scheme with Magic Breakfast (who offer a membership scheme) or another charity.

Costs are presented as annual cost per pupil at the school, averaged over three years, to give a reasonable picture of the longer-term investment required by schools. Costs are presented for the school breakfast clubs that operated in the trial, which typically had relatively low take-up. The costs presented here therefore do not reflect the cost of providing breakfast to *all* children. A school with higher take-up than the schools in the trial would incur higher costs.

Any revenue from pupils being charged for the breakfast club (in the schools that imposed a charge) has not been deducted from school costs as this was not in line with the delivery model tested.

Cost data was collected from headteachers in the intervention schools and from Magic Breakfast records of food provided. Headteachers were asked what they had actually spent in a number of categories to set up and run the breakfast club.

Cost data was converted to figures for a full school year by:

- using the cost given when the headteacher indicated it was an annual figure;
- multiplying the cost given by 8.5 if given as a monthly figure;
- multiplying the cost given by 38 if given as a weekly figure; or

- multiplying the cost given by 190 if given in days.

In cases where the headteacher did not indicate the period that the cost pertains to, the period was imputed based on the period used in the headteacher’s other responses. If the period was still unclear, ranges of other schools’ responses and the implied annual costs were used to determine the most plausible period based on the school’s total cost/total number of supervisory hours.

The cost figures presented here exclude one major outlier, the reported spending of £20,000 on staff recruitment and training for the breakfast club in one school. This is more than three times as large as the next-highest expenditure in this category and also far exceeded that school’s reported spending on staff time, so it was concluded that this was likely to be an erroneous report and the data point was recoded to missing.

Schools used different approaches to remunerate staff for supervising the breakfast club that depended on the availability of volunteers (used by ten schools responding to the headteacher survey) and breakfast club arrangements (for example, before school or during school). Costs of staff time have *not* been included in the ‘annual cost per pupil at the school’ figure, in line with EEF guidance, and instead are reported in units of hours, with hours for paid and unpaid staff presented separately. The intention is that this will help schools to calculate the actual financial cost of the intervention to them, based on whether or not they would pay for extra staff, move staff time from elsewhere, or recruit volunteers. The process evaluation provides further information on the different approaches used by schools to supervise the breakfast club.

The average annual cost per pupil over three years is £11.86 for each pupil attending the school. This is calculated from ongoing and fixed costs in the first year of the intervention, as shown in Table 16. It is important to note that this cost excludes the cost of staff time, which was a significant cost that was incurred by two-thirds of the intervention schools who responded to a survey on their expenditure.<sup>18</sup>

It is important to stress that £11.86 is the average annual cost *per pupil enrolled in the school*. Per-pupil cost was calculated by dividing the average cost per school by the average pupil count for the relevant years (in schools that responded to the survey). If take-up were higher, the average cost per school would increase (more breakfasts being eaten) but the average pupil count for the relevant years would stay the same. So the cost per pupil per year figure of £11.86 would be higher in a school that had higher take-up than the schools in this trial. The figure we present is the likely cost for a similar school that has similar take-up and which could therefore expect a similar impact.

**Table 16: Cumulative and average cost per pupil**

| Number of years using programme | Fixed cost per pupil (£) | Ongoing cost per pupil (£) | Cumulative cost per pupil (£) | Average annual cost per pupil (£) |
|---------------------------------|--------------------------|----------------------------|-------------------------------|-----------------------------------|
| <b>1 year</b>                   | £3.31                    | £10.75                     | £14.07                        | £14.07                            |
| <b>2 years</b>                  | -                        | £10.75                     | £24.82                        | £12.41                            |
| <b>3 years</b>                  | -                        | £10.75                     | £35.57                        | £11.86                            |

*Fixed costs include spending on furniture, the physical environment (e.g. painting a room), catering facilities (e.g. a freezer), and resources (e.g. toys and games, cutlery). Ongoing costs include spending on staff time, food and drinks (from Magic Breakfast and additional items purchased by schools), and other purchases. Costs per pupil are calculated as the average school-level cost divided by the average number of pupils in the intervention schools that responded to the survey or had Magic Breakfast costs reported.*

<sup>18</sup> This survey was part of the follow-up survey to headteachers where 42 of the 53 intervention schools responded.

As Table 16 shows, the majority of this cost—£10.75 per pupil per year—can be classified as an ongoing expenditure for food. Although some schools ran the breakfast club without incurring any costs for staff time, the average per-pupil expenditure on staff time was £5.46 (not shown in the following tables), which would also be an ongoing cost. Set-up costs were concentrated on catering facilities and staff recruitment and training, though for those schools that incurred costs to improve their physical environment (such as buying new curtains or painting a room) this formed a significant part of their up-front expenditure. These figures capture schools’ estimated expenditure during the year. If parents or local businesses contributed to the set-up or running of the breakfast club, it is possible that these donations will not be captured in the cost data presented here.

Table 17 demonstrates that per-pupil costs varied depending on whether the school had previously offered any ad-hoc breakfast provision. For example, among schools that spent money on catering facilities (for example buying a freezer)—schools that had never previously offered any breakfast—paid more than twice as much on average per pupil.

**Table 17: Average cost per type of cost over one academic year**

| Cost type   | Number of schools incurring | Average per-pupil cost for incurrers | Average per-pupil cost across survey respondents with previous provision | Average per-pupil cost across all survey respondents with no previous provision | Average per-pupil cost for all survey respondents |
|---|-----------------------------|--------------------------------------|--|---|---|
| <b>Furniture</b>                                  | 15                          | £1.73                                | £0.58  | £0.72   | £0.60   |
| <b>Physical environment</b>                       | 3                           | £3.47                                | £0.47  | £0.09   | £0.25   |
| <b>Catering facilities</b>                        | 34                          | £1.40                                | £0.68  | £1.47   | £1.00   |
| <b>Resources</b>                                  | 33                          | £0.70                                | £0.16  | £0.82   | £0.46   |
| <b>Staff training and recruitment</b>             | 18                          | £2.32                                | £0.41  | £1.68   | £0.98   |
| <b>Other</b>                                      | 2                           | £0.73                                | £0.08  | N/A   | £0.04   |
| <b>Upfront Costs</b>                              | 41                          | £3.75                                | £2.38  | £4.78   | £3.31   |
| <b>Food &amp; drinks (from Magic Breakfast)</b>   | 44                          | £10.16                               | £8.99  | £10.89  | £9.84   |
| <b>Food &amp; drinks (provided by the school)</b> | 23                          | £1.93                                | £0.75  | £1.24   | £0.92   |
| <b>Ongoing costs</b>                              | 46                          | £10.75                               | £9.74  | £12.12  | £10.75  |
| <b>TOTAL</b>                                      | <b>46</b>                   | <b>£14.07</b>                        | <b>£12.13</b>  | <b>£16.90</b>   | <b>£14.07</b>                                     |

*Average per-pupil costs in this table are found by dividing the cost averaged across the relevant schools (incurrers or responders, depending on the column) by the average pupil count of those schools. Because of the different denominators, numbers in the ‘incurrers’ column will not add.  
Costs are presented for a single school year.*

*Due to rounding, some numbers might not add perfectly.  
 Column 3 gives the average per-pupil cost only for those schools which actually incur that type of cost. Columns 4–6 give costs averaged over all schools that responded to the headteacher survey at follow-up or who had food costs reported by Magic Breakfast.  
 Cost information for the food provided by Magic Breakfast was obtained directly from Magic Breakfast. Other cost information comes from the survey of headteachers.*

In addition to their spending on food costs, schools reported relatively high staff time requirements. Table 18 explores how schools chose to supervise their breakfast clubs and what demands this placed on staff time and on parents and community volunteers.

**Table 18: Hours of staff and volunteer time required per academic year**

| Staff type                | Number of schools incurring | Average hours for incurrees | Average hours across all survey respondents | Average per-pupil hours across all survey respondents |
|---------------------------|-----------------------------|-----------------------------|---|---|
| Teachers                  | 13                          | 280.83                      | 86.92                                       | 0.27  |
| Teaching assistants       | 37                          | 538.38                      | 448.65                                      | 1.41  |
| <b>All teaching staff</b> | <b>39</b>                   | <b>576.77</b>               | <b>535.57</b>                               | <b>1.69</b>   |
| Catering staff            | 6                           | 220.83                      | 31.55                                       | 0.1   |
| Caretaking & maintenance  | 4                           | 546.25                      | 52.02                                       | 0.16  |
| Pastoral staff            | 10                          | 247.45                      | 58.92                                       | 0.19  |
| Office staff              | 10                          | 100.39                      | 21.51                                       | 0.07  |
| <b>All support staff</b>  | <b>26</b>                   | <b>264.92</b>               | <b>164</b>                                  | <b>0.52</b>   |
| Volunteers                | 10                          | 419.75                      | 99.94                                       | 0.32  |
| Other                     | 4                           | 214.38                      | 20.42                                       | 0.06  |
| <b>Total hours</b>        | <b>42</b>                   | <b>819.93</b>               | <b>819.93</b>                               | <b>2.58</b>   |

*Average per-pupil hours are found by dividing the average hours of the relevant schools (incurrees or responders) by the average pupil count of those schools. Because of the different denominators, numbers in the 'incurrees' column will not add.  
 Hours are presented for a single school year.  
 Due to rounding, some numbers might not add perfectly.  
 Column 3 gives the average per-pupil number of hours only for those schools which actually use that type of staff time. Columns 4–5 give hours averaged over all 42 intervention schools that responded to the headteacher survey at follow-up.*

Almost all schools used teaching assistants for supervision. On average, schools used 535 hours of teaching staff time, or 1.7 hours per pupil per year. Schools relied on a variety of other staff to deliver the Magic Breakfast intervention, including pastoral staff, office staff, and volunteers.

On average over the course of the year, schools used 820 person-hours to deliver the Magic Breakfast intervention (2.6 hours per pupil). The evidence from the case study schools suggests that a large part of this time was supervisory, but other major activities included food preparation and setting up and cleaning the room where breakfast was served. Since schools were not explicitly asked about the time spent on staff training, it is possible that these figures slightly underestimate the actual time costs in the

short run as staff training hours are excluded from Table 18 (the cost of these hours is reported in Table 17).

It should be emphasised that these figures for cost and supervision time are calculated based on a single year of data and assumes that provision and running costs would be the same for the following two years. Estimating the cost over a three-year period is difficult based on this data, since both the headteacher survey and the case studies indicate that schools would consider changing a number of facets of their delivery if they continue with the Magic Breakfast programme, all of which would impact cost.

## Process evaluation

### Implementation

#### Setting up Magic Breakfast

Although schools shared a common rationale (see Introduction) for introducing a breakfast club they varied in terms of how they chose to deliver breakfast in practice.

In each school, the set-up process was supported by a dedicated member of Magic Breakfast called a school change leader. School change leaders helped schools find solutions to particular problems as they were setting up their provision, as well as sharing learning gained from other schools. School change leaders encouraged schools to think flexibly, to keep their designs simple and manageable and, from the outset, to consider how to ensure that delivery models would be sustainable beyond the Magic Breakfast supported pilot phase. This support would be unlikely to be available to other schools unless they were supported by Magic Breakfast or a similar charity.

All schools originally signed up to the Magic Breakfast programme with the understanding that the model of delivery should be universal and free. In practice, however, the approach adopted in schools varied and two of the schools in our case study sample introduced a charge for breakfast (of £2 or £5 per pupil per week, with some flexibility for families with multiple children and more vulnerable families). Aside from cost to parents, across the four case study schools the nature of the breakfast offer also differed considerably in terms of start time, location, staffing arrangements, food choices, and charging structure.

Our sample of four schools included two distinct delivery models:

- a traditional 'breakfast club model' that involved breakfast being delivered in a single location before the official start of the school day—various food options were available for children to choose from which were laid out in a buffet-style format; and
- an 'in classroom model' that involved multiple breakfasts being prepared in every classroom in the school with children only offered one breakfast option each morning.

The different approaches to developing a delivery model were informed by three related considerations.

#### Funding

The need to top up the Magic Breakfast grant of £300 for setting up and delivering the provision resulted in staff seeking donations of items, or approaching private companies for financial donations. One large school reflected that,

*'The money they gave us would never have matched what we needed because we are just too big.'*

Initially the funding was needed to purchase items to help with food preparation and storage resources. Elsewhere, governors or senior colleagues ring-fenced pupil premium funds in the school budget to cover any outstanding set-up costs.

Otherwise, schools needed to find a way to cover ongoing costs for food items not provided by Magic Breakfast (such as milk and spreads) and to reimburse staff for the time they spent preparing and supervising breakfast. Where delivery was integrated within the existing school day through an in-classroom model, teaching staff viewed the provision of breakfast as part of their morning routine. In contrast, where a breakfast club model was adopted, breakfast delivery and supervision was a more substantial task and required staff to work significantly more than their contracted hours. Breakfast



supervisors, therefore, needed to be compensated for the additional time they spent at work. One solution to reimbursing staff took the form of providing staff with time off in lieu instead of payment. An alternative way of overcoming the challenge of paying staff was to charge parents for the breakfast provision.

### Access

The desire to facilitate equal access had an impact on how schools organised the timing and delivery of their breakfast offer and whether they charged for the provision.

- **Charging:** The decision to offer a free breakfast was underpinned by the belief that cost had acted as a deterrent to take-up with previous school breakfast clubs. One Magic Breakfast School Change Leader said:

*'Like many schools, we have tried a breakfast club in the past and in conversation with [the School Breakfast Co-ordinator] we established that parents said they didn't want it... and it was the charge that was the obstacle. So, something we talked about a lot was the importance of having a free breakfast provision with no stigma and no barriers to entry.'*

Conversely, in the two schools where families were charged for their breakfast, staff said that the funds were needed to cover ongoing costs, such as staffing. Another reason given was that it would encourage a greater commitment from children who opted to have breakfast, enabling staff to estimate the likely demand and to plan the delivery of provision and staff needed more effectively.

- **Timing and delivery:** Where it was hoped that breakfast would help to address problems around attendance and punctuality, schools designed models which were integrated into the start of the school day, with a later start-time which was seen as 'realistic' for families. One School Breakfast Co-ordinator commented,

*'Partly we were thinking if we can have a breakfast club to entice the children who were normally late, then that [just before the start of the school day] would be the best time to do it.'*

Conversely, where there was a demand from parents for early morning childcare provision, schools designed models with earlier start times, creating a greater burden in terms of staff time. Access considerations also inspired parallel initiatives designed to make it as easy as possible for children to make use of the breakfast provision. One approach was to integrate breakfast delivery with fun and sociable 'soft start' activities taking place in every classroom each morning. This involved children being invited to come into the classroom in a gradual, relaxed way from 8:45 each morning. Another initiative designed to facilitate access provided a 'walking bus' in conjunction with a breakfast club offer. This involved members of staff picking children up from their homes and walking with them, as a group, to school in time for them to have breakfast. Both initiatives were seen as a way of maximising engagement with breakfast provision.

### Practical

Practical considerations had a bearing on the location of breakfast. Where catering facilities were located off-site, or were being used by external catering staff, there was a need to set-up self-contained food preparation areas and find room for larger pieces of equipment such as freezers. Compromises had to be made, such as changing the location of a breakfast club from a more comfortable library space to an assembly hall with an adjacent room for food preparation and storage. Similarly, practical considerations influenced the design of breakfast menus and the food that was on offer to children. For example, in classrooms staff felt bagels were the only practical option to minimise mess.

## Barriers to take-up

A number of barriers were identified as having deterred children from taking up the breakfast offer in their school. These included:

- **Being charged for breakfast.** This was perceived to have resulted in a lower take-up of breakfast compared to charging for breakfast. Delivery staff expressed the view that ‘if it was free, we’d have more parents... of course we would’. Similarly, in the case study schools that did charge for breakfast, rates of take-up were observed to decrease when the price was increased. Once the price was returned to the original level then take-up rates started to increase again.
- **Earlier start times** to cater for breakfast provision. Earlier start times appeared to have a bearing on take-up rates. The ease with which parents were able to take children to school even earlier than normal depended on their personal circumstances and, specifically, whether they or their children had any mobility issues.
- **A lack of positive promotion** about breakfast provision. This was also felt to have contributed to disappointing take-up rates. Within our sample, there was a perception from some that more could have been done to highlight the benefits of the breakfast offer and create a ‘buzz’ around it. However, where there was felt to be a stigma attached to the offer of Magic Breakfast the process of actually broaching the subject with particular parents was identified as a challenge in itself.

## Delivery challenges

In addition to the barriers to take-up there were a number of other challenges that staff identified when setting up and delivering breakfast provision in the four case study schools. These included:

- **Difficulty managing supply and demand**, particularly in schools that offered children a choice of food items each morning. One reason for this was that children’s preferences regularly seemed to change and staff explained that it was difficult to anticipate this. Schools tried to avoid this situation by stocking particular food items. This also helped safeguard against problems resulting from the wrong items being delivered by suppliers, which was reported as happening from time to time.
- **A lack of engagement from governing bodies** resulted in considerable challenges for management especially in relation to accessing budgets and resources.
- **Poor levels of delegation within teams** resulted in staff not feeling involved in decision-making or feeling ‘part of a team’ or supported by managers. Instead, staff wanted to ‘feel part of the decision making... rather than finding out what’s been decided’.
- **Over-reliance on the goodwill of staff** to deliver the breakfast club—where staff went above and beyond their responsibilities—was not felt to be a sustainable approach to breakfast delivery. Indeed, some delivery staff said:

*‘We claim the time as half an hour, but we’re actually here for an hour [extra]. So it’s a lot of good will, but I won’t have it any other way.’*

## What is key to effective delivery?

Stakeholders identified a range of important steps and essential elements that they considered to be key to the effective delivery of breakfast club provision. These were in part driven by the barriers to take-up and delivery challenges identified in the previous sections. They largely centred on a communication strategy to inform parents about the breakfast offer and promote take-up, as well as the importance of having an established breakfast routine and a well-functioning Magic Breakfast team within the wider school community.

### Encouraging the take-up of the breakfast club

Staff highlighted the importance of promoting and raising awareness about breakfast by creating a buzz about it and encouraging take-up via word of mouth. A number of different approaches were used to alert parents about the breakfast offer including: advertising in newsletters and on websites, the use of social media such as a school app, running initiatives such as 'Family Fridays' where parents were invited to come along and enjoy a free breakfast with their children, and holding themed breakfasts on particular days which were designed to 'add something a bit different' to the breakfast offer.

Key to successful promotion was the need to avoid any feeling among children and parents that the breakfast offer at school was 'just for the poor families'. Encouraging all children—from a wide variety of socio-economic backgrounds—to eat breakfast at school was seen as a good way of removing any stigma. However, the importance of targeting any promotion at pupils who would be most likely to benefit was equally recognised as being an effective approach, as long as staff were able to identify potential vulnerability. In some circumstances, consideration was given to those families who were known to be receiving support from social services and had children on child protection plans, or the children in the need register. Speaking about this, one School Breakfast Co-ordinator commented:

*'Because quite often... it's your children who are on plans who come to school late and don't have their uniform and they haven't got their reading book and they haven't had breakfast, and it's not that they're not cared for, or they're neglected, there's just a bit of a lack of organisation.'*

In these circumstances, it was felt that children should be sensitively targeted through personal contact with parents, sending personalised letters and, more innovatively, through proactive efforts to get children into school on time. Where this approach was adopted, parents described feeling that staff cared about them and recognised some of the individual challenges that they faced encouraging their child to take part.

### A well-functioning Magic Breakfast team

The key features of a well-functioning Magic Breakfast team included:

*A governing body who were supportive of the breakfast club initiative* and helped to maintain momentum, obtain budgetary approval for ongoing costs, and contribute to the take-up and school-wide 'buy-in'. Where this had been achieved, the school community was more willing to support the breakfast provision by, for example, assisting with breakfast clubs when staff were sick, or teachers being accommodating if children occasionally arrived late to class. This level of 'buy-in' was attributed to the provision of clear information about the rationale behind the breakfast club at the beginning of the academic year, regular feedback and opportunities to provide feedback on training days, and through forums such as the staff council.

*A passionate School Breakfast Co-ordinator*—a post usually fulfilled by senior staff, including business managers and assistant or deputy headteachers—was judged to be vital for the successful internal management of the breakfast club. Drawing on their experiences of working in multiple schools, Magic Breakfast school change leaders felt the right person to act as School Breakfast Co-ordinator was someone who was close to the day-to-day running of the school but who held enough 'seniority to make the decisions and get things done'. In addition to this, passion, and a willingness to trust and delegate to the delivery staff, are important attributes. Certain tasks, such as thinking through the daily logistics of delivery and liaising with suppliers about orders, were often handed over to delivery staff who were seen as best placed to do this. One School Breakfast Co-ordinator expressed the view that the 'secret' was to identify key individuals who were 'reliable and good at organising things' and then to delegate responsibility to them. They felt that not only did this serve the purpose of relieving some of the administrative burden from School Breakfast Co-ordinators, it also helped to foster ownership among the team.

*Dedicated and committed staff* were seen as necessary for effective delivery. Delivery staff commented that they saw their role as breakfast club supervisors as among their most important responsibilities and felt extremely dedicated to supporting the children and the delivery team. This enthusiasm and 'can do' attitude of delivery staff was widely recognised by their colleagues throughout their schools.

*Well-established routines* contributed to effective delivery and the creation of the 'right' environment. This depended on having routines structured around clearly defined staff roles, where each member of staff took responsibility for a different task. Simple, time-saving strategies that allowed daily tasks such as the food preparation and clear-up to be completed as efficiently as possible also proved to be very important for delivery staff.

### Children's views and experiences

Children reported generally positive views and experiences of having breakfast in their school. These experiences were echoed by both school staff and Magic Breakfast staff. Children reported really enjoying the food provided and acknowledged that it was 'healthy' and 'nutritious'. When children were asked to create a typical breakfast plate they selected stickers of bagels, porridge, cereal, and fruit juice. However, Figure 3 shows that in an ideal world children would have liked substantially more choice. While their wish list included some 'treat' foods (such as pancakes, muffins and pastries), a popular request made by the children we spoke to was for fruit to be offered as a breakfast option.

Figure 3. What children have and would like for breakfast.



All schools worked hard to find out what food children liked and it appeared that children liked the food they were being served. Staff had successfully encouraged children to try new foods by describing them in ways children would understand, for example, bagels were described as 'round bread' for those who were not familiar with them. Such engagement appeared to have worked as staff reported that children were willing to try new foods, which they appeared to enjoy.

Delivery staff also tried to ensure that breakfast was a fun activity and they encouraged children to play with the games on offer, which helped to encourage cross-age playing, facilitating interaction between children in different friendship groups, classes and year groups. One member of delivery staff expressed the view:

*'It's not about them sitting down and being completely silent for an hour and a half. It's a breakfast club, they've got six and a half hours of doing that. It's about them enjoying themselves, having fun and having a bit of a giggle and stuff isn't it? Being creative, drawing and playing games with each other.'*

Children liked the opportunity to make new friends across the school and valued the social side of breakfast and having more time with their friends. They also appeared to value having the opportunity to talk to teachers in a more informal way, outside of their learning environment; children particularly

liked how the staff ‘had their own personalities at breakfast’ and described them as ‘some of the best staff in school’. As a result of these opportunities to develop positive relationships, children saw the delivery staff as people who could offer emotional support and who they might turn to for help.

While children were generally positive about their experience of breakfast, they did express some dislike about its timing and location. Children in schools that opted to have a later start time, or incorporated breakfast into their soft start activities, reported that they would like more time to eat their food and socialise with their friends. Furthermore, children who ate their breakfast in the classroom did not feel as though this was anything special and were quick to cite other locations as more exciting or beneficial, for example, the communal hall where they would interact with others, or the playground where they could enjoy some fresh air with their breakfast.

## Outcomes

Staff, children, and parents spoke of many benefits of the breakfast club intervention on the pupils, the classroom, and the general school environment. In line with the theory of change, participants perceived, in the four case study schools, that children were less likely to be coming into lessons hungry. As a consequence they felt that classrooms were more settled as children were less disruptive and more able to concentrate. Teachers also reported that there was an overall decrease in disruption during lesson time. One School Breakfast Co-ordinator observed:

*‘Children were unfocused and fidgety because they were hungry or they’d had the wrong breakfast, like some of them come and they’ve had a packet of Wotsits or something that they’ve got from the shop on the way, or they’re hyped up when they’ve got here and they’ve slumped very quickly. I think rather than noticing that their children are concentrating more, they’re [the teachers] noticing that they can get through the lesson without having to stop and say “what’s the matter?”, “Are you hungry?”, “Do you need some breakfast?”.’*

Indeed, teachers viewed classes as being much calmer after the introduction of the breakfast club, and this was most noticeable where a soft start model was adopted. School staff also perceived that there was a reduction in children coming in late, possibly as a result of offering the incentive of breakfast.

As well as helping to make children more independent and introducing them to new foods, the breakfast club was also believed to positively impact on children’s relationships, social skills, and eating habits. It gave children the opportunity to develop relationships with staff and pupils outside of their formal learning environment and thus positively influenced their view of the school.

Speaking of this, one School Breakfast Co-ordinator commented,

‘It’s nice for the kids, they go “oh hello Sir”... they feel positive enough in themselves to go up to an adult, especially a man in a suit. They’re brave. Any other time they would just shy away.’

Participants found it much harder to speculate whether the breakfast club was having an impact on performance and attainment. However, schools said they were actively encouraging Year 6 pupils to take advantage of the breakfast offer during SATs week as it was perceived that this would help them to concentrate.

In addition to having an impact on pupils, the classroom, and schools, the breakfast club was also perceived to help families. Parents spoke of benefits, including having less stress in the morning and fewer arguments with their children over breakfast. Parents felt that a school breakfast club helped to ease their morning routine and allowed them to get off to work or appointments earlier.

The impacts identified by the different participants were observed to have been mediated by a number of factors related to the school, as well as the age and background of the children who ate school breakfasts. These included:

- **The breakfast model adopted.** It was suggested that there was a greater potential for impact when schools served breakfast in a communal location where children of different ages could interact with each other and staff.
- **The wider school and policy context.** Any reflections about the impact of the breakfast club need to be considered alongside the other initiatives operating in schools to improve attendance and behaviour.
- **Age of children eating school breakfasts.** For older children, the impact was perceived in terms of an increase in independence and children getting themselves up and into school on time, whereas for younger children it helped to ease them into school life and helped them feel comfortable in new environments.
- **The children being targeted.** Schools that targeted children who they viewed as being most vulnerable felt more able to cite impacts around a perceived reduction in hunger, more settled behaviour, and improvements in concentration. For these children, schools felt that the impacts appeared to be much more pronounced than for children who used the breakfast club as a supplementary form of breakfast, in addition to the one they received at home.

### Sustainability of Magic Breakfast

Staff across the sample expressed a strong desire to continue providing an in-school breakfast in the future. Parents who took part in the research also echoed this enthusiasm for the continuation of breakfast provision. However, while commitment to the principle of giving children breakfast in school was unwavering, concerns about the feasibility of continuing, once the practical and financial support provided by Magic Breakfast was withdrawn at the end of the pilot phase, were widely expressed.

Specifically, there were concerns about how to remunerate staff and cover the cost of food without introducing a charge or raising the price of breakfast. It was acknowledged by some that an elevated charge would defeat the very purpose of providing breakfast in the first place, by making it inaccessible to the many of the children it was designed to support. Even where staff were reimbursed through time off in lieu, there was ambiguity over whether or not this arrangement could continue.

Other concerns related to sustainability were linked to the unique role played by individual 'breakfast champions', with fears being expressed that the momentum might be lost if particular individuals were to leave a school. Related to this were fears that schools might be veering towards an over-dependence on the good will of delivery staff. This ran the risk of causing staff to burn out and would also be difficult to replicate following any changes in personnel.

Overall, however, schools appeared to be optimistic that they could find a way to ensure there was a continuing breakfast offer at their schools. Schools were open to novel ideas, alternative funders, and different types of foods. In terms of generating new sources of income, one approach that was encouraged by Magic Breakfast school change leaders was to seek sponsorship from local businesses. Elsewhere, school change leaders had worked with staff to complete a costing analysis exercise to show how much money would be needed to sustain the breakfast provision. It was hoped that this could be used to make a case to governing bodies for ring-fencing a proportion of school funds so as to protect the provision.

## Conclusion

### Interpretation

#### Key conclusions

1. Year 2 children in breakfast club schools experienced around two months' additional progress compared to Year 2 children in the other schools in the trial. These positive results would be unlikely to occur by chance.
2. For Year 6 children in breakfast club schools, results for the main outcomes, reading and maths, were positive but could have occurred by chance. However, on other measures of writing and English they experienced around two months' progress compared to the other Year 6 children. These positive results would be unlikely to occur by chance.
3. The findings suggest that it is not just eating breakfast that delivers improvements, but attending a breakfast club. This could be due to the content of the breakfast itself, or to other social or educational benefits of the club.
4. Pupil behaviour, as measured by a teacher survey, improved in breakfast club schools. This is interesting because it shows that breakfast clubs may improve outcomes for children who do not even attend breakfast club, by improving classroom environments.
5. Activities thought to increase take-up of the breakfast provision included promoting it to parents and encouraging all children to attend while sensitively targeting pupils most likely to benefit. The project required additional staff time which some schools found difficult to provide without charging for breakfast.

Magic Breakfast was introduced in the context of concerns about pupils arriving at school hungry and the impact that this may be having on their wellbeing, behaviour, and experiences in the classroom. The introduction of Magic Breakfast provision was seen by staff as a way to tackle a number of pre-existing problems:

- to address problems of pupil hunger and, by virtue of this, improve children's wellbeing, concentration, and behaviour in class;
- to improve attendance and punctuality by creating an added incentive for children to arrive at school on time;
- to support parents by reducing the stress of morning routines and providing early morning childcare; and
- to establish and reinforce a welcoming and inclusive community ethos within the school.

At the end of the intervention around 70% of participating intervention schools (that responded to the follow-up headteacher survey) reported that they planned to continue breakfast provision, which reflects the positive experience of the intervention and perceived positive impacts. For example, the majority of headteachers felt that concentration, behaviour, attendance, and attainment had improved.

The impact evaluation findings are largely consistent with headteachers' perceptions. There was evidence to suggest that attainment at the end of Key Stage 1 improved more than would otherwise be expected by chance, equivalent to around two months' progress. Attainment at the end of Key Stage 2 also improved, but this improvement could have been due to chance. However, on some *secondary* attainment outcomes (teacher-assessed ability in writing and English) KS2 pupils improved more than would otherwise be expected.

There are multiple mechanisms through which this effect occurred that support the Theory of Change outlined in Figure 1. First, breakfast consumption at school increased, although breakfast consumption overall increased only marginally. This suggests that the school context (which was typically more social) or school food (which was potentially more nutritious) contributed to the improvement in attainment, rather than whether or not breakfast was eaten.

Second, there were large improvements in teacher-reported levels of concentration and behaviour in the classroom. This means that the classroom environment improved for teachers, and presumably pupils, which may have been beneficial for learning. Through this mechanism, even pupils who did not change their breakfast consumption patterns may have benefitted from the intervention.

Attendance and punctuality recorded in administrative data improved slightly, which could be a result of improved health (fewer days of sickness) or a greater incentive to arrive at school on time. An increase in time in school for pupils, and reduced class disruption due to lateness and absence, might therefore have also played a role in improving outcomes.

Finally, there was no evidence to suggest that Year 6 pupils' Body Mass Index (a proxy for health) was affected. This data is less reliable, however, as only school-level averages for Year 6 pupils for a limited set of schools were available, with no individual-level control for BMI before the intervention.

The effect sizes for Key Stage 1 attainment are of a similar magnitude to the effect size in the evaluation of universal free school meals in two pilot areas (Brown *et al.*, 2012) that led to a national roll-out of this policy. The breakfast club intervention occurred at a time when all KS1 pupils were eligible for free school lunches, which suggests that the provision of a school breakfast has an effect over and above that of provision of school lunch. This was particularly noticeable in teachers' reports of classroom behaviour and concentration, where impacts were larger for KS1 than for KS2, despite KS1 pupils already being eligible for a free school meal.

In contrast to the evaluation of universal free school meals, the main effect for KS2 was smaller than for KS1, despite the intervention's significantly larger effect on school breakfast consumption at KS2. This may be because there was a high presence of breakfast clubs in control schools during the week of national tests (above 90% of schools), or because the improvement in concentration and behaviour translated less into KS2 externally-marked assessments than KS1 teacher assessments.

The findings are broadly consistent with evidence for Wales, where a national breakfast club policy was made a duty for maintained schools in 2013. An early randomised controlled trial found that absences reduced and concentration improved as a result of the breakfast club, although in that case conduct also worsened.

Outcomes discussed by case study participants were generally positive, for example, breakfast was perceived to have reduced hunger, increased exposure to new foods, and improved behaviour, attendance, and social skills. It was evident that schools in the case study and follow-up surveys adapted the model of breakfast club provision to their context. Access and funding considerations exerted the biggest influence over the design of breakfast provision, with some schools deciding that the need to cover the cost of staff supervision outweighed access considerations for keeping the club free of charge.

Data from the four case study schools suggested that delivering breakfast effectively appeared to depend upon a proactive and innovative approach to getting children through the door, the establishment of a solid daily routine built around simple time-saving strategies, and the existence of a well-functioning, mutually-supportive team driving forward and championing the breakfast provision. Delivery challenges arose in the form of a number of different barriers to children accessing the provision, difficulties in balancing supply and demand, and problematic team dynamics. Overall, it was these team dynamics, particularly the extent to which the wider school community had bought in to the



importance of Magic Breakfast, which appeared to be most instrumental in determining the success of the breakfast club because they set an overarching tone which either supported or undermined other aspects of delivery.

While commitment to the principle of giving children breakfast in school was unwavering, concerns were widely expressed about the feasibility of continuing the provision once the practical and financial support provided by Magic Breakfast was withdrawn.

Specifically, there were concerns about how to remunerate staff and cover the cost of food without introducing a charge or raising the price of breakfast. It was acknowledged by some that an elevated charge would defeat the very purpose of providing breakfast in the first place, by making it inaccessible to many of the children it was designed to support. Even where staff were reimbursed through time off in lieu, there was ambiguity over whether or not this arrangement could continue.

Other sustainability-related concerns were linked to the unique role played by individual 'breakfast champions', with fears being expressed that the momentum might be lost if particular individuals were to leave a school. Related to this were fears that schools might be veering towards an overdependence on the goodwill of delivery staff. Not only did this run the risk of causing staff to burn out, it would also be difficult to replicate following any changes in personnel.

Overall, however, schools appeared to be optimistic that they could find a way to ensure there was a continuing breakfast offer at their schools. Schools were open to novel ideas, different types of food, and alternative funders, for example approaching local businesses for sponsorship.

## Limitations

The results of this evaluation are relevant to schools in a similar social and economic context to the schools in the trial—those that have a higher proportion of pupils eligible for free school meals than the national average, and are more likely to be in urban areas. The estimated impacts are generalizable to schools in these similar circumstances as the trial was an effectiveness (or 'real world') trial.

Schools varied their approach to breakfast club provision depending on their constraints and objectives, despite signing up to provide a universal and free breakfast club, which means that fidelity to the tested model was sometimes limited.

The majority of outcomes are accurately and reliably recorded, and available for almost all pupils of interest. Primary outcomes at Key Stage 1 are based on teacher assessments rather than externally-marked tests, however, which could introduce some bias if teachers' perceptions are skewed by knowledge of participation in the trial. However, teacher assessments and externally-marked tests at Key Stage 2 show a similar effect of the breakfast club where it is possible to compare, which suggests that teacher assessments are reliable indicators of attainment in these schools.

The evaluation has not been able to adequately explore the effect of the intervention on health. The reduction in overall absences from school is consistent with fewer days of illness, but not conclusive. The measurement of body mass index used in the evaluation is a school-level measure for Year 6 pupils only, where an individual baseline measure is not available, making the estimates more imprecise and subject to other variation.

The pupil survey had a lower response rate than the teacher and headteacher surveys and therefore has the most potential to be influenced by non-random response bias. For example, average points at both KS1 and KS2 were significantly lower among those students whose NPD records did not link to a pupil survey compared to those who could be linked. There is little difference in the estimated effectiveness of the programme on academic outcomes among those who can and cannot be linked to a pupil survey, however, providing some reassurance that the effects estimated using the pupil surveys should be reasonably representative of the effects amongst all trial pupils.

### **Future research and publications**

A large number of schools in England provide breakfast clubs of some form, and the direction in U.K. policy seems to be for this to increase. Further research is therefore required to inform schools about the most effective ways of delivering breakfast provision and improving pupil attainment. Through case study visits, this evaluation has suggested that the school context is an important determinant of the feasible model of breakfast club provision, but further quantitative research should provide more evidence on the direct effects of different choices schools can make, for example whether to provide breakfast before school or during school, or with or without charge, for some or all pupils.

## References

- AITB (1944) *Army Individual Test Battery, Manual of Directions and Scoring*, Washington, DC: War Department, Adjutant General's Office.
- Crawford, C. and Skipp, A. (2014) 'LIT Programme. Evaluation Report and Executive Summary', London: Education Endowment Foundation. Available at: [https://educationendowmentfoundation.org.uk/public/files/Projects/EEF\\_Project\\_Report\\_LITProgramme.pdf](https://educationendowmentfoundation.org.uk/public/files/Projects/EEF_Project_Report_LITProgramme.pdf) [accessed 5 July 2016].
- Department for Education (2015a) 'KS1 and KS2 test and examination point scores used in the 2015 school performance tables'. Available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/516898/Key\\_stage\\_1\\_and\\_2\\_test\\_and\\_examination\\_point\\_scores.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/516898/Key_stage_1_and_2_test_and_examination_point_scores.pdf).
- Department for Education (2015b) 'Schools, pupils, and their characteristics: January 2015'. Available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/433680/SFR16\\_2015\\_Main\\_Text.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/433680/SFR16_2015_Main_Text.pdf)
- Gadernann, A. M., Guhn, M. and Zumbo, B. D. (2012) 'Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide', *Practical Assessment, Research & Evaluation*, 17 (3). Available at: <http://pareonline.net/pdf/v17n3.pdf>
- Grantham-McGregor, S. (2005) 'Can the provision of breakfast benefit school performance?', *Food and Nutrition Bulletin*, 26, pp. 144–58.
- Goodman, R. (1997) 'The Strengths and Difficulties Questionnaire: A Research Note', *Journal of Child Psychology and Psychiatry*, 38, pp. 581–586.
- Brown, V., Crawford, C., Dearden, L., Greaves, E., Kitchen, S., Payne, C., Purdon, S. and Tanner, E. (2012) 'Evaluation of the Free School Meals Pilot: Impact Report', London: Department for Education. Available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/184047/DFE-RR227.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/184047/DFE-RR227.pdf) [accessed June 2016].
- Mahoney, C. R., Taylor, H. A., Kanarek, R. B. and Samuel, P. (2005) 'Effect of breakfast composition on cognitive processes in elementary school children', *Physiology & Behavior*, 85, pp. 635–645.
- Meyers, A. F., Samson, A. E., Weitzman, M., Rogers, B. L. and Kayne, H. (1989) 'School Breakfast Program and School Performance', *American Journal of Diseases of Children*, 143, pp. 1,234–1,239.
- Minnesota Department of Children, Families and Learning (1998) 'School Breakfast Programs Energizing the Classroom'. Available at: <http://www.californiahealthykids.org/articles/energize.pdf> [accessed 22 June 2016].
- Murphy, J. M., Pagano, M., Nachmani, J., Sperling, P., Kane, S. and Kleinman, R. (1998) 'The Relationship of School Breakfast to Psychosocial and Academic Functioning', *Archives of Pediatric and Adolescent Medicine*, 152, pp. 899–907.
- Murphy, S., Moore, G. F., Tapper, K., Lynch, R., Clarke, R., Raisanen, L., Desousa, C. and Moore, L. (2011) 'Free healthy breakfasts in primary schools: a cluster randomised controlled trial of a policy intervention in Wales, UK', *Public Health Nutrition*, 14 (2), pp. 219–26.
- Pollitt, E., Cueto, S. and Jacoby, E.R. (1998) 'Fasting and Cognition in Well- and Undernourished Schoolchildren: A Review of Three Experimental Studies', *The American Journal of Clinical Nutrition*, 67 (4), pp. 779S–784S.

Powell, C. A., Walker, S. P., Chang, S. M. and Grantham-McGregor, S. M. (1998) 'Nutrition and education: a randomized trial of the effects of breakfast in rural primary school children', *The American Journal of Clinical Nutrition*, 68 (4), pp. 873–879.

Rampersaud, G. C., Pereira, M. A., Girard, B. L., Adams, J. and Metz, J. D. (2005) 'Breakfast Habits, Nutritional Status, Body Weight, and Academic Performance in Children and Adolescents', *Journal of the American Dietetic Association*, 105 (5), pp. 743–760.

Ritchie, J. and Spencer, L. (1994) 'Qualitative data analysis for applied policy research' in A. Bryman and R. G. Burgess (eds) *Analyzing qualitative data*, Abingdon, Oxon: Routledge, pp.173–194.

Rodda, M., Hallgarten, J., and Freeman, J., 2013. 'Between the cracks: Exploring in-year admissions in schools in England.' Royal Society for the Encouragement of Arts, Manufactures, and Commerce. Accessed online at <<https://www.thersa.org/globalassets/pdfs/reports/education-between-the-cracks-report.pdf>> [Accessed October 2016].

Shaw, M. E. (1988) Adolescent breakfast skipping: an Australian study', *Adolescence*, 33, pp. 851–61.

Shemilt, I., Harvey, I., Shephstone, L., Swift, L., Reading, R., Mugford, M., Belderson, P., Norris, N., Thoburn, J. and Robinson, J. (2004) 'A national evaluation of school breakfast clubs: evidence from a cluster randomized controlled trial and an observational analysis', *Child: Care, Health and Development*, 5, pp. 413–427.

Stevens, L., Oldfield, N., Wood, L. and Nelson, M. (2008) 'The impact of primary school breakfast clubs in deprived areas of London', School Food Trust. Available at: [http://www.schoolfoodtrust.org.uk/download/documents/pdf/sft\\_breakfast\\_club\\_findings\\_dec08.pdf](http://www.schoolfoodtrust.org.uk/download/documents/pdf/sft_breakfast_club_findings_dec08.pdf) [accessed 22 June 2016].

## Appendix A: EEF cost rating

Cost ratings are based on the approximate cost per pupil per year of implementing the intervention over three years. More information about the EEF’s approach to cost evaluation can be found [here](#). Cost ratings are awarded as follows:

| Cost rating | Description   |
|-------------|---|
| £ £ £ £ £   | <i>Very low:</i> less than £80 per pupil per year.    |
| £ £ £ £ £   | <i>Low:</i> up to about £200 per pupil per year.      |
| £ £ £ £ £   | <i>Moderate:</i> up to about £700 per pupil per year. |
| £ £ £ £ £   | <i>High:</i> up to £1,200 per pupil per year.         |
| £ £ £ £ £   | <i>Very high:</i> over £1,200 per pupil per year.     |

## Appendix B: Security classification of trial findings

| Rating | Criteria for rating   |            |            | Initial score | Adjust  | Final score |
|--------|---|------------|------------|---------------|---|-------------|
|        | Design  | Power      | Attrition* |               |   |             |
| 5      | Well conducted experimental design with appropriate analysis  | MDES < 0.2 | 0-10%      | 5             |   |             |
| 4      | Fair and clear quasi-experimental design for comparison (e.g. RDD) with appropriate analysis, or experimental design with minor concerns about validity | MDES < 0.3 | 11-20%     |               | Adjustment for Balance<br>[ 0 ]                       |             |
| 3      | Well-matched comparison (using propensity score matching, or similar) or experimental design with moderate concerns about validity                      | MDES < 0.4 | 21-30%     |               | Adjustment for threats to internal validity<br>[ -1 ] |             |
| 2      | Weakly matched comparison or experimental design with major flaws   | MDES < 0.5 | 31-40%     |               |   |             |
| 1      | Comparison group with poor or no matching (E.g. volunteer versus others)  | MDES < 0.6 | 51-50%     |               |   |             |
| 0      | No comparator   | MDES > 0.6 | <50%       |               |   |             |

- **Initial padlock score:** lowest of the three ratings for design, power and attrition = 5 padlocks
- **Reason for adjustment for balance** (if made): None required as this is well- balanced
- **Reason for adjustment for threats to validity** (if made): The intervention was not implemented as planned as some schools charged for the intervention and some control schools had introduced a breakfast club.
- **Final padlock score:** initial score adjusted for balance and internal validity = 4 padlocks

\*Attrition should be measured at the pupil level, even for cluster trials.

## Appendix C: Consent form and Memorandum of Understanding



Tel: +44 (0) 20 7291 4800  
Fax: +44 (0) 20 7323 4780

mailbox@ifs.org.uk  
www.ifs.org.uk

7 Ridgmount Street  
London WC1E 7AE

8<sup>th</sup> September 2014

Dear Parent/Guardian,

Your child's school is participating in the Magic Breakfast project. Your school will have a universal free breakfast club from September 2014. This project will be the first major study in England about the impact of the provision of a universal free breakfast club on pupils' behaviour, concentration, and academic attainment in school.

This project is funded by the Department for Education and the Education Endowment Foundation (EEF), whose aim is to find out what works best to raise the attainment of pupils in England.

To help this research, your child will be asked to complete a short survey about what they had for breakfast one day in September 2014 and one day in June 2015.

Named data will be matched with the National Pupil Database and shared with Magic Breakfast, the evaluator (Institute for Fiscal Studies), the EEF data archive, and UK Data Archive for research purposes.

We will not use your child's name or the name of the school in any report arising from the research.

If you prefer for your child NOT to be part of this research, please inform the evaluation team: [mb\\_evaluation@ifs.org.uk](mailto:mb_evaluation@ifs.org.uk). If you would like more information, please contact [mb\\_evaluation@ifs.org.uk](mailto:mb_evaluation@ifs.org.uk).

Yours sincerely

**Alex Cunningham**  
Magic Breakfast Interim CEO

**Ellen Greaves**  
Senior Research Economist,  
Institute for Fiscal Studies

Director:  
Paul Johnson

Research Director:  
Richard Blundell

The Institute for Fiscal Studies  
Limited by Guarantee,  
Registered in England: 954616  
7 Ridgmount Street  
London  
WC1E 7AE

Registered Charity: 258815  
VAT no: GB 394 5830 17

## Magic Breakfast: Evaluating School Breakfast Provision

Magic Breakfast is a registered charity working to ensure no child starts the day too hungry to learn. With over a decade working with schools to set up breakfast clubs, we have seen the positive impact it can have on children's concentration and energy in class.

Now Magic Breakfast are launching an exciting new study to evaluate the impact of breakfast provision on pupil academic attainment. A randomised controlled evaluation will find out whether the provision of a school breakfast club has a beneficial impact on pupils' concentration, behaviour and academic attainment by reducing hunger and associated learning problems. The evaluation will also find out the best method of breakfast club provision. We have taken steps to ensure there is no unnecessary administrative burden on schools participating.

If you agree to take part and accept the terms and conditions for receiving the resources and training then please sign a copy of this form and return it the address provided at the end of this letter.

**All results will remain confidential, with no school being identified by name.** Participating schools will be randomly assigned either as a 'phase one' school (for one of two treatment groups) with food deliveries beginning in September or as a 'phase two' school whose food deliveries will begin in September 2015, with 50 schools being assigned to each group. Magic Breakfast will provide free healthy breakfast foods (as much as required) and equipment to set up the provision will be provided to all schools.

- Phase One Schools in the first treatment group will receive support to establish a **universal free breakfast club before the school day**, (i.e. offered for free to all pupils).
- Phase One Schools in the second treatment group will receive support to establish a **breakfast club before the school day, with a charge but offering free breakfast to pupils eligible for free school meals**.
- Phase Two Schools will receive 18months of food support deliveries from September 2015 and support to establish a breakfast provision which will be most effective for their school, beginning in September 2015.

All of the models have been tested in Magic Breakfast partner schools and shown to be successful.

As part of the evaluation KS1 and KS2 teacher assessments and test results will be obtained from the National Pupil Database to observe whether attainment improves, on average, in each treatment group relative to the control group. Differences between treatment groups will also be observed and tested.

In a small number of schools a researcher will be in contact up to twice in the year to see how the breakfast club operates within your school, and the views of school staff, pupils and parents about impact of the breakfast club.

### Requirements for schools

I confirm that:

- At least 35% of pupils in the school are currently eligible for free school meals.
- The school either does not currently have a breakfast club in place or has less than 6% of the school roll attending the breakfast club.



- The school will work with Magic Breakfast to establish a breakfast provision of the method that is randomly allocated by the evaluation team.

#### Requirements for schools: independent evaluation

I agree that:

- All pupils receiving breakfast at school will complete a 3 minute breakfast paper survey in September 2014 and June 2015.
- Classroom teachers of Year 2 and Year 6 pupils will complete a 5 minute online survey in September 2014.
- Classroom teachers of Year 2 and Year 6 pupils will complete a 5 minute online survey in June 2015.
- Headteacher will complete a 3 minute survey in June 2014 and 10 minute online survey in June 2015.
- Year 2 and Year 6 classroom teachers will provide valid email addresses to participate in their short survey.
- Teachers and schools will accommodate case study visits conducted by the National Children's Bureau. These will take place in a sub-sample of treatment schools (around 6% of schools) and involve confidential interviews and discussion groups with school leaders, teachers, catering staff, pupils and parents about their experiences and perceptions of impact of the breakfast club.
- The school agrees to the evaluation team (or other evaluating bodies commissioned by the Education Endowment Foundation) obtaining data on the KS1, KS2 and GCSE results from the National Pupil Database, and will provide the UPNs (as part of the pupil survey) to enable this to be achieved.

#### Commitments of Magic Breakfast

- Provide free food (as much as required) to support the breakfast club from September 2014 in phase one schools.
- Support with setting up the breakfast provision in the school, including a freezer for storage of breakfast food.
- Provide ongoing support to all schools from September 2014.
- Provide phase two with the most suitable method of breakfast club delivery in September 2015.
- Work towards sustainable breakfast clubs in the future with all schools.

#### Commitments of the evaluation team

- All data provided by schools will be stored in accordance with the Data Protection Act (1998).
- All results will be anonymised so that no schools will be identifiable in the report or dissemination of results. Confidentiality will be maintained and no one outside the evaluation team will have access to the database.

#### **Headteacher agreement**

I agree for my school \_\_\_\_\_ to take part in the Magic Breakfast evaluation and I accept the eligibility terms and conditions.

**Signature of Headteacher:** \_\_\_\_\_

**Name of Headteacher:** \_\_\_\_\_

**Email address of Headteacher:** \_\_\_\_\_

**Date:** \_\_\_/\_\_\_/\_\_\_\_\_

## Appendix D: Pupil, teacher and headteacher surveys

### Pupil Survey (baseline and follow-up)



Name:

Year group:

(Class teacher to complete)

Child's birthday - d/m/y:

Child's full name:

Child's UPN:

Did you have  
breakfast today?

Yes

No



Did you have breakfast  
at school today?

Yes

No

Did you like what you had  
for breakfast today or not?

Yes I liked it

It was OK

No, I didn't like it



Turn over!



magic breakfast  
fuel for learning



Did you feel hungry when you started your first lesson this morning, or not?

- Yes
- No
- Can't remember

What will you eat at lunchtime?

- A packed lunch
- A school lunch
- Nothing



I ate \_\_\_\_\_  
\_\_\_\_\_

What did you eat for breakfast?

Write and draw a picture



### Teacher Survey: Baseline

1. What is the name of your school?  
Text response:
  
2. What is your email address?  
Text response:
  
3. What local authority is your school in?  
Text response:
  
4. How long have you worked at your school?
  - a. Less than one year
  - b. One to two years
  - c. Two to five years
  - d. Five to ten years
  - e. More than ten years
  
5. What national curriculum Year Group do you teach?
  - a. Year 2
  - b. Year 6
  - c. Mixed (including Year 2)
  - d. Mixed (including Year 6)
  
6. Does your class have streaming (grouping children by general ability)?
  - a. Yes
  - b. No
  
7. Does your class have setting (grouping by subject-specific ability) for English?
  - a. Yes
  - b. No
  
8. Does your class have setting (grouping by subject-specific ability) for maths?
  - a. Yes
  - b. No
  
9. When the first lesson begins, I have to wait quite a long time for students to quiet down.
  - a. Strongly disagree
  - b. Disagree
  - c. Agree
  - d. Strongly agree

10. Students in this class take care to create a pleasant learning atmosphere.
- Strongly disagree
  - Disagree
  - Agree
  - Strongly agree
11. I lose quite a lot of time because of students interrupting the lesson.
- Strongly disagree
  - Disagree
  - Agree
  - Strongly agree
12. There is much disruptive noise in this classroom
- Strongly disagree
  - Disagree
  - Agree
  - Strongly agree
13. Thinking about the first lesson your class had today, what percentage of children in your class do you think had a **poor** level of concentration?
- 0-10%
  - 10-20%
  - 20-40%
  - 40-60%
  - 60-80%
  - 80-90%
  - 90-100%
14. Thinking about the first lesson your class had today, what percentage of children in your class do you think had a **good** level of concentration?
- 0-10%
  - 10-20%
  - 20-40%
  - 40-60%
  - 60-80%
  - 80-90%
  - 90-100%

15. Please give the percentage of your class that are usually ready to learn at the start of the first lesson of the day
- a. 0-10%
  - b. 10-20%
  - c. 20-40%
  - d. 40-60%
  - e. 60-80%
  - f. 80-90%
  - g. 90-100%
16. Please give the percentage of your class that are usually ready to learn at the start of the first lesson after lunch
- a. 0-10%
  - b. 10-20%
  - c. 20-40%
  - d. 40-60%
  - e. 60-80%
  - f. 80-90%
  - g. 90-100%

---

## Teacher Survey: Follow-up

Classroom teacher survey (for those who replied at baseline)

---

1. What is the name of your school?  
Text response:
  
2. What is your email address?  
Text response:
  
3. What local authority is your school in?  
Text response:
  
4. What national curriculum Year Group do you teach?
  - a. Year 2
  - b. Year 6
  - c. Mixed (including Year 2)
  - d. Mixed (including Year 6)
  
5. When the first lesson begins, I have to wait quite a long time for students to quiet down.
  - a. Strongly disagree
  - b. Disagree
  - c. Agree
  - d. Strongly agree
  
6. Students in this class take care to create a pleasant learning atmosphere.
  - a. Strongly disagree
  - b. Disagree
  - c. Agree
  - d. Strongly agree
  
7. I lose quite a lot of time because of students interrupting the lesson.
  - a. Strongly disagree
  - b. Disagree
  - c. Agree
  - d. Strongly agree
  
8. There is much disruptive noise in this classroom
  - a. Strongly disagree
  - b. Disagree
  - c. Agree
  - d. Strongly agree



9. Thinking about the first lesson your class had today, what percentage of children in your class do you think had a **poor** level of concentration?
- a. 0-10%
  - b. 10-20%
  - c. 20-40%
  - d. 40-60%
  - e. 60-80%
  - f. 80-90%
  - g. 90-100%
10. Thinking about the first lesson your class had today, what percentage of children in your class do you think had a **good** level of concentration?
- a. 0-10%
  - b. 10-20%
  - c. 20-40%
  - d. 40-60%
  - e. 60-80%
  - f. 80-90%
  - g. 90-100%
11. Please give the percentage of your class that are usually ready to learn at the start of the first lesson of the day
- a. 0-10%
  - b. 10-20%
  - c. 20-40%
  - d. 40-60%
  - e. 60-80%
  - f. 80-90%
  - g. 90-100%
12. Please give the percentage of your class that are usually ready to learn at the start of the first lesson after lunch
- a. 0-10%
  - b. 10-20%
  - c. 20-40%
  - d. 40-60%
  - e. 60-80%
  - f. 80-90%
  - g. 90-100%

## Headteacher Survey: Baseline

### 1. What is the name of your school?

Text response:

### 2. What local authority is your school in?

Text response:

### 3. How long have you worked at this school?

- a. Less than one year
- b. One to two years
- c. Two to five years
- d. Five to ten years
- e. More than ten years

### 4. What is your role in your school?

- a. Headteacher
- b. Deputy headteacher
- c. Assistant headteacher
- d. Early years coordinator
- e. Key stage coordinator
- f. Special educational needs coordinator
- g. Other (please specify)

### 5. How many classes does your school have for Year 2 pupils?

Text response:

### 6. How many classes does your school have for Year 6 pupils?

Text response:

### 7. Please give the email addresses of the Year 2 class teachers this academic year

Text response:

### 8. Please give the email addresses of the Year 6 class teachers next academic year

Text response:

### 9. What model of breakfast club provision would you prefer for your school?

- a. Universal provision before school
- b. Targeted free provision before school

### 10. What are your reasons for signing up to the Magic Breakfast pilot?

- a. Address problem of pupil hunger in the school
- b. Improve cohesion in the school
- c. Reduce inequality in the school

- d. Improve pupil concentration
- e. Improve pupil health and wellbeing
- f. Improve academic attainment
- g. Other (please specify)

**11. What is your main reason for signing up to the Magic Breakfast pilot?**

- a. Address problem of pupil hunger in the school
- b. Improve cohesion in the school
- c. Reduce inequality in the school
- d. Improve pupil concentration
- e. Improve pupil health and wellbeing
- f. Improve academic attainment
- g. Other (please specify)

**12. Would your current facilities be able to accommodate universal breakfast provision before school?**

- a. Yes
- b. No

**13. What is the current capacity of the school dining area?**

Text response:

**14. Does your school currently offer ad hoc breakfast provision?**

- a. Yes
- b. No

**15. If the school has some ad hoc breakfast provision, how many pupils access this on a usual day?**

Text response:

**16. Has your school had a breakfast club in the past which has closed down within the last 12 months?**

- a. Yes
- b. No

**17. Thinking about all the behaviour you encounter around school, how many pupils do you find generally badly behaved and / or difficult to deal with?**

- a. All / almost all
- b. Most
- c. Some
- d. Few
- e. None / almost none

**18. In your experience, when is indiscipline most likely to occur in your school?**

- a. Before the school day begins
- b. During morning classes
- c. Morning break
- d. Lunchtime
- e. During afternoon classes
- f. Afternoon break
- g. After the end of the school day

### Headteacher Survey: Follow-up (Control schools)

- 1. Please confirm your e-mail address.**  
Text response:
- 2. What is the name of your school?**  
Text response:
- 3. What local authority is your school in?**  
Text response:
- 4. Thinking about all the behaviour you have encountered around the school over the last term, how many pupils do you find generally badly behaved and / or difficult to deal with?**
  - a. All / almost all
  - b. Most
  - c. Some
  - d. Few
  - e. None / almost none
- 5. In your experience, when has indiscipline been most likely to occur in your school over the last term?**
  - a. Before the school day begins
  - b. During morning classes
  - c. Morning break
  - d. Lunchtime
  - e. During afternoon classes
  - f. Afternoon break
  - g. After the end of the school day
- 6. Has your school started offering any breakfast provision in the last 12 months?**
  - a. Yes
  - b. No
- 7. What type of provision have you started offering?**  
Text response:
- 8. Was breakfast offered to all Year 6 students at school during this year's SATs week?**
  - a. Yes
  - b. No
- 9. Is this the first year you have offered breakfast to Year 6 students during SATs week, or have you also done so in previous years?**
  - a. First year

- b. Done previously

**Headteacher Survey: Follow-up (Intervention schools)**

**1. Please confirm your e-mail address.**

Text response:

**2. What is the name of your school?**

Text response:

**3. What local authority is your school in?**

Text response:

**4. Does your school usually offer food BEFORE the school day as part of the Magic Breakfast pilot?**

- a. Yes
- b. No

**5. Where is this provision delivered?**

- a. In classrooms
- b. Somewhere else, e.g. dining hall

**6. Please specify where this provision is delivered:**

Text response:

**7. Which children can take advantage of this provision?**

- a. All children
- b. All children but with a limit on numbers
- c. Targeted children only (e.g. those eligible for free school meals)

**8. Which children are defined as "targeted" in terms of eligibility for breakfast provision before the school day.**

Text response

**9. How many children on average per day would you estimate take advantage of this type of provision?**

- a. 0-20 pupils on average per day
- b. 21-50 pupils on average per day
- c. More than 50 pupils on average per day

**10. Which of the following best describes the charging policy for this element of your provision?**

- a. Free to all children
- b. Free to targeted children with a charge for other children
- c. Free breakfast for all children but a charge for other facilities (e.g. childcare, activities)
- d. Free breakfast to targeted children but a charge for other facilities
- e. Other (text response)

**11. Which children are defined as "targeted" in terms of your charging policy for breakfast before the school day.**

Text response

**12. Does your school usually offer food DURING THE SCHOOL DAY (e.g. at registration or break-time) as part of the Magic Breakfast pilot?**

- a. Yes
- b. No

**13. Where is this provision delivered?**

- a. In classrooms

- b. Somewhere else, e.g. dining hall

**14. Please specify where this provision is delivered:**

Text response:

**15. Which children can take advantage of this provision?**

- a. All children
- b. All children but with a limit on numbers
- c. Targeted children only (e.g. those eligible for free school meals)

**16. Which children are defined as "targeted" in terms of eligibility for breakfast provision during the school day?**

Text response

**17. How many children on average per day would you estimate take advantage of this type of provision?**

- a. 0-20 pupils on average per day
- b. 21-50 pupils on average per day
- c. More than 50 pupils on average per day

**18. Which of the following best describes the charging policy for this element of your provision?**

- a. Free to all children
- b. Free to targeted children with a charge for other children
- c. Free breakfast for all children but a charge for other facilities (e.g. childcare, activities)
- d. Free breakfast to targeted children but a charge for other facilities
- e. Other (text response)

**19. Which children are defined as "targeted" in terms of your charging policy for breakfast during the school day?**

Text response

**20. Before the Magic Breakfast pilot could be implemented, did the school have to make any changes, in terms of the following? Please select all that apply.**

- a. Improvements to the physical environment where breakfasts will be received
- b. Increase the number of staff / adjust staff hours to cover breakfast delivery
- c. Training of staff
- d. Buying new furniture
- e. Improvement to catering facilities
- f. Improvement to food storage facilities
- g. Accessing additional funding
- h. Other (please specify)

**21. Did you have to spend additional money on FURNITURE in order to successfully deliver the breakfast club?**

- a. Yes
- b. No

**22. Please estimate approximately how much you spent on furniture in order to successfully deliver the breakfast club.**

Text response

**23. Did you have to spend additional money on IMPROVEMENTS TO THE PHYSICAL ENVIRONMENT, e.g. buying new curtains or painting a room, in order to successfully deliver the breakfast club?**

- a. Yes
- b. No

**24. Please estimate approximately how much you spent on improvements to the physical environment in order to successfully deliver the breakfast club.**

Text response

**25. Did you have to spend additional money on CATERING FACILITIES, e.g. buying a freezer to store the food, in order to successfully deliver the breakfast club?**

- a. Yes  
b. No
- 26. Please estimate approximately how much you spent on catering facilities in order to successfully deliver the breakfast club.**  
Text response
- 27. Did you have to spend additional money on RESOURCES, e.g. plates/bowls, cutlery, activities, etc., in order to successfully deliver the breakfast club?**  
a. Yes  
b. No
- 28. Please estimate approximately how much you spent on resources in order to successfully deliver the breakfast club.**  
Text response
- 29. Did you have to spend additional money on STAFF RECRUITMENT AND TRAINING in order to successfully deliver the breakfast club?**  
a. Yes  
b. No
- 30. Please estimate approximately how much you spent on staff recruitment and training in order to successfully deliver the breakfast club.**  
Text response
- 31. Did you have to spend additional money on INCREASES IN STAFF TIME in order to successfully deliver the breakfast club?**  
a. Yes  
b. No
- 32. Please estimate approximately how much you spent on increases in staff time in order to successfully deliver the breakfast club.**  
Text response
- 33. Did you have to spend money on ADDITIONAL FOOD/DRINKS DURING BREAKFAST PROVISION in order to successfully deliver the breakfast club?**  
a. Yes  
b. No
- 34. Please estimate approximately how much you spent on additional food/drinks during breakfast provision in order to successfully deliver the breakfast club.**  
Text response
- 35. Did you have to spend additional money on something that has not previously been asked about in the questions concerning additional expenditure, in order to successfully deliver the breakfast club?**  
a. Yes  
b. No
- 36. Please estimate approximately how much you spent on things that were not covered in the questions regarding additional expenditure, in order to successfully deliver the breakfast club.**  
Text response
- 37. Is your breakfast club staffed/supervised by any CATERING STAFF?**  
a. Yes  
b. No
- 38. Please give the total number of hours spent staffing/supervising the club by catering staff.**  
Text response
- 39. Is your breakfast club staffed/supervised by any TEACHING ASSISTANTS?**  
a. Yes  
b. No
- 40. Please give the total number of hours spent staffing/supervising the club by teaching assistants.**  
Text response
- 41. Is your breakfast club staffed/supervised by any TEACHERS?**

- a. Yes
- b. No

**42. Please give the total number of hours spent staffing/supervising the breakfast club by teachers.**

Text response

**43. Is your breakfast club staffed/supervised by any PASTORAL STAFF, e.g. learning mentors, family support workers?**

- a. Yes
- b. No

**44. Please give the total number of hours spent staffing/supervising the breakfast club by pastoral staff.**

Text response

**45. Is your breakfast club staffed/supervised by any OFFICE STAFF, e.g. business managers?**

- a. Yes
- b. No

**46. Please give the total number of hours spent staffing/supervising the breakfast club by office staff.**

Text response

**47. Is your breakfast club staffed/supervised by any CARETAKING or MAINTENANCE STAFF?**

- a. Yes
- b. No

**48. Please give the total number of hours spent staffing/supervising the breakfast club by caretaking or maintenance staff.**

Text response

**49. Is your breakfast club staffed/supervised by any VOLUNTEERS?**

- a. Yes
- b. No

**50. Please give the total number of hours spent staffing/supervising the breakfast club by volunteers.**

Text response

**51. Is your breakfast club staffed/supervised by ANY OTHER TYPE OF STAFF not previously mentioned in the questions about staffing/supervision?**

- a. Yes
- b. No

**52. Please give the total number of hours spent staffing/supervising the breakfast club by any other type of staff not previously mentioned in the questions about staffing/supervision.**

Text response

**53. Some aspects of delivering breakfast clubs in schools can be more challenging than others. How well do you think the following aspects of delivery have worked in practice in your school during the Magic Breakfast pilot? [Monitoring and managing the supply and demand for breakfasts]**

- a. Very well
- b. Quite well
- c. Not very well
- d. Not well at all
- e. Not sure
- f. Not applicable

**54. Some aspects of delivering breakfast clubs in schools can be more challenging than others. How well do you think the following aspects of delivery have worked in practice in your school during the Magic Breakfast pilot? [Delivering breakfast: preparing and serving food]**

- a. Very well



- b. Quite well
  - c. Not very well
  - d. Not well at all
  - e. Not sure
  - f. Not applicable
- 55. Some aspects of delivering breakfast clubs in schools can be more challenging than others. How well do you think the following aspects of delivery have worked in practice in your school during the Magic Breakfast pilot? [Providing a suitable, safe and welcoming environment where the breakfast takes place]**
- a. Very well
  - b. Quite well
  - c. Not very well
  - d. Not well at all
  - e. Not sure
  - f. Not applicable
- 56. Some aspects of delivering breakfast clubs in schools can be more challenging than others. How well do you think the following aspects of delivery have worked in practice in your school during the Magic Breakfast pilot? [Supervision of children]**
- a. Very well
  - b. Quite well
  - c. Not very well
  - d. Not well at all
  - e. Not sure
  - f. Not applicable
- 57. Some aspects of delivering breakfast clubs in schools can be more challenging than others. How well do you think the following aspects of delivery have worked in practice in your school during the Magic Breakfast pilot? [Clearing up after breakfast club]**
- a. Very well
  - b. Quite well
  - c. Not very well
  - d. Not well at all
  - e. Not sure
  - f. Not applicable
- 58. Some aspects of delivering breakfast clubs in schools can be more challenging than others. How well do you think the following aspects of delivery have worked in practice in your school during the Magic Breakfast pilot? [Developing a sustainable breakfast club for the next academic year]**
- a. Very well
  - b. Quite well
  - c. Not very well
  - d. Not well at all
  - e. Not sure
  - f. Not applicable
- 59. How successfully do you think the following aspects of engagement and buy-in have been achieved in your school during the Magic Breakfast pilot? [Informing and engaging children about breakfast provision]**
- a. Very successfully
  - b. Quite successfully
  - c. Not very successfully
  - d. Not at all successfully
  - e. Not sure
  - f. Not applicable

- 60. How successfully do you think the following aspects of engagement and buy-in have been achieved in your school during the Magic Breakfast pilot? [Informing and engaging parents about breakfast]**
- Very successfully
  - Quite successfully
  - Not very successfully
  - Not at all successfully
  - Not sure
  - Not applicable
- 61. How successfully do you think the following aspects of engagement and buy-in have been achieved in your school during the Magic Breakfast pilot? [Targeting those children who need breakfast the most]**
- Very successfully
  - Quite successfully
  - Not very successfully
  - Not at all successfully
  - Not sure
  - Not applicable
- 62. How successfully do you think the following aspects of engagement and buy-in have been achieved in your school during the Magic Breakfast pilot? [Gaining buy-in from staff as to the importance of providing free breakfast provision]**
- Very successfully
  - Quite successfully
  - Not very successfully
  - Not at all successfully
  - Not sure
  - Not applicable
- 63. Thinking about all the behaviour you have encountered around the school over the last term, how many pupils do you find generally badly behaved and/or difficult to deal with?**
- All / almost all
  - Most
  - Some
  - Few
  - None / almost none
- 64. In your experience, when has indiscipline been most likely to occur in your school over the last term?**
- Before the school day begins
  - During morning classes
  - Morning break
  - Lunchtime
  - During afternoon classes
  - Afternoon break
  - After the end of the school day
- 65. Since the Magic Breakfast pilot started, what do you think has happened to the following aspects of school life? [Pupil behaviour]**
- Improved a lot
  - Improved a bit
  - Stayed the same
  - Got a little worse
  - Got a lot worse

- f. Not sure
- 66. Since the Magic Breakfast pilot started, what do you think has happened to the following aspects of school life? [Pupil concentration]**
- Improved a lot
  - Improved a bit
  - Stayed the same
  - Got a little worse
  - Got a lot worse
  - Not sure
- 67. Since the Magic Breakfast pilot started, what do you think has happened to the following aspects of school life? [Pupil attainment]**
- Improved a lot
  - Improved a bit
  - Stayed the same
  - Got a little worse
  - Got a lot worse
  - Not sure
- 68. Since the Magic Breakfast pilot started, what do you think has happened to the following aspects of school life? [Pupil attendance]**
- Improved a lot
  - Improved a bit
  - Stayed the same
  - Got a little worse
  - Got a lot worse
  - Not sure
- 69. Since the Magic Breakfast pilot started, what do you think has happened to the following aspects of school life? [Pupil cohesion / collegiality]**
- Improved a lot
  - Improved a bit
  - Stayed the same
  - Got a little worse
  - Got a lot worse
  - Not sure
- 70. Since the Magic Breakfast pilot started, what do you think has happened to the following aspects of school life? [The school facilities (e.g. catering facilities or dining area)]**
- Improved a lot
  - Improved a bit
  - Stayed the same
  - Got a little worse
  - Got a lot worse
  - Not sure
- 71. Since the Magic Breakfast pilot started, what do you think has happened to the following aspects of school life? [Other (if no other aspect is relevant, please select 'Not sure')]**
- Improved a lot
  - Improved a bit
  - Stayed the same
  - Got a little worse

- e. Got a lot worse
- f. Not sure

**72. If you selected a response other than 'Not sure' against 'Other' in the above question, please specify the aspect of school life that you have responded about.**

Text response

**73. How likely is it that the school will continue with free breakfast provision once the support from Magic Breakfast ends?**

- a. Very likely
- b. Quite likely
- c. Neither likely nor unlikely
- d. Quite unlikely
- e. Very unlikely
- f. Not sure

**74. Would the school change any aspects of breakfast club delivery next year?**

- a. Yes
- b. No
- c. Not sure

**75. Please specify the aspect(s) of breakfast club delivery that you would change.**

Text response

**76. In your view, what aspects, if any, do schools most need advice and support on to make set-up and implementation of breakfast clubs successful? Please select all that apply.**

- a. Developing menus to provide nutritious and appealing food that is practical and affordable to deliver
- b. Catering facilities
- c. Sourcing food
- d. Monitoring and managing the supply and demand for breakfasts
- e. Staffing
- f. Preparing and delivering the breakfasts
- g. Providing a suitable, safe, welcoming environment where breakfast takes place
- h. Supervising children
- i. Informing and engaging children about breakfast provision
- j. Informing and engaging parents about breakfast provision
- k. Encouraging attendance at the club by difficult to reach and vulnerable children
- l. Thinking through different models of offering food and which would fit best with the school's situation
- m. Deciding on a charging policy and evaluating the implications of that decision
- n. None of these

**77. Was breakfast offered to all Year 6 students at school during this year's SATs week?**

- a. Yes
- b. No

**78. Is this the first year you have offered breakfast to Year 6 students during SATs week, or have you also done so in previous years?**

- a. First year
- b. Done previously

## Appendix E: Stata code for randomisation

```
// set seed
set seed 77200848

cap drop _u
cap drop _treat

// gen random number by stratification variable (strat3)
qui bys strat3: gen _u = uniform()

// number of schools within strata
qui bys strat3: gen N_strat = _N

// sorted position within strata
sort strat3 _u

qui bys strat3: gen n_strat = _n

// gen whether odd number within strata
qui gen odd_N_strat = mod(N_strat,2) == 1

// assign randomly selected half to the treatment group
qui gen _treat = n_strat <= (N_strat / 2)

// randomly assign the last school in a strata with an odd number
qui sum _u if odd_N_strat == 1 & n_strat == N_strat, de

local med_u = r(mean)

qui replace _treat = 1 if _u < `med_u' & odd_N_strat == 1 & n_strat
== N_strat

qui replace _treat = 0 if _u >= `med_u' & odd_N_strat == 1 & n_strat
== N_strat
```

## Appendix F: Distribution of attainment for those with and without demographic information in NPD

**Table AF1: Attainment data for those with and without full demographic information in the NPD**

| Outcome            | Has Demographics |       | Missing Demographics |        | Equality of Means |          |
|--------------------|------------------|-------|----------------------|--------|-------------------|----------|
|                    | n                | mean  | n                    | mean   | Difference        | p-value  |
| <b>KS1 Reading</b> | 4586             | 0.145 | 62                   | -0.76  | 0.905             | 0.000*** |
| <b>KS1 Writing</b> | 4586             | 0.168 | 62                   | -0.713 | 0.881             | 0.000*** |
| <b>KS1 Maths</b>   | 4586             | 0.135 | 62                   | -0.601 | 0.735             | 0.000*** |
| <b>KS2 Reading</b> | 3907             | 0.018 | 65                   | -0.833 | 0.851             | 0.000*** |
| <b>KS2 Maths</b>   | 3907             | 0.017 | 65                   | -0.705 | 0.722             | 0.000*** |

*Stars for statistical significance: p <= 0.10 (\*) p <= 0.05 (\*\*) p <= 0.01 (\*\*\*)*  
*Demographic characteristics: Sex, ethnicity, ever FSM, English known or believed to be an additional language, any Special Educational Needs.*  
*Outcomes are standardised point scores.*  
*The sample is pupils who would be in the analysis sample if the demographic requirement were removed.*

Table AF1 indicates that there are large and significant differences in attainment between pupils who do and do not have a full set of demographic data recorded in the NPD, with those who have a full record significantly outperforming the students with missing information. This is potentially concerning, since we are excluding a non-random subsample from the analysis based on the definition of our sample. However, due to the very small proportion of the sample excluded (1.3% at KS1, 1.6% at KS2) and to the robustness of our results to various alternative sample specifications (see Table 7), we judged that the risk of introducing bias was small relative to the rewards of controlling for another source of bias, the lack of balance on some demographic characteristics. For this reason we imposed the demographic restriction and controlled for demographic traits in our main analysis.

## Appendix G: Achieved sample for case studies

Table AG1: Achieved sample for case study schools

| School characteristic                                 | Value            | Number of schools |
|---|------------------|-------------------|
| School Size   | Small            | 2                 |
|   | Medium           | 1                 |
|   | Large            | 1                 |
| School Type   | Community        | 2                 |
|   | Academy          | 2                 |
| School Area   | Rural            | 1                 |
|   | Urban            | 3                 |
| Age Range   | 3-7              | 1                 |
|   | 3-11             | 3                 |
| English as another language percentage <sup>1</sup>   | Low              | 1                 |
|   | Medium           | 1                 |
|   | High             | 2                 |
| Free school meals eligibility percentage <sup>2</sup> | Low              | 1                 |
|   | Medium           | 2                 |
|   | High             | 1                 |
| Start Date  | Autumn term 2014 | 3                 |
|   | Spring term 2015 | 1                 |
| Delivery model  | Breakfast club   | 3                 |
|   | Soft start       | 1                 |

Notes:

<sup>1</sup> This was used to represent the level of ethnic diversity within the school

<sup>2</sup> To be eligible for subsidised breakfast provision from Magic Breakfast, schools had to have 35 per cent of their pupils eligible for free school meals. As such, low is considered 35 – 40 per cent.

## Appendix H: Analysis code

This appendix provides the analysis code to create Table 5, the headline table of results. For the code behind the rest of the tables, please contact the authors directly.

### \*\*\* TABLE 5 - PRIMARY ACADEMIC OUTCOMES \*\*\*

\*\*\*\*\*

tempname file

file open `file' using `"\$ptables\Analysis 7 (NPD effect size)\reportable tables.csv", write text replace

// Headline results - tests

file write `file' "Table A: Headline academic outcomes" \_new

file write `file' " , Raw means, , , , Effect size, , " \_new

file write `file' " , Intervention group, , Control group, , , " \_new

file write `file' "Outcome, n (missing), Mean (95% CI), n (missing), Mean (95% CI), n in model (intervention; control), Effect size (95% CI), p-value" \_new

foreach yvar in ks1\_readpoints\_z ks1\_writpoints\_z ks1\_matpoints\_z ks2\_readfine\_z ks2\_matfine\_z {

    if "`yvar'" == "ks1\_readpoints\_z"      local name = "KS1 reading"

    if "`yvar'" == "ks1\_writpoints\_z"      local name = "KS1 writing"

    if "`yvar'" == "ks1\_matpoints\_z"      local name = "KS1 maths"

    if "`yvar'" == "ks2\_readfine\_z"      local name = "KS2 reading"

    if "`yvar'" == "ks2\_matfine\_z"      local name = "KS2 maths"

    if "`yvar'" == "ks1\_readpoints\_z" | "`yvar'" == "ks1\_writpoints\_z" | "`yvar'" == "ks1\_matpoints\_z" {

        local ctrlattain = "ifspattain"

        local stage "ks1"

    }

    if "`yvar'" == "ks2\_readfine\_z" | "`yvar'" == "ks2\_matfine\_z" {

        local ctrlattain = "iks1attain"

        local stage "ks2"

    }

    // Treatment Group raw means

    sum `yvar' if sample == 1 & hasnpd == 1 & \_treat == 1 & fullresults == 1 & hasdemo == 1 & stage == "`stage'"

    cii mean r(N) r(mean) r(sd)

        local TN = trim("": display %6.0f r(N)")

        local Tmean = trim("": display %6.3f r(mean)")



```

    local Tcilo = trim("": display %6.3f r(lb)"" )
    local Tcihi = trim("": display %6.3f r(ub)"" )

count if sample == 1 & hasnpd == 1 & stage == "`stage'" & _treat == 1
    local Tmiss_temp = r(N) - `TN'
    local Tmiss = trim("": display %6.0f `Tmiss_temp'"" )

// Control Group raw means
sum `yvar' if sample == 1 & hasnpd == 1 & _treat == 0 & fullresults == 1 & hasdemo
== 1 & stage == "`stage'"
cii mean r(N) r(mean) r(sd)
    local CN = trim("": display %6.0f r(N)"" )
    local Cmean = trim("": display %6.3f r(mean)"" )
    local Ccilo = trim("": display %6.3f r(lb)"" )
    local Ccihi = trim("": display %6.3f r(ub)"" )

count if sample == 1 & hasnpd == 1 & stage == "`stage'" & _treat == 0
    local Cmiss_temp = r(N) - `CN'
    local Cmiss = trim("": display %6.0f `Cmiss_temp'"" )

// Formatting
local paro (
local parc )
local semicn ;

local col2 `TN' `paro`Tmiss`parc'
local col3 `Tmean' `paro`Tcilo`semicn`Tcihi`parc'
local col4 `CN' `paro`Cmiss`parc'
local col5 `Cmean' `paro`Ccilo`semicn`Ccihi`parc'

// Number of schools
gen tempflagT = (sample == 1 & hasnpd == 1 & _treat == 1 & sample == 1 & stage ==
"`stage'" & fullresults == 1 & hasdemo == 1 & !missing(`yvar'))
    bys tempflagT urn: gen urnflagT = _n
    count if urnflagT == 1 & tempflagT == 1
    local NschIT = r(N)

gen tempflagC = (sample == 1 & hasnpd == 1 & _treat == 0 & sample == 1 & stage ==
"`stage'" & fullresults == 1 & hasdemo == 1 & !missing(`yvar'))
    bys tempflagC urn: gen urnflagC = _n
    count if urnflagC == 1 & tempflagC == 1
    local NschIC = r(N)

local Nschltot = `NschIT' + `NschIC'

local col6 `Nschltot' `paro`NschIT`semicn`NschIC`parc'
drop tempflagT tempflagC urnflagT urnflagC

// Effect size
reg `yvar' i._treat i.strat3 `ctrlattain' i.female i.everfsm_6_spr15 i.ethnicshort i.anysen
i.eal i.effectiveness_2014 imd2010rank i.urbrur2 census_totpupsendn ///
    if sample == 1 & stage == "`stage'" & fullresults == 1 & hasdemo == 1,
cluster(urn)

assert `Nschltot' == e(N_clust)

```

```

local Ecoef = trim("` : display %6.3f _b[1._treat]`")
local Ecilo = trim("` : display %6.3f _b[1._treat] - invttail(e(df_r),0.025) * _se[1._treat]`")
local Ecihi = trim("` : display %6.3f _b[1._treat] + invttail(e(df_r),0.025) * _se[1._treat]`")
local Epval = trim("` : display %6.3f 2*(1-normprob(abs(_b[1._treat]/_se[1._treat])))`")

    if      `Epval' <= 0.10 local sigstars = "*"

    if      `Epval' <= 0.05 local sigstars = "***"
    if      `Epval' <= 0.01 local sigstars = "****"
    if      `Epval' > 0.10 local sigstars = ""

// Formatting
local col7 `Ecoef' `parc' `Ecilo' `semicn' `Ecihi' `parc'
local col8 `Epval' `sigstars'

file write `file' "`name" " " `col2" " " `col3" " " `col4" " " `col5" " " `col6" " " `col7" " " `col8"
_new
}

```

## Appendix I: Pupil Hunger Sensitivity Analysis

The ‘analysis’ sample includes children with the relevant outcome variable, a baseline survey response, and a full set of both demographic characteristics and outcome variables in the NPD. The ‘no baseline’ sample relaxes these conditions by removing the requirement that the child has a linked baseline pupil survey. The ‘Universal & Free’ sample uses the analysis sample criteria but restricts the analysis to pupils in schools which implemented breakfast provision that was both universal and free of charge. This analysis was not specified in the evaluation protocol but has been conducted to investigate whether the results are being driven by sample choice. Overall, the results are very similar across the three samples, suggesting that they are driven by real effects rather than by the choices made in defining the analysis sample.

**Table AI1: Sensitivity analysis – breakfast consumption and hunger**

| Outcome                         | Sample           | Raw means          |                     |               |                     | Effect size                        |                      |          |
|---------------------------------|------------------|--------------------|---------------------|---------------|---------------------|------------------------------------|----------------------|----------|
|                                 |                  | Intervention group |                     | Control group |                     | n in model (intervention; control) | Effect size (95% CI) | p-value  |
| n (miss)                        | Mean (95% CI)    | n (miss)           | Mean (95% CI)       |               |                     |                                    |                      |          |
| <b>Ate breakfast</b>            | Analysis         | 1842 (2357)        | 0.902 (0.888;0.915) | 1531 (1839)   | 0.894 (0.878;0.909) | 71 (38;33)                         | 0.032 (-0.001;0.066) | 0.045**  |
|                                 | No baseline      | 2162 (2037)        | 0.907 (0.894;0.919) | 2117 (1253)   | 0.888 (0.875;0.901) | 77 (38;39)                         | 0.037 (0.006;0.068)  | 0.014**  |
|                                 | Universal & Free | 785 (3414)         | 0.878 (0.855;0.901) | 1531 (1839)   | 0.894 (0.878;0.909) | 49 (16;33)                         | 0.022 (-0.023;0.067) | 0.317    |
| <b>Breakfast at school</b>      | Analysis         | 1835 (2364)        | 0.236 (0.217;0.255) | 1526 (1844)   | 0.077 (0.063;0.090) | 71 (38;33)                         | 0.146 (0.073;0.220)  | 0.000*** |
|                                 | No baseline      | 2153 (2046)        | 0.230 (0.213;0.248) | 2114 (1256)   | 0.071 (0.060;0.082) | 77 (38;39)                         | 0.140 (0.073;0.207)  | 0.000*** |
|                                 | Universal & Free | 782 (3417)         | 0.320 (0.287;0.352) | 1526 (1844)   | 0.077 (0.063;0.090) | 49 (16;33)                         | 0.205 (0.088;0.321)  | 0.000*** |
| <b>Hungry at start of class</b> | Analysis         | 1602 (2597)        | 0.280 (0.258;0.302) | 1408 (1962)   | 0.249 (0.226;0.271) | 71 (38;33)                         | 0.032 (-0.022;0.085) | 0.219    |
|                                 | No baseline      | 1894 (2305)        | 0.280 (0.260;0.301) | 1941 (1429)   | 0.273 (0.253;0.292) | 77 (38;39)                         | 0.012 (-0.050;0.073) | 0.698    |
|                                 | Universal & Free | 686 (3513)         | 0.273 (0.239;0.306) | 1408 (1962)   | 0.249 (0.226;0.271) | 49 (16;33)                         | 0.040 (-0.021;0.101) | 0.172    |

Stars for statistical significance: p <= 0.10 (\*) p <= 0.05 (\*\*) p <= 0.01 (\*\*\*)

Reported effect sizes are average marginal effects following logistic regression.

Standard errors clustered at school level.

Analysis sample: Children with the relevant outcome variable, a baseline survey response, a full set of demographics in the NPD, and a full set of academic outcomes in the NPD.

No baseline: A larger sample that does not require (or control for) a baseline response to the pupil survey.

Universal & Free: Children in the analysis sample whose headteachers reported offering a universal, free breakfast club as part of the intervention.

Controls: randomisation strata; average of baseline responses to the three outcome questions; demographics (sex, ever FSM, ethnic group, SEN, EAL); pre-intervention school characteristics (Ofsted rating; IMD rank; urban-rural; class size).

## Appendix J: Cost distribution

Table AJ1 explores the variation in different types of cost across the entire distribution of costs. It shows that there are often substantial differences in per-pupil expenditure between the lowest- and highest-spending schools. For upfront costs, the mean is substantially larger than the median, indicating that a small number of schools spent a relatively large amount of money initially. However, the very similar mean and median for overall expenditure suggest that per-pupil costs are not particularly strongly skewed by a small number of outliers at one end of the distribution.

**Table AJ1: Summary statistics of average per-pupil costs**

| Cost type                    | Number of schools incurring | Mean          | Median        | Minimum      | Maximum       |
|------------------------------|-----------------------------|---------------|---------------|--------------|---------------|
| Furniture                    | 15                          | £0.71         | £0.00         | £0.00        | £6.85         |
| Physical environment         | 3                           | £0.20         | £0.00         | £0.00        | £6.20         |
| Catering facilities          | 34                          | £1.39         | £1.03         | £0.00        | £6.91         |
| Resources                    | 33                          | £0.75         | £0.43         | £0.00        | £4.11         |
| Staff training & recruitment | 18                          | £1.08         | £0.00         | £0.00        | £20.20        |
| Other                        | 2                           | £0.03         | £0.00         | £0.00        | £1.03         |
| <b>Upfront Costs</b>         | <b>41</b>                   | <b>£3.73</b>  | <b>£2.00</b>  | <b>£0.00</b> | <b>£22.39</b> |
| Food & drinks (Magic B'fast) | 44                          | £12.15        | £11.38        | £2.47        | £35.17        |
| Food & drinks (additional)   | 23                          | £1.48         | £0.46         | £0.00        | £13.83        |
| <b>Ongoing costs</b>         | <b>46</b>                   | <b>£12.97</b> | <b>£12.32</b> | <b>£1.12</b> | <b>£39.03</b> |
| <b>TOTAL</b>                 | <b>46</b>                   | <b>£16.71</b> | <b>£14.94</b> | <b>£2.47</b> | <b>£49.45</b> |

*These summary statistics are based on the average per-pupil cost of all respondent schools (rather than the total cost averaged across schools, then divided by the average pupil count). This means that figures will not correspond directly to Table 15.*

*Individual cost-type means do not add to totals and subtotals since totals and subtotals are calculated by first summing a school's expenditure in that category, then dividing by pupils, then averaging the result over respondent schools.*

*Costs are presented for a single school year.*

*Cost information for the food provided by Magic Breakfast was obtained directly from Magic Breakfast. Other cost information comes from the survey of headteachers.*

You may re-use this document/publication (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence v2.0.

To view this licence, visit [www.nationalarchives.gov.uk/doc/open-government-licence/version/2](http://www.nationalarchives.gov.uk/doc/open-government-licence/version/2) or email: [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk)

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned. The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

This document is available for download at [www.educationendowmentfoundation.org.uk](http://www.educationendowmentfoundation.org.uk)



Education  
Endowment  
Foundation

The Education Endowment Foundation  
9th Floor, Millbank Tower  
21–24 Millbank  
London  
SW1P 4QP  
[www.educationendowmentfoundation.org.uk](http://www.educationendowmentfoundation.org.uk)