SCOPING A FULL ECONOMIC IMPACT ANALYSIS

SHINE on Saturday

Claire Brinkman
September 2016

SHINE

PRO BONO ECONOMICS

ACKNOWLEDGEMENTS

THANK YOU…

The author would like to thank the following people for their generous time and assistance in carrying out this project:

- Federico Cilauro & Thomas Baily, Economic consultants, who performed all of the econometric analysis and helped identify issues with the various econometric models as well as calculating the minimum sample size to robustly identify the treatment effect – all in a voluntary capacity.
- Ron Smith, Professor of Applied Economics at Birkbeck, University of London, for his help and advice on the econometric analysis – also entirely in a voluntary capacity.
- Karen Hancock, Economist at PBE, for helping to think through some of the economic issues and putting me in touch with helpful contacts in relation to the literature review.
- Fiona Spellman, Senior Programme Manager at SHINE, for her ongoing help and assistance in carrying out this project and her continued patience when we encountered challenges along the way.
- The Deputy Head of the anonymous primary school who kindly helped us to obtain the data we needed about the relevant pupils at her school over the last 8 years.
- Chris Douglass, NPD team at Department for Education, for his help in understanding the requirements for requesting data from the NPD.
- Andy Cawthera, Freelance economic consultant (previously at EEF), for his initial help with understanding SHINE's RCT and pointing me in the direction of the literature on non-cognitive skills.

CONTENTS

EXECUTIVE SUMMARY

KEY MESSAGES

- This report provides detailed guidance on how to perform a full economic impact study into the SHINE on Saturday programme. The economic impact of any such social intervention should be measured as the present value of all of the economic benefits generated as a result of the intervention, net of the costs associated with its provision.
- It is proposed that the main route through which the programme results in economic benefits is via the effect on pupils' non-cognitive skills[1]. According to the work of economist Jim Heckman, interventions which improve children's non-cognitive skills, unlike those that simply target cognitive skills, are far more likely to have a lasting impact on academic attainment.
- The initial findings from a pilot study into the effect of SHINE on Saturday on pupils' academic performance, based on data from one London school, seems promising. On average, the intervention appears to have improved KS2 results by 0.22 of one level, relative to comparable pupils who did not attend the programme. It is estimated that those pupils would have gone on to gain 1.4 additional GCSEs at A*-C, on average, compared to their counterparts. Based on research by the DfE, this can be translated into an improvement in economic productivity (in the form of greater lifetime earnings) of approximately £6.3 million across all pupils who had received the programme at this school from 2005/06 to 2013/14. However, it must be emphasized that this result is specific to this individual school, does not allow for the costs of provision and assumes that we have captured all relevant explanatory factors. Further research is now needed, using significantly more data drawn from multiple schools.
- In terms of next steps, the most time efficient approach would be to fill gaps in existing data SHINE holds from other London based primary schools that have offered the SHINE on Saturday programme, to develop a broader dataset. The analysis should then be rerun to identify the impact on KS2 results and expected KS4 results which can then be linked to potential future earnings. In addition, the costs incurred by SHINE, the schools, teachers and parents need to be measured. This will allow the *net* economic benefit of the intervention to be more accurately assessed.

[1] Non-cognitive skills includes "softer" skills, attitudes and strategies which can be vital for success in learning and life in general, Examples include motivation, time management, problem solving abilities, etc. Cognitive skills on the other hand would include, for example, numeracy and literacy.

SHINE (Support and Help in Education) is a charity which supports educational initiatives that encourage children and young people to raise their academic attainment. In 2013, SHINE approached Pro Bono Economics to help determine whether a full economic impact analysis could be performed based on the outcomes of an RCT (randomised controlled trial) which it had recently commissioned. This RCT related to one of its programmes "Hallé SHINE" which was launched in Manchester in 2012. If so, the charity wanted to understand what such an analysis would entail and how much data would be required. After looking into the RCT in more detail, it became clear that the results would not be suitable for these purposes. Nevertheless, historic data that the charity had been collecting from the longer running London based version of this programme, "SHINE on Saturday", could be used to perform an ex-post analysis instead.

The basic format of an economic impact analysis is to calculate the present value of all of the economic and social benefits generated by the intervention and offset this by the present value of all of the economic and social costs incurred as a result of providing the intervention. As explained below, the discount rate proposed by the Treasury for assessing public policy projects, is commonly used when performing such analysis for charities. This would therefore be applied to convert the benefits and costs into a present value. The first question is therefore what economic benefits does this intervention generate and via what mechanisms and secondly, what costs are incurred as a result?

Based on anecdotal evidence from schools and teachers, as well as evidence collected from a survey of SPMs which was run as part of this study, it seems likely that SHINE on Saturday helps children predominantly by enhancing their non-cognitive skills. According to the research of Heckman and others, improvements in children's non-cognitive skills appear to generate a lasting impact in terms of academic performance and subsequent labour market outcomes. Therefore, the underlying assumption taken in this analysis is that the programme is primarily helping pupils via this mechanism. This leads (both directly and indirectly) to an improvement in KS2 results, relative to the "counterfactual" (which can be measured by considering the performance of pupils with similar characteristics, who did not receive the intervention). This in turn leads to an improvement in KS4 (GCSE) results than would otherwise occur and therefore greater economic productivity which is proxied by lifetime earnings.

KS2 data can be obtained from the schools SHINE has worked with and the charity has also collected other pupil characteristic data which can help to explain pupils' academic performance. Therefore, using an econometric model, it should be possible to identify firstly the incremental effect of SHINE on Saturday on pupils' KS2 results (i.e. by comparing the treatment and comparator group). Although KS2 data can be obtained from the schools SHINE has worked with, subsequent KS4 data is more difficult to get hold of. Although, in theory, this is available from the National Pupil Database, due to data protection laws, unless parental consent has been obtained, such data cannot be released. Therefore, an alternative approach is presented in this report, using conditional probabilities of different KS4 results (depending on KS2 results and various disadvantage indicators) to produce expected (probability-weighted) KS4 results for different sub-sets of pupils. The incremental impact on KS4 results is then traced through to additional productivity, using the results of an academic paper by the Department for Education.

Using data from one school we ran a pilot of the KS2 impact analysis, in order to estimate the minimum sample size for performing a full economic impact study. This indicated that at a bare minimum, approx. 160 pupils should be sampled. However, following sensitivity analysis, it became clear that a sample of approximately 2000 – 2200 would be a more appropriate aim. In addition, it would be wise to spread this data collection evenly across multiple schools – 10-12 if possible. Based on information provided by the charity, it seems likely that they would be able to obtain this number of observations relatively easily. As the pilot included 257 observations, the impact of the intervention calculated was statistically significant at the 95% level[1]. It was therefore possible to conclude, that in this specific school, the intervention, on average, had improved KS2 results by 0.22 of one level. This is associated with an additional 1.4 GCSEs at A*-C, relative to the average achievement of 4 GCSEs at A*-C for those who did not receive the intervention. Applying research by DfE, it is estimated that this would have translated into an improvement in economic productivity[2] of approximately £6.3 million across all pupils who had received the programme at this school over the period 2005/06 – 2013/14. However, due to the fact that only one school was used in this analysis, it cannot be assumed that these results apply across all the schools SHINE has worked with. In addition, this is based on a minimal sized sample, and with more data the accuracy could be improved[3].

The next step is to consider the economic and social costs which must be offset against these benefits. Three main types of costs are identified and methods for calculating them are proposed. Firstly, the relevant share of SHINE's internal costs which are attributable to running this programme should be taken into account. The most appropriate method would be to use a "fully-allocated-costing" (FAC) approach and appropriate cost drivers. Secondly, there are costs for the schools and teachers which are involved in offering this programme. For example, the additional pay for teachers & SPMs running the Saturday school and the costs of arts materials & trips. Although

---

[1] This means that we can have confidence that the results drawn from this sample are indicative of the results for all children who received the intervention at that specific school and that the impact is positive.

[2] Measured in the form of increased lifetime earnings of those receiving the intervention.

[3] For example, based on our analysis of this sample, we can be 95% confident that the true effect size (on KS2 results) fell somewhere between 0.04 & 0.39, which is a reasonably large range. If the real result was different from the mean result (0.22), this would have significant implications for the effect on KS4 results and lifetime earnings referred to above. Therefore these tentative results should be treated with care.

the exact amounts spent will vary from school to school and year to year, a reasonable estimate can be obtained from the annual budgets which the schools submit to SHINE as part of their grant applications. In addition, there are more implicit costs which must be captured, such as a share of the pay of the headteachers of those schools hosting SHINE on Saturday, for the time spent managing the programme. This can be estimated bottom-up using data published by DfE on average headteachers' salaries. A share of the running costs of a primary school should also be captured, to reflect the additional time it must be opened on a Saturday. Again this can be estimated bottom-up using DfE data on primary school running costs. Finally, there is the opportunity cost of the time parents must spend transporting their children to and from the Saturday school, which can be measured based on the Department for Transport's estimates of the value of time spent travelling.

Once the gross economic and social benefits and gross economic and social costs have been estimated, the net economic impact can then be calculated. This is done by discounting / compounding the annual net benefits and costs to determine the present value of the overall net economic impact of the intervention. Given the inherent uncertainty in trying to accurately predict the benefits and costs of any such intervention, it is important to carry out sensitivity analysis around this and therefore generate a range of potential outcomes. Another useful technique is to consider whether there is likely to be a reasonable margin between the present value of the benefits (which can be harder to estimate accurately) and the present value of the costs.

The next steps for SHINE, if they decide that they want to initiate a full economic impact study, would be to decide how to estimate the economic benefits. One option would be to use existing historic data combined with the conditional KS4 probabilities set out in Appendix 5 of this report; the other would be to collect new data and subsequently obtain these pupils' actual KS4 results. Under the first option, the KS2 results and any other gaps in the existing data would need to be filled by the relevant schools. Under the second, consent from the parents of the pupils sampled would need to be collected to ensure that their KS4 results could be subsequently obtained from the National Pupil Database. However, due to the time which would need to elapse before the pupils reached KS4, taking such an approach would mean that the results of the study would not be available until at least 2023.

Economic impact analysis has a wide range of potential uses for a charity. It can be very helpful when seeking new sources of funding or maintaining existing sources. This can be particularly true in relation to funding provided by local or national government. Other stakeholders can also find this analysis useful, including the current and potential future beneficiaries, which in this case would include the primary schools offering the programme and the associated local education authorities. However, it can also be of great benefit internally, in terms of helping the charity really focus the use of its resources, support strategic decision making and consequently improve effectiveness.

INTRODUCTION

SHINE (Support and Help in Education) was set up in 1999 to support educational initiatives which encourage children and young people to raise their academic attainment. The majority of SHINE's funding goes directly to schools to support additional learning, at the weekends, in the holidays and after school. SHINE on Saturday is the most significant of the programmes facilitated by the charity and was first trialled at a primary school in Hackney in 2001. Since then a further 46 projects have been launched at schools across London[4] and more recently in Manchester. The programme is targeted at primary school pupils who are in some way disadvantaged and at risk of failing to fulfil their potential. The hypothesis is that by providing them with this additional Saturday curriculum, which is based on a much more creative learning approach, the children who attend will perform better during weekday schooling. The annual programme runs for between 25 and 30 Saturdays and is part funded by SHINE and the participating schools[5]. SHINE also provides detailed instructions on how to run the programme and they review and sign off on the school's curriculum plan. In 2012, a new version of SHINE on Saturday was set up in Manchester in collaboration with the Hallé Orchestra. The curriculum for this programme is centred on musical themes and enrichment activities but otherwise the underlying principles remained unchanged.

SHINE approached Pro Bono Economics to help determine whether a full economic impact analysis could be performed based on the outcomes of an RCT (randomised controlled trial) which it had commissioned in relation to the new Hallé SHINE programme. Having looked into the RCT in more detail, it became clear that it would not be suitable for that purpose. Instead, after discussing with SHINE the detailed monitoring & evaluation data that they have collected over the years from the London-based SHINE on Saturday projects, I proposed that this data could form a more useful basis for such an analysis.

The objective of this report is therefore to explain in some detail the steps required to run a full economic impact analysis on the London-based SHINE on Saturday programme, including a discussion of the relevant data sources. In addition, calculations have been performed to determine an appropriate sample size to obtain a reasonably robust result. This should help in ascertaining the potential scale and timeframe of the project and hence the likely resources it would require.

The report is structured into five main sections. The first discusses the overall approach to performing an economic impact analysis and explains at a high level what would be required in this specific case. This is followed by a section focusing on the economic benefits which describes, in some detail, how these could be measured. It also includes a calculation of the sample size required to perform the necessary robust regression analysis which would underlie this. The next section focusses on the costs associated with running SHINE on Saturday and discusses the different elements of the full economic & social costs and how these could be measured. The calculation and interpretation of the overall net economic impact is then discussed. Finally, a set of next steps are proposed which SHINE would need to take if it wanted to pursue a full economic impact analysis.

APPROACH

In this section an overview of the broad approach to measuring the economic impact of a charitable intervention is outlined. In addition, a "theory of change" relating specifically to SHINE on Saturday is proposed, to help understand the elements of the full economic impact of this programme.

OVERVIEW OF ECONOMIC IMPACT

The aim of an economic impact analysis in the context of the third sector is to compare the economic & social costs and benefits associated with the provision of a particular intervention and therefore assess how "effective" that intervention is. This requires consideration of the "counterfactual" scenario – i.e. in order to isolate the

[4] Note that each project normally involves multiple schools.

[5] Historically, schools received full funding from SHINE for the first few years, but more recently SHINE has moved to a co-funding model with schools and other partners. In addition, with one exception, SHINE continues to provide some level of partial funding to all of the schools that continue to run SHINE on Saturday.

incremental benefits from an intervention, it is necessary to consider what would happen in a state of the world where everything remained the same, but the intervention had not been provided. The difference between the benefits generated under the actual and the counterfactual scenarios is therefore the economic benefit which is attributable to the intervention alone. Since the counterfactual scenario is assumed to generate zero costs (as the intervention is not provided), the costs of the intervention can be measured more directly.

In order to make such a comparison between the economic & social costs and benefits associated with an intervention, it is important that adjustments are made to allow for differences in timing. The method commonly used for this purpose is "net present value" (NPV) which makes the comparison more economically meaningful. In the private sector, many organisations will use an NPV approach in deciding what activities to invest in. Only if the net present value of the cash flows associated with that activity exceed zero, is investment considered to be worthwhile – i.e. the firm will at least break-even on that particular project. The diagram below depicts the NPV calculation (see Figure 1) - the costs and benefits from year 1 onwards have to be "discounted" to the present value, which means reducing the value by taking into account the "time value of money". In other words, the value today of a payment or receipt which is expected in the future is lower than its actual nominal value.

Assume that the intervention is run for just three consecutive years in a school (i.e. years 0, 1 & 2) for a group of year 4, 5 & 6 pupils. In the first three years, the intervention just generates costs, but after seven years, economic benefits start to flow – i.e. those pupils who were in year 6 at the start of the period will have finished their GCSEs and, we presume, entered the workforce (for the purpose of exposition). The NPV would then be the sum of the discounted value of both the costs and the benefits which would consequently be shown at today's value. This scenario is depicted in Figure 1 below.

The "discount factors" used to adjust the costs and benefits (or in- and outflows) to their "present" value, in order to be able to compare them needs to be appropriate. Note that the generic term for a discount factor is as follows, where r = the discount rate:

$$Discount\ factor\ at\ time\ 't' = \frac{1}{(1+\ r)^t}$$

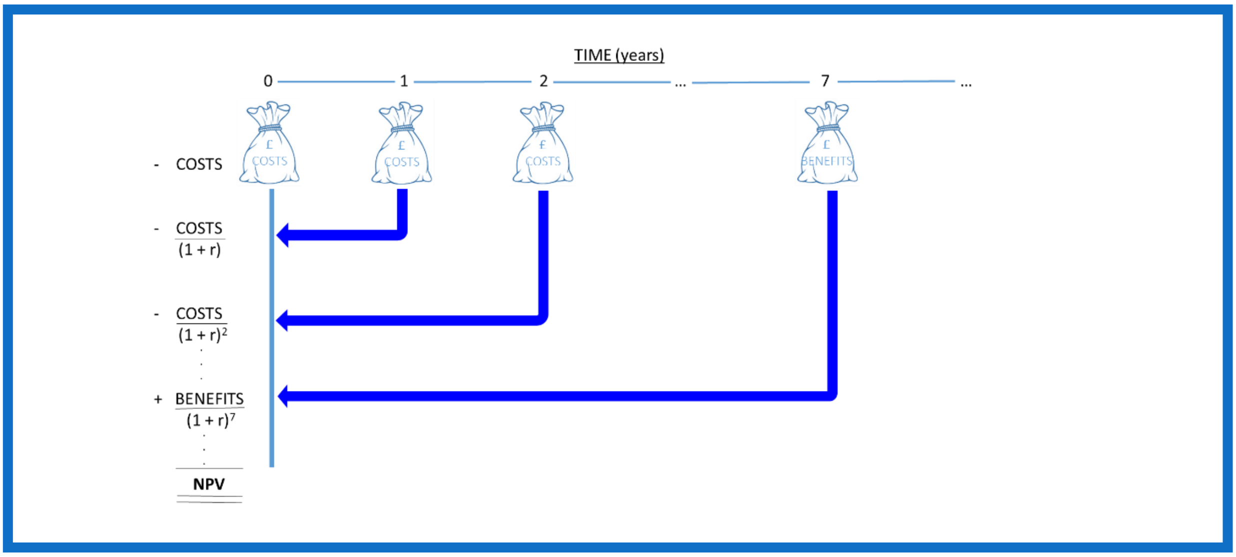

Figure 1: NPV calculation diagram

A common method for determining an appropriate value for "r" the discount rate, or "time value of money" is to use that rate applied by the government in public policy assessment. The "Green Book" produced by the Treasury sets out how projects should be evaluated and appraised within central government and specifies the "social time preference rate" which should be used for discounting benefits and costs in order to trade-off the value society attaches to present as opposed to future consumption. Currently, this is set at 3.5% (other than for periods beyond

30 years) for all public sector projects – see "The Green Book - Appraisal and Evaluation in Central Government"[6]. Therefore, unless the Green Book is subsequently updated, it would seem appropriate to apply a discount rate of 3.5% to this project.

THEORY OF CHANGE

Below we set out the hypothetical theory of change which sits behind our analysis. This is helpful in trying to identify all of the economic and social costs and benefits associated with SHINE on Saturday. Note that this is presented in a fairly simplistic way, but the routes via which children are affected by the intervention could be more complex.

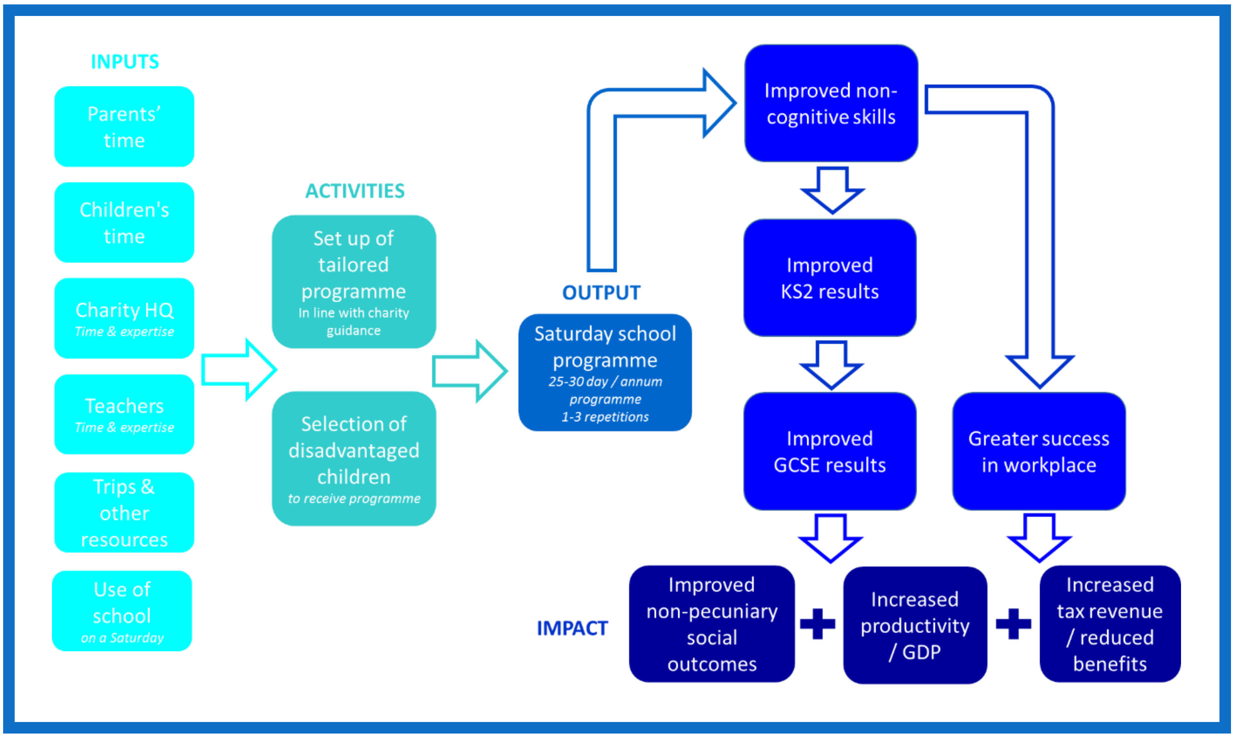

Figure 2: Hypothetical theory of change for SHINE on Saturday

As shown, there are a range of inputs (which each has its own associated costs), including both the parents' and pupils' time; the time and expertise of the teachers who run the Saturday school; use of the school on a Saturday; and all of the materials used in delivering the curriculum and any off-site trips. In addition, there is the work performed by the charity's head office which contributes to the running of this intervention. The initial activities required to prepare for delivery of the intervention include designing the tailored Saturday school curriculum which is developed by the schools' teachers and then signed off by SHINE. They also have to select the disadvantaged pupils at their school to whom the intervention will be offered, which again is reviewed & challenged by SHINE. For the children who take up this offer and attend the annual SHINE on Saturday programme (which historically has been between one & three times during their primary school career) it seems very likely an improvement in their non-cognitive skills occurs and it is this which results in their improved performance during weekday school (the evidence for this is discussed more in the next section below). Consequently, their KS2 results improve (relative to the counterfactual of not attending the programme) and subsequently, their GCSE (KS4) results are also better than they would otherwise have been. Given both the direct and more indirect benefits of improved non-cognitive skills, the overall impact is predominantly increased economic productivity and GDP. This also leads to greater tax revenue (and reduced benefit claims) along with other improved social outcomes such as reduced crime & inequality, along with improved health & wellbeing.

The benefits and costs associated with the SHINE on Saturday programme are discussed in greater detail below.

[6] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220541/green_book_complete.pdf

BENEFITS

This section begins by explaining why the RCT which SHINE had recently commissioned at the start of this project was considered unsuitable for determining the economic impact of the SHINE on Saturday programme. It then goes on to describe an alternative ex-post framework for understanding the benefits of the intervention, using existing historic data. This includes linking the potential improvement in academic performance generated by the intervention through to an increase in predicted lifetime earnings. Details are then given of the calculations which have been performed to determine the minimum sample size required to identify the economic benefits of the intervention under a full economic impact analysis. Finally, the steps required to estimate the total benefits of SHINE on Saturday over a specific period using an econometric model, in conjunction with information on the necessary sample size are described.

WHY THE RCT IS NOT SUITABLE FOR OUR PURPOSES

CEM (Centre for Evaluation & Monitoring) at Durham University have recently carried out a full, long-term RCT on the SHINE Hallé programme in Manchester, with funding from The Education Endowment Fund (EEF). SHINE wondered if the results of this test could be used to determine the economic benefits of their Saturday School intervention. After reviewing the detailed plans for this study, I concluded that it would not be suitable. The reasons for this are set out below.

USE OF THE INCAS TOOL

CEM have used their own "InCAS" computerised assessment tool to measure the effect of the Hallé SHINE programme on the pupils' literacy and numeracy. Currently, there is no detailed understanding of the association between the outcomes from InCAS assessments for children of different ages and their subsequent KS2 or GCSE results. In order to be able to link the intervention with economic impact, it must be possible to trace the immediate effect through to later academic attainment. Given these limitations surrounding InCAS, it would therefore be difficult to use the outcomes from the RCT for this purpose.

RISK OF FADE-OUT EFFECT

The RCT is designed to assess the effect of one year of the Hallé SHINE programme on literacy & numeracy. However, given that we can't be completely sure at this stage if and how the intervention is improving academic performance, it seems possible that there could be a risk of "fade-out" effect on these measures of performance. This means that a one off experience of the programme may not have a lasting impact on these specific cognitive skills. This is based on the academic literature which suggests that only repeated cognitive focused interventions or interventions that target the development of non-cognitive skills over cognitive skills will tend to have a more enduring effect.

EFFECT OF IMPROVING NON-COGNITIVE VS. COGNITIVE SKILLS

The Nobel Prize winning economist Jim Heckman has been investigating the benefits of developing disadvantaged children's non-cognitive skills for many years[1]. His research found that, based on various long term studies, following young American children into adulthood, the cognitive benefits of early intervention programmes (such as "Head start" & the "Perry pre-school programme") soon faded away. On the other hand improvements in what he termed "character skills" as a result of these interventions, were more long-lasting and explained a significant part (30-40%) of subsequent performance in academic achievement tests.

A paper looking into which literacy programmes were most effective in helping pupils of primary age that were struggling with their reading, concluded that intervention over many years was necessary for any lasting impact[2]. It was found that brief, intensive tutoring interventions were unlikely to have a longer term effect. This paper reviewed 96 other studies into methods of improving literacy and therefore seems to support the idea that interventions which purely target the improvement of cognitive skills tend to exhibit some "fade out effect".

A more recent study from the UK identifies the link between higher levels of emotional, behavioural, social and school wellbeing (non-cognitive skills by another name) and higher academic attainment at the time of measurement and in later years as well as improved school engagement. This highlights the feedback mechanism from non-cognitive to cognitive skills[3].

[1] Heckman J., Kautz T. (2013), 'Fostering and measuring skills: Interventions that improve character & cognition', NBER Working paper no. 19656

[2] Slavin R., Lake C., Davis S. & Madden N.A. (2009), 'Effective programs for struggling readers: A best evidence synthesis', Best Evidence Encyclopedia

[3] Morrison Gutman L. & Vorhaus J. (2012), 'The impact of pupil behaviour and wellbeing on educational outcomes', Childhood wellbeing research centre, Institute of Education, University of London

Even if the results of the RCT imply that there is an immediate benefit, it is difficult to know whether this will be maintained or not. If the long-term effect of the programme via the route of a pupil's cognitive skills is minimal, then there cannot be a wider economic impact as our model relies on the intervention affecting longer term academic attainment and earning potential[7].

MORE BESPOKE PROGRAMME?

The Manchester programme is being run in conjunction with the Hallé Orchestra and focuses on music as the main topic. There could be a question over the relevance of this programme for understanding the more generic version of SHINE on Saturday which has been run in London for almost a decade, where children learn about a wide variety of topics. However, it could also be argued that the subject specific content is less important than the general teaching approach which all schools running SHINE on Saturday must apply.

RCT COULD NOT BE CHANGED / ADAPTED AT THIS STAGE

The RCT had been carefully designed by CEM to ensure that the experiment was "controlled" and therefore it was not possible at this stage to amend or adapt it in any way. This meant that the number of times the pupils receive the Saturday school programme was fixed and could not be increased.

[7] It should be noted that The Educational Endowment Foundation (EEF) are planning to do a long-term follow up on a number of interventions that they are involved with, including the Hallé SHINE programme (i.e. the RCT). In six years' time, they will be able to collect data on the GCSE results of the pupils in both the treatment and control groups to assess whether there is a significant difference between them and hence whether a single experience of the programme appears to improve performance at GCSE level – i.e. whether the intervention has a lasting effect.

The other problem with the RCT was that we would not have been permitted to survey the pupils, parents or teachers taking part. This is important to determine the time they have had to give up due to their involvement in the Hallé SHINE programme and the associated opportunity costs. This would have fed into the assessment of the full costs of running the programme.

For all these reasons, rather than relying on the outcomes from the RCT, an alternative approach was considered, using historic data that SHINE has collected from the SHINE on Saturday programmes run in primary schools in London[8]. This ex-post framework is described in more detail below.

PROPOSED EX-POST FRAMEWORK FOR UNDERSTANDING THE BENEFITS

The proposed ex-post framework involves using historic data about pupils who received the intervention (termed the "treatment group"), collected from schools which have provided SHINE on Saturday, to retrospectively assess the effectiveness of the programme. A specially constructed comparator group is required to represent the counterfactual – i.e. the outcome, given the assumption that these pupils had not received the intervention. The difference in outcome between these two states of the world allows the effect of the intervention to be isolated. Although, in theory, a randomized control trial would be more robust as the comparator pupils would be selected in a more controlled fashion, the difficulties and costs associated with setting up RCTs means that in reality, ex-post analyses are a common method used by charities for measuring the impact of an intervention.

The amount of historic data available is potentially quite large as the SHINE on Saturday London based programme has been running at a reasonable scale for the last ten years across a total of approximately 150 primary schools[9]. SHINE estimate that if data was taken from 3-4 of the schools that have been participating the longest, a sample of approximately 1000 treated pupils could be obtained. In addition, some pupils have received the intervention on two or even three occasions during years 4 – 6, meaning that it should be possible to test the hypothesis that repeated interventions improve the longer-term academic benefits[10]. One potential concern is the completeness of the dataset which would be required for this analysis. In relation to the relevant metrics, there are some gaps or errors in what has been historically provided by the schools. While SHINE could go back to the schools to request more data, this could prove time-consuming.

As shown in the theory of change diagram, we are assuming (based on Heckman's research – see "Effect of improving non-cognitive vs. cognitive skills") that the intervention improves a pupil's non-cognitive skills and hence

[8] Note that the findings of the RCT were published by EEF in August 2016, just prior to the publication of this report. The main conclusion drawn from the RCT was that Hallé SHINE on Manchester had no impact on pupils' attainment in reading & maths, or their attitudes towards reading, maths, music or school in general. There are several differences between that study and the proposed analysis set out in this report and the potential resulting findings. First, the curricula of the two schemes differ. The Hallé SHINE on Manchester programme offers a music-based curriculum in conjunction with the Hallé orchestra, solely in schools in Manchester. Our analysis proposes to evaluate the SHINE on Saturday programme offered in London schools that have offered curricula covering a broader range of subjects. Second, the RCT used an "intention to treat" approach to control for potential selection bias in the pupils who took up the opportunity to attend Hallé SHINE on Manchester, which could influence their propensity to benefit from it. Therefore, all pupils who were offered a place on the programme under the RCT, regardless of whether they took it up, formed part of the treatment group. This approach is not feasible under the ex-post analysis recommended here, given the use of historic data. Finally, pupil uptake of Hallé SHINE on Manchester under the RCT was particularly low, as was the attendance of those who did attend. In part, this may have been due to the formulation of the trial, which prevented schools from approaching parents about this new programme, before the trial had officially begun. However, unlike the London programme, the fact that the Manchester programme wasn't yet well established would always have made recruiting pupils more challenging. Again, due to the ex-post nature of the analysis set out here, such problems should not arise in this case. In conclusion therefore, the two studies are estimating quite different effects. In the case of the RCT, the study attempts to determine the impact of offering a music-based Saturday school curriculum to pupils in schools in Manchester; in the case of our analysis, we are proposing to estimate the impact of taking part in a more established Saturday school programme with a broader curriculum in primary schools in London.

[9] The total number of discrete schools that have ever run a SHINE on Saturday project (i.e. for at least one year) is approximately 150, however, not all of these schools repeated the programme over a longer timeframe. For example, during the academic year 2015/16, there were 92 schools offering SHINE on Saturday.

[10] Based on our survey of School Project Managers, we found that almost 40% of those who responded felt that it was optimal for pupils to attend SHINE on Saturday twice (i.e. over two years) and the remaining 60% felt that it should be three times.

has a lasting effect on academic attainment relative to a situation without the intervention. The programme is provided to pupils in years 4, 5 or 6, so shortly before KS2. However, our theory is that due to the resulting improvement in non-cognitive skills, not only will they perform better at KS2 but also in their GCSE exams in year 11. Better academic attainment means improved employment prospects and greater value output for the economy, relative to the counterfactual. Therefore we need to be able to assess the incremental lifetime productivity of pupils who receive the intervention. Economic theory suggests that employers would be willing to pay no more to hire an employee than the value of the output they will generate. Therefore, we can proxy this additional output using the difference in the costs of hiring those from the treatment group (who are expected to earn more) and those from the comparator group. The cost of hiring is measured in terms of wages and other non-wage costs (i.e. NICs, pension contributions, etc.). Each of these steps in the cascade of benefits is discussed further below but first the role of non-cognitive skills is explored in more depth.

IMPROVEMENTS IN NON-COGNITIVE SKILLS

Non-cognitive skills and cognitive skills are defined as follows by The Institute of Education[11]:

> *'Non-cognitive skills', such as motivation, perseverance and self-control, are the attitudes, behaviours and strategies that are thought to underpin success in school and at work. They are usually contrasted with the 'hard skills' of cognitive ability in areas such as literacy and numeracy, which are measured by academic tests.*

As mentioned above, the academic literature suggests that to have a prolonged impact on a child's academic attainment, interventions which target non-cognitive skills are particularly successful.

The idea is that by improving a child's non-cognitive skills, they are being better prepared for learning and study – they obtain a skill which will stand them in good stead in a variety of circumstances, including educational ones. Attempts to just target cognitive skills (such as numeracy or literacy), without also improving non-cognitive skills may not have a prolonged effect. This can be likened to the saying "Give a man a fish and you feed him for a day. Teach a man to fish and you feed him for a lifetime". Unfortunately, it has been found that, relative to those born into more advantageous circumstances, children from more disadvantaged backgrounds will not acquire these non-cognitive skills so easily. Their luckier contemporaries will tend to pick up such skills in the home from their parents or carers and therefore are automatically at an advantage. However, as discussed below, this is not necessarily just down to growing up in a household with greater financial income.

[11] Morrison Guttman L., Schoon I., 'The impact of non-cognitive skills on outcomes for young people – Literature review', Institute of Education, 21 November 2013

THE ROLE OF HOME ENVIRONMENT ON NON-COGNITIVE SKILLS

A study from the LSE in 2007, found that both family background and home environment were extremely important for explaining the development of "social" (non-cognitive) skills. In terms of home environment, it was factors such as how much the parents (especially the father) reads each week; whether the parents show an interest in their children's education; whether there are serious difficulties (such as alcoholism, mental health issues & divorce) that were found to be important[1].

According to McLanahan, a sociologist, children in affluent homes are bathed in financial and cognitive resources. Those in less advantaged circumstances are much less likely to receive cognitive and socioemotional (or non-cognitive) stimulation and other family resources[2].

According to the child psychologist Rutter, a substantial body of evidence suggests that a major determinant of child disadvantage is the quality of the nurturing environment rather than just financial resources available or presence or absence of parents[3]

A DfE paper made use of data from the EPPSE 3-16 longitudinal study to explore reasons why some disadvantaged children still perform well academically, against the odds. In general, low socio-economic status had been found to be significantly related to child outcomes, but the home learning environment was also relevant. Where parents actively fostered academic achievement and held it in high esteem; set high standards for behaviour; and provided both practical and emotional support, pupils were more likely to thrive at school. These children also presented certain social and behavioural characteristics which improved their ability to cope with school in a self-fulfilling manner. For example, having a positive attitude & an internal locus of control meant that they had a more positive image within school which was continually reinforced, leading to a positive self-image and a strong sense of self efficacy[4].

Finally, a paper reviewing other studies, assessed which factors are associated with a higher risk of experiencing social and emotional (non-cognitive) and cognitive difficulties, in order to help identify those children that could really benefit from early intervention. Indicators related to socio-economic status and income but also parenting measures were all positively associated with social, emotional and cognitive development[5]

[1] Carneiro P., Crawford C. & Goodman A. (2007), 'The impact of early cognitive and non-cognitive skills on later outcomes', Centre for the Economics of Education, London School of Economics

[3] McLanahan S. (2004), 'Diverging destinies: How children are faring under the second demographic transition', Demography 41(4)

[3] Rutter M. (2006), 'Genes and Behavior: Nature-Nurture interplay explained', Blackwell Publishers

[4] Siraj-Blatchford I. et. al (2010), 'Performing against the odds: developmental trajectories of children in the EPPSE 3-16 study', Department for Education

[5] Blank L., et. al (2011), 'Promoting the social and emotional wellbeing of vulnerable pre-school children', ScHARR, The University of Sheffield

According to SHINE, anecdotal evidence from teachers and headteachers seemed to indicate that the main reason that they sent children to SHINE on Saturday or saw it as beneficial was because of its ability to improve a child's overall confidence or their social skills or their general interest in and aptitude for learning. We ran a survey of all current school project managers[12] to see if this was a common view[13]. We asked whether they felt that the course helped to improve various non-cognitive skills (which for ease of understanding we referred to as "social & emotional skills")[14] and the majority felt that they had seen improvements in all of the skills we quoted – see Figure

[12] School project managers (or SPMs) are the members of school staff who are responsible for overseeing the running of SHINE on Saturday. They can be a teaching or non-teaching member of staff.

[13] See Appendix 2 for more details about the SPM survey and the questionnaire used.

[14] This was based on a set of social and emotional indicators which are referred to in the "Strengths and Difficulties" Questionnaire (SDQ) which is used as a behavioural screening tool for children – see: Goodman R (1997) The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, 38, 581-586.

3 below. Interestingly, the respondents seemed to be less sure about improvements in "resilience" but this could be due to the fact that this is a skill which is less easy to observe.

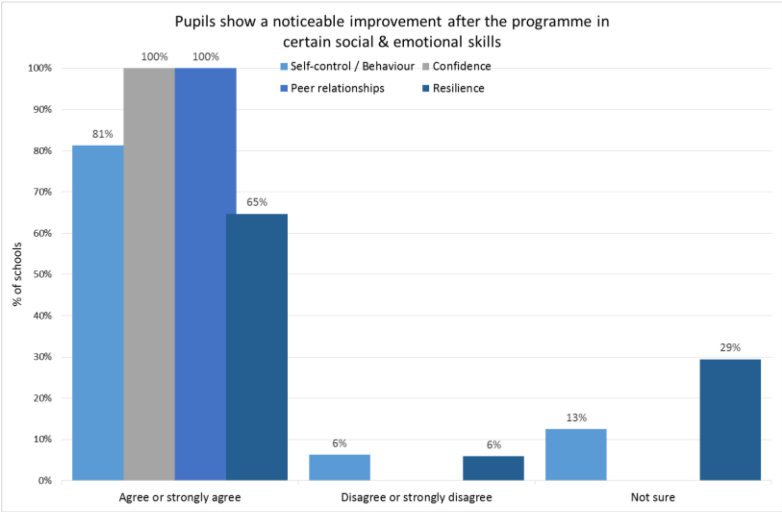

Figure 3: SPMs response (% of respondents) about whether pupils showed a noticeable improvement in certain social & emotional skills after SHINE on Saturday – Survey of SPMs

Schools were also asked whether they actively aimed to develop certain other potential social & emotional skills as part of their SHINE on Saturday programmes. Based on the responses of the SPMs, the most common other skills identified included: communication, teamwork skills & presentation skills as well as enjoyment of and attitudes towards learning – see Figure 4 below.

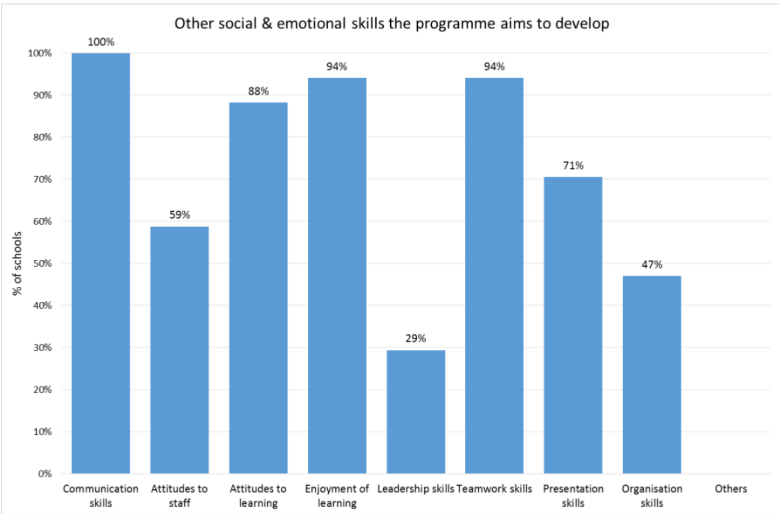

Figure 4: SPMs response (% of respondents) about which other social & emotional skills their SHINE on Saturday curriculum aims to develop - Survey of SPMs

In addition, of those who responded, 82% of SPMs felt that any improvement in social & emotional skills was transferred back into the classroom during weekday school and the remaining 18% weren't sure. This suggests that any improvements in a child's non-cognitive skills, as a result of the SHINE on Saturday programme could be feeding through to their academic performance. Other positive comments made by SPMs about the improvements in pupil's performance at school during the week are set out below.

“Pupils originally deemed to be troubled are now viewed as good achievers and positive contributors in class.”

“Feedback from weekday school and data [provides] evidence of a significant drop in negative behaviour and negative social interactions. Students go on to become leaders and role models within weekday school and support and encourage their peers”

“[Children] are more active as opposed to passive learners. The children have become more confident in expressing themselves and more resilient in approaching new challenges.”

Therefore, based on all of the above, the framework underlying this analysis assumes that SHINE on Saturday is operating predominantly through an improvement in non-cognitive skills. A useful next step for SHINE, would be to produce more robust evidence of this by introducing a method for assessing the non-cognitive skills of every child who receives SHINE on Saturday, both before the programme and at the end to be certain that it is having the positive effect proposed here. SHINE are actually running a pilot of the “Strengths and Difficulties Questionnaire” (or SDQ referred to above) with a sample of schools to see what this suggests about changes in pupils’ non-cognitive skills[15]. The results of this should be reviewed and taken into account, but further research into other surveys and questionnaires which would be suitable for assessing non-cognitive skills would be advisable, to ensure that the best method is used in the longer term, given the particular characteristics of this intervention.

MEASURING EFFECT OF INTERVENTION ON KS2 / KS4 RESULTS

The historic data collated by SHINE for pupils who have received the SHINE on Saturday programme in London can be used to understand the benefits for the treatment group. Schools determine their own selection criteria but we understand that these criteria are likely to be very similar. During our survey of SPMs, we asked about the selection criteria used to target pupils to attend SHINE on Saturday. The results were as follows:

Selection criteria	% of SPMs who use this criteria
Free school meals (FSM) status	100%
English as an additional language (EAL) status	78%
Special educational needs (SEN) status	72%
Academic performance below expectations	100%

Table 1: Selection criteria used to recruit pupils for SHINE on Saturday - SPM survey

Other criteria referred to included additional indicators of deprivation such as poor housing, overcrowding, having parents with mental health issues; children identified as vulnerable at the weekends; children lacking in confidence & social skills; and children who would not otherwise benefit from some of the stimulating experiences and opportunities that they would experience at SHINE on Saturday. To the extent that some schools don’t

[15] See: www.sdqinfo.com for the questionnaires and more details on their use. Some points to consider when making use of this questionnaire would be firstly, the benefit of collecting responses from both teachers and parents, to improve the reliability of the results. Secondly, in theory, given the potential length of time which would pass between the pre and post-intervention questionnaires, there is a risk that there could be some age-related change, which is unrelated to the intervention. Some “norms” data, generated using a large sample of British children, is presented on the SDQ website referred to above. Mean SDQ results are only provided for children aged 5-10 and those aged 11-15, but as the differences in results between these age groups are relatively small, it seems that any age-related change in the results over a one year period is unlikely to be material.

predominantly rely on FSM, SEN & EAL status as the basis of their selection though, it would seem appropriate to consider excluding these.

A "comparator group" – which should represent a group of pupils who did not receive the intervention but are otherwise similar to the treatment group (i.e. exhibit very similar characteristics) needs to be set up. Their academic performance is then considered to represent the "counterfactual", or the outcome if the treatment group had not received the intervention.

SELF-COLLECTED HISTORIC DATA – TREATMENT GROUP

SHINE has collected data on the SHINE on Saturday projects consistently since the programme was first trialled in Hackney in 2001. In the past they used progress in the results of annual teacher assessment scores (i.e. before and after the SHINE on Saturday programme) to monitor the impact on the pupils taking part but did not consistently collect KS2 results for all pupils (other than those who were in year 6). In addition, they have obtained the following anonymised historic data on each pupil:

- School
- Year of intervention
- Level of attendance (i.e. number of sessions out of 30)
- Gender
- Month & year of birth
- Year group
- Indicators of "disadvantage" – i.e.
    - FSM (free school meals)
    - SEN (special educational needs)[16]
    - EAL (English as an additional language)

METHODS FOR SELECTING THE COMPARATOR GROUP

Ideally pupils from the same schools would be used to ensure greater consistency. With help from the schools themselves, the aim would be to identify those pupils who had similar characteristics to those who took part in SHINE on Saturdays and could also have benefitted from the intervention – i.e. those who would have also met the same eligibility criteria at the same point in time. This will only work however, if the number of children who received the intervention was less than the number who were eligible for the programme.

If it is not feasible to construct a robust comparator group using school specific data – i.e. not all of the schools used in constructing the treatment group can provide data on comparator pupils, then the following alterative approach could be considered. Data from the National Pupils Database (NPD) maintained by the Department for Education could be used to generate an appropriate "benchmark". Data extracts from the NPD, which holds a wide range of information and data about pupils at English state schools, are available to bodies and organisations which are conducting research and analysis. This would involve selecting pupils from similar types of schools (e.g. London based, similar levels of deprivation) which would have met the schools' eligibility criteria and with other similar characteristics (i.e. controlling for other potential factors which could affect their academic performance). This data could then be used to determine an average level of academic performance for each type of pupil, against which to compare the treatment group.

[16] This is a binary indicator of whether the pupil has any special educational needs provision or not, rather than any details of what that entails.

OBTAINING ACADEMIC PERFORMANCE DATA

Historic KS2 results for the pupils in the treatment group and the comparator group should be available directly from the relevant schools[17]. The subsequent GCSE (KS4) results achieved by these pupils are, in theory, available from the NPD. However, as discussed in Appendix 3, due to data protection rules, without parental consent to link a child's academic performance records in this way, this data cannot be released from the NPD. Historically, SHINE have not obtained this consent and therefore it is not possible to obtain actual KS4 results for pupils who have received SHINE on Saturday prior to the academic year 2015/16.

One method for resolving this would be to "predict" the KS4 results of these pupils, based on their KS2 results, using more aggregated data from the NPD, which does not allow individual pupils to be identified. However, the fact that the children selected to receive the programme are disadvantaged means that applying linkages based on national averages here could be inappropriate. Based on a review of the literature, it is clear that more disadvantaged pupils tend to follow a different academic trajectory, relative to their more advantaged contemporaries.

[17] Ideally, this should be KS2 "sub-levels" (i.e. 4a, 4b, 4c, etc.) rather than "levels" (i.e. 2, 3, 4, etc.) as this allows more fine improvements in performance to be captured. Note that this is only relevant if historic data is used, as from 2015/16 onwards, the KS2 levels system is being removed and therefore this data will no longer be available.

DISADVANTAGE AND ACADEMIC TRAJECTORY

Analysis for the Social Mobility Commission compared the academic trajectories of high achieving disadvantaged pupils with their more advantaged peers from KS1 through to university. The main findings were that not only did most of the more disadvantaged pupils perform more poorly at every stage of academia but those from the more disadvantaged backgrounds who were initially performing well were more likely to fall off a high achieving trajectory. By KS4, these pupils were likely to have been "overtaken" by those of their more advantaged peers who were achieving lower results than them at KS1 – see Figure 5 below.

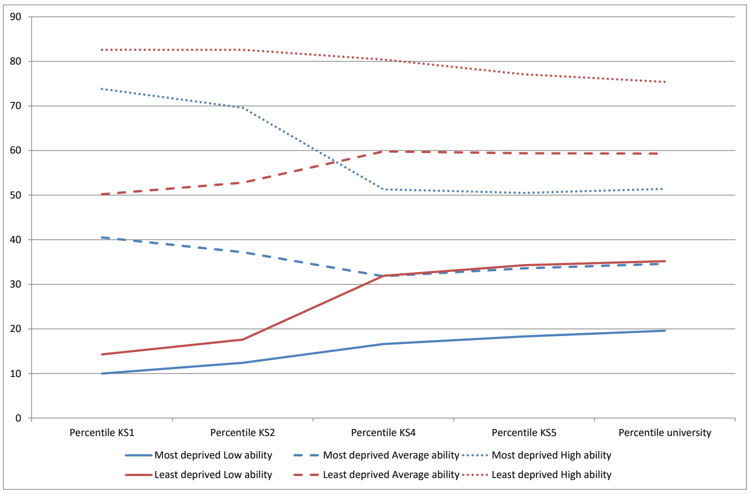

Figure 5: 'Trajectories across Key Stages by early achievement (defined using KS1 maths) for the most deprived and least deprived quintiles of socio-economic status (state school only) – Figure A8'[1]

A very recent paper published by the Sutton Trust highlighted how 15% of highly able pupils who score in the top 10% nationally at age 11 fail to achieve in the top 25% at GCSE, pupils that they refer to as "missing talent". Not surprisingly, approximately a quarter of highly able girls who are eligible for the pupil premium (i.e. those that have ever had FSM status) and approximately a third of highly able boys who are also eligible for the pupil premium fall into this "missing talent" group.

[1] Crawford C., Macmillan L., Vignoles A. (2014), 'Progress made by high attaining children from disadvantaged backgrounds', Centre for analysis of youth transitions, Social mobility & child poverty commission, June 2014

[2] Allen R., Bibby D. & Parameshwaran M. (2015), 'Missing talent', Research brief, Edition 5, The Sutton Trust, 2015

As outlined above, even when two pupils achieve the same level at KS2, the likelihood is that, for the more disadvantaged pupil, performance at KS4 will decline, relatively. Therefore, it is very important that any predictive linkages applied between KS2 and GCSE results are based on the performance of children with similar levels of disadvantage. An appropriate approach would be to use data from the NPD to generate conditional probabilities indicating the likelihood of a pupil with a certain KS2 result, reaching a certain KS4 result, given certain characteristics (especially relating to indicators of disadvantage). For the purposes of this report, the data required to calculate such probabilities has been obtained from FFT[18]. They have provided SHINE with high level, aggregated data from the NPD relating to London primary schools only, which has enabled us to calculate conditional KS4 probabilities for the years 2010/11 – 2014/15. This data could be used in any future studies that SHINE were to undertake using historic data and has been used below in estimating the impact of SHINE on Saturday on the pupils used to determine the minimum sample size. See Appendix 5 for more details including the resulting "expected" probability weighted KS4 predictions. Therefore, it should be possible to at least determine the predicted KS4

[18] FFT is a non-profit organisation which has been processing the NPD for the DfE since 2004 and providing high level data from this database to schools to help them compare the performance of their pupils.

results (i.e. the “expected value” based on probabilities of different outcomes) for each pupil in the intervention and comparator groups.

LINKING KS4 RESULTS TO PREDICTED LIFETIME EARNINGS

Ideally, this economic impact analysis would involve a long-range study following pupils who received the intervention & the comparators into adulthood, controlling for all relevant factors, however this is unlikely to be feasible. The next best alternative is to use findings in the literature to estimate the subsequent effect on productivity of improved KS4 results. My literature review covered a number of academic papers which looked into the economic impact of education. However, in many cases the method of measurement meant that they could not be directly applied to this analysis. The most relevant paper is a recent study published by the Department for Education (DfE).

DFE STUDY ON LIFETIME PRODUCTIVITY GAINS FROM ADDITIONAL GCSES

In a recent study produced by the DfE, the economic value of intermediate qualifications is considered[1]. This includes estimating the lifetime productivity gains from GCSEs, A-levels & apprenticeships. In particular, the paper provides estimates of the present value of the marginal lifetime productivity gains of additional “good” GCSEs at grade A*-C depending on the number of existing GCSEs at that level – e.g. 1-2 vs. no good GCSEs, 3-4 vs. 1-2, etc. It also presents separate results for men & women. Using these results, it should be possible to determine the additional productivity gains generated by a certain number of additional “good” GCSEs. However, given that the incremental effect of the intervention is likely to be less than the step up between numbers of GCSEs shown in this paper, for our purposes some adjustment is required. Due to the difficulty of dealing with ranges, the first step is to convert them into mid-points (i.e. 1-2 = 1.5 and 3-4 = 3.5), therefore the differential can be more easily defined (in this case two additional good GCSEs). As an approximation, the marginal return to the increase from 1-2 to 3-4 good GCSEs can then be divided by two to determine the marginal return per GCSE within that range. This assumes that there is a linear relationship between the marginal return and the number of GCSEs, which may not actually be the case. It is possible that we are therefore overstating the true return. One way to take this into account, would be to use the “low estimate” of the returns quoted in the paper rather than the central estimate. All of these figures are presented in Appendix 6.

[1] Hayward H., Hunt E., Lord A. (2014), ‘The economic value of key intermediate qualifications: estimating the returns and lifetime productivity gains to GCSEs, A levels and apprenticeships’, Department for Education

Because these are marginal returns, they presume that GCSE is the highest level of academic qualification achieved. In actual fact, some of these children may go on to further study (A levels or a university degree), potentially as a result of their better than otherwise level of performance at GCSE. In such cases, the economic benefit of the intervention could be understated. However, the likelihood is lower than average, given that we are dealing with children who are already disadvantaged (as discussed in the academic papers referred to in the previous section).

In addition, as already mentioned (and discussed further below), the literature on non-cognitive skills suggests that disadvantaged children who received interventions aimed at improving their non-cognitive skills went on to earn over and above the income that their level of academic attainment would predict. Clearly, the approach proposed would also exclude this further benefit. However, it is better to be conservative in such analyses and understate benefits / overstate costs rather than the reverse.

THE INDIRECT BENEFITS OF NON-COGNITIVE SKILLS ON LABOUR MARKET OUTCOMES

From a review of other academics' research, Heckman concludes that non-cognitive ("character") skills predict a wide range of outcomes, including labour market outcomes. Such "softer skills" are as powerful in their predictive ability as measures of more standard cognitive skills. Conscientiousness is considered the most widely successful of all of the personality measures identified[1]. In another paper, from his own analysis, he finds that employment status, work experience, occupational choice and wages are strongly affected by non-cognitive skills, or specifically self-esteem and locus of control, over and above their impact on academic achievement[2].

Other studies have focused on the reverse – the negative impact of inadequate non-cognitive skills on work and earnings. One UK base study focused on just one non-cognitive skill "social skills" (or rather degree of social maladjustment). It found that social maladjustment at the age of 11 is associated with lower employment probabilities and lower wages at age 33 & 42[3].

Another study considered the effect of various psychological and behavioural attributes evident by age 10 on later life. The existence of conduct disorder at age 10 was found to predict male adult unemployment very well. However, existence of self-esteem was a good indicator of a man's earnings. For women, the existence of a locus of control was particularly significant in understanding the probability of long-term unemployment and earnings[4].

[1] Heckman J. & Kautz T. (2012), 'Hard evidence on soft skills', Labour Economics 19(4), 451-464

[2] Heckman J., Pinto R., & Savelyev P. (2013), 'Understanding the mechanisms through which an influential early childhood program boosted adult outcomes, American Economic Review, 103(6), 2052-2086

[3] Carneiro P., Crawford C., Goodman, A. (2006), 'Which skills matter', Centre for the economics of education, LSE

[4] Feinstein L. (2000), 'The relative importance of academic, psychological and behavioural attributes developed in childhood', Centre for economic performance, LSE

OTHER POTENTIAL BENEFITS

Although, in theory, there are other potential knock-on benefits from increased academic attainment, they are either less significant than increased productivity and / or quite difficult to estimate. Therefore, also considering the need to take a conservative approach to ensure that the gross benefits are not overstated, we would recommend not attempting to include these additional benefits in the overall economic impact. However, we outline these additional benefits at a high level below.

TAX REVENUES & BENEFITS

Clearly, if these pupils end up earning more than they would otherwise be expected to, then they will generate greater tax revenues for the government and are likely to claim lower benefits. The additional tax would already be captured in the value of additional productivity as any extra tax will be paid out of any extra income earned. However, in theory, benefit payments foregone should be incorporated in the sum of economic benefits. If this were to be estimated, some form of linkage between GCSE results and levels of benefit claims would have to be identified.

OTHER SOCIAL BENEFITS

There are a range of other benefits of improved academic attainment which may have more indirect economic benefits. While many of these could be considered as very important in their own right, attempting to robustly measure these and the potential feedbacks generated would be challenging. Examples include reduced crime, improvements in health outcomes, improved psychological wellbeing, reduced income inequality and improved social mobility. In addition, there can be positive spillover effects from those whose academic attainment improves to those who are less well educated, in the form of knowledge transfer.

SAMPLE SIZE ESTIMATION

It is important to determine whether SHINE currently has sufficient historic data to reliably estimate the main productivity benefits of SHINE on Saturday. This will help the charity decide whether the economic impact analysis could be performed now or if it would have to be delayed until sufficient evidence has been collected. In order to ensure that the results of the analysis underlying an economic impact analysis will be statistically robust, the sample size used must be adequate.

In order to calculate the minimum sample size a "power calculation" has to be performed. This is based, amongst other things, on an "estimate" of the impact of the intervention on academic results relative to the comparator group as well as the level of statistical significance or "confidence" one wishes to achieve. At first sight this may seem rather circular – i.e. a sample must be gathered and the impact of the intervention tested, in order to determine the size of the sample required. However, a reasonable rule of thumb is that providing the sample size for this analysis is at least 30 or more, then the outcome of the power calculation should be reasonably reliable for this purpose[19].

In order to determine the appropriate sample size, we have performed a "pilot" KS2 impact analysis. Data from one London primary school, which has been involved with SHINE over a long period has been used, for the purpose of "estimating" the impact of the intervention[20]. The expectation was that enough pupils at this school would have received the intervention to provide sufficient data for this analysis. In addition, by focusing on one school, there was no need to worry about the effect of variation in school-related characteristics on the pupils' performance. In addition, we know that many pupils at this school have repeated the SHINE on Saturday programme twice or three times and we hoped to be able to glean some initial indication of the relevance of this.

As discussed above, because SHINE has not historically obtained parental consent for subsequently linking pupils' KS2 results with their later academic performance in the NPD, it was not possible to obtain actual KS4 results for the treatment and comparator groups for this specific school. Instead, we performed our analysis using KS2 data. Therefore, the output of the analysis identified the number of pupils required to reliably determine the impact of SHINE on Saturday on pupils' KS2 results.

Full details of the regression analysis to determine the estimated impact of SHINE on Saturday and the power calculation are set out in Appendix 4. Here we summarise the results:

- Approximately **2,000 – 2,200** observations are required to generate a reliable assessment with a 95% confidence level of the impact of the intervention on KS2[21].
- The sample should cover multiple schools, in order to allow for potential school specific effects – we would propose an even number of pupils selected from 10-12 schools.
- SHINE has access to data from approximately 150 schools. However, data is most easily accessible for the more recent period 2011/12 – 2015/16 – the distribution of this data is set out below (note that this does not take into account duplicated school or pupil records - i.e. schools running / pupils attending SHINE on Saturday in more than one year):

[19] More details of the power calculation formula is provided in the section "Sample size calculation" in Appendix 4.
[20] This school wished to remain anonymous.
[21] The minimum number is closer to 160, however, as explained in more detail in Appendix 4, sensitivity analysis suggests that this result is quite sensitive to the impact on KS2, which based on the confidence interval, is quite variable. Therefore, it was considered appropriate to increase the sample size to reflect the risk of a potentially lower effect size.

Academic year	SHINE on Saturday	
	No. of schools	No. of pupils
2011-12	74	900
2012-13	90	1,260
2013-14	98	1,740
2014-15	95	1,860
2015-16	92	1,740

Table 2: Numbers of schools & pupils taking part in SHINE on Saturday, by year

Therefore, it seems clear that the analysis could be performed using existing data, providing predicted rather than actual KS4 results were relied upon.

FURTHER IMPLICATIONS OF THIS ANALYSIS

As the sample collected from this one primary school was found to be large enough to generate statistically significant results, we can draw some conclusions regarding the impact of the intervention for this school[22]. However, it is important to note that these results are specific to this school and in particular the number of times the programme was repeated for these pupils. Therefore it cannot be extended to the intervention more generally. As discussed above, a broader study covering more schools is required. However, these are nonetheless very interesting and quite promising results at this stage.

We can conclude that at this specific school, SHINE on Saturday has historically improved the average KS2 results of those disadvantaged children who received the programme at least once by 0.22 of one "level" on average, relative to those that would have been eligible but did not receive it[23]. Given the conditional probability data provided by FFT, we are therefore able to predict that this increase in KS2 scores has led to an average improvement in the GCSE results of these pupils of at least 1.4 additional "good" GCSEs (i.e. at grade A*-C). This represents the difference between the probability weighted expectation of the GCSE performance of the treatment group (5.4 GCSEs at A*-C) & the comparator group (4 GCSEs at A*-C), given that they have attended a primary school in London, and conditional on their gender, the combination of disadvantage indicators they exhibited (i.e. FSM, EAL & SEN) and their KS2 results. Finally, using the DfE paper referred to in the section above ("Linking KS4 results to predicted lifetime earnings") in conjunction with the average expected improvement in "good" GCSEs of this specific sample of 148 children who attended the SHINE on Saturday programme between the academic years 2005/06 and 2013/14, we have estimated the effect on economic productivity. In conclusion, the intervention run at this one school could have resulted in an overall improvement in economic productivity (in the form of additional lifetime earnings) in the region of £6.3million[24].

We now turn to how this framework can be applied in practice.

[22] In other words, we can have confidence that the true relationship between the variables of interest (i.e. the effect of receiving the SHINE on Saturday intervention on a pupil's KS2 results) is positive this is termed "statistical significance". The confidence interval sets out how close the true relationship is likely to be to the estimated relationship.
[23] More specifically, we are 95% confident that the real result falls somewhere within the range of 0.04 – 0.39 (of which 0.22 is the mean). This is a reasonably broad "confidence interval" but 0.22 is our "best estimate".
[24] As discussed in the section "Linking KS4 results to predicted lifetime earnings" above, we have applied the DfE's "low estimates" to allow for the fact that we are assuming a linear relationship between the return to an increase in the full dataset number of good GCSEs achieved.

CONSTRUCTING AN ECONOMETRIC MODEL

Once appropriate data has been obtained and a sufficiently large sample has been selected to reflect the treatment group and the comparator group, the next step is to construct an econometric model to determine the impact of the intervention.

An econometric model is required, similar to that used to estimate the impact of SHINE on Saturday for the purpose of determining the minimum sample size. This will be used to identify the average incremental effect of receiving SHINE on Saturday, relative to the comparator group, while controlling for any other measurable factors which could be influencing the academic performance of the pupils within the sample. As discussed already, ideally, the model should be run using KS4 data for the treatment group and comparator group. This will help to determine what the actual longer term impact of the intervention is. However, as explained, without parental consent, it is not possible to obtain KS4 results from the NPD. Even if SHINE took the decision to start a full economic impact analysis from this academic year onwards and parental consent was collected, it would take some time to build up a large enough sample of pupils, given that it takes 6-8 years for a pupil who has received SHINE on Saturday (in years 4, 5 or 6) to reach year 11, when pupils sit their GCSEs. In addition, the sample size calculation has been based on analysis using KS2 data. To the extent that there is some sort of fade out between then and KS4, the sample would have to be somewhat larger, if KS4 data was to be relied upon.

Assuming that the econometric model is run using KS2 data, the process should be as follows:

1. Test different potential models to explain pupils' KS2 results, including some of the following potential explanatory variables, data for which should have already been collected, or available directly from the schools:
   - Gender
   - KS1 results or other measure of pre-KS2 performance[25]
   - School indicator
   - Month of birth & year of birth
   - Level of attendance
   - Cohort (e.g. year of KS2)
   - Indicators of disadvantage:
     - FSM status
     - SEN status
     - EAL status
   - Indicator for concern that academic performance would fall below expectations[26]
2. Identify the "treatment effect" which is what is left after controlling for all other factors, using the econometric model with the "best fit" (or rather that most closely meets the accepted criteria for model specification) – this reflects the average improvement in KS2 results for a pupil in the treatment group relative to a pupil in the comparator group. This, in itself, is a proxy for the incremental improvement in KS2 results just from receiving SHINE on Saturday.
3. Use this best-fit model to "predict" the average KS2 results for pupils in the treatment group and for pupils in the comparator group which also display different combinations of the pupil characteristics (i.e. a subset). This is in order to match up with the sub-sets defined in the set of KS4 conditional probabilities – for example:
   - Female + FSM only
   - Male + FSM + EAL
   - etc.

---

[25] For example, Teacher assessment results at the end of year 3.

[26] The pilot school didn't record anything on this, but it would be worth confirming with other schools whether they do.

CALCULATING THE TOTAL IMPACT ON EARNINGS / PRODUCTIVITY

Finally, given everything discussed above, it should be possible to trace through from the impact of the intervention on the sampled pupils' KS2 results, to the impact on their expected KS4 results and hence to the overall productivity benefits, scaling up to the total number of pupils who received the intervention during that period.

Using the predicted average KS2 results for each sub-set of pupils, based on the relevant outcome of the best-fit econometric model, the relevant average "expected", probability-weighted KS4 results can be identified in each case[27]. This would need to be repeated for each sub-set in both the treatment group and comparator group. Consequently, the difference between the expected KS4 results in the treatment & comparator groups for each sub-set of pupils can be calculated – see below[28].

$$Average\ EV\ (KS4_{T,X}) -\ Average\ EV\ (KS4_{C,X}) = Impact\ at\ KS4\ for\ subset\ 'X'$$

Using these differentials, in conjunction with the relevant level of additional lifetime earnings from additional GCSEs at grade A*-C (from the academic paper referred to above – see Appendix 6), the increase in earnings, or equivalently productivity, for each sub-set of pupils can be calculated – see below.

$$Additional\ earnings/GCSE \times \left(Average\ EV\ (KS4_{T,X}) -\ Average\ EV\ (KS4_{C,X})\right)$$
$$= Impact\ on\ earnings\ for\ subset\ 'X'$$

Finally, grossing up these results for the number of pupils which fall into each of the sub-sets in the whole "population" (i.e. all pupils who received SHINE on Saturday during the relevant time period), will produce the overall economic benefit of this intervention, over the period of interest.

$$\sum_X Impact\ on\ earnings\ for\ subset\ \ 'X' \times No.of\ pupils\ in\ population\ in\ subset\ 'X'$$
$$= Total\ impact\ on\ productivity\ for\ population$$

COSTS

In this section, a proposed approach is set out for measuring the total economic & social costs of providing the SHINE on Saturday programme. This includes both the explicit and implicit costs of offering this intervention incurred by all parties involved.

OVERVIEW

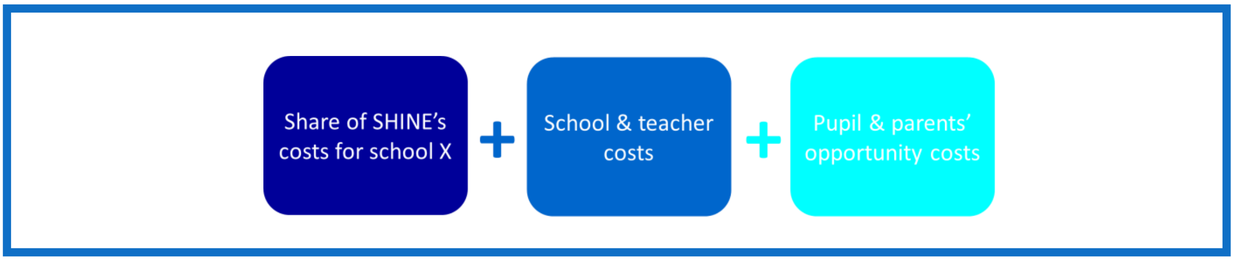

Figure 6: Total economic costs of SHINE on Saturday for school X

---

[27] See Appendix 5 for these expected probability-weighted KS4 results, calculated using conditional KS4 probabilities which are based on data provided by FFT.

[28] Average EV(KS4) is the average expected, probability-weighted KS4 result; T = treatment group; C = comparator group; and 'X' is a sub-set of pupils with the same combination of gender and disadvantage characteristics.

As the diagram above shows, the total costs of SHINE on Saturday include a share of SHINE's running costs – i.e. based on a cost allocation process, a proportion of the costs of running SHINE's head office are associated with the SHINE on Saturday programme. In addition, there are the costs of opening the school on a Saturday and paying the teachers as well as determining the opportunity costs of the time sacrificed by both the pupils attending SHINE on Saturday and their parents for the purpose of transporting them to the school.

In calculating the overall net benefits, a suitable sample of pupils will be identified and then the results will be grossed up to reflect the total numbers of pupils which have received the intervention over each year. Therefore, the costs should be calculated on the same basis – i.e. they will reflect the full scale of provision of SHINE on Saturday.

SHINE'S INTERNAL COSTS

The aim here, is to calculate the proportion of the charity's internal running costs which can be allocated to SHINE on Saturday, as the charity is involved in a variety of different activities. The most appropriate way of doing this is to apply the "fully-allocated costing" (or FAC) approach. Guidance produced by ACEVO (in conjunction with KPMG & NPC) for charities on the importance of allocating all of their costs when costing their activities provides suggestions on how to apply this technique.

Under the FAC approach, costs are separated into "direct costs" which can be directly allocated to services & "indirect costs" which are generated by some intermediary activity and which therefore require allocation across the services provided. For each indirect cost a cost driver or measure which accurately reflects the scale of the identified activity driving that cost can then be used to allocate a proportion of that indirect cost to the various services. SHINE's services currently include the following:

- SHINE on Saturday
- Serious Fun on Saturday
- SHINE in Secondaries
- SHINELabs
- Student leaders
- Let teachers SHINE
- London Teacher Innovation Fund
- Grant making to other charities

In line with ACEVO's guidance[29], indirect costs (more commonly termed "overheads") should be first assigned to an appropriate cost centre and then allocated across other cost centres and the organisation's services using the recommended "cost drivers". This process continues until all costs have been allocated to final services. For example, as shown in the diagram below, a proportion of "Premises & office costs" should be allocated on the basis of headcount (see Table 3 below). A proportion of these costs should therefore be allocated to the cost centre "Finance", according to the number of finance staff relative to all staff at head office (x_2% in the diagram below). Similarly, if there are staff at head office who spend all or some proportion of their time working on SHINE on Saturday, then a proportion of these costs should also be directly allocated to that final service, again based on head count relative to total headcount at head office (x_9%). Note that the size of the cost centres at the top of the diagram and the charities various services at the bottom of the diagram are not shown to scale. In addition, the proportions (x_1 %, etc.) are currently unknown, but given the small size of the charity, it seems likely that headcount would have to be measured in terms of a proportion of an FTE, given that many of the staff roles include multiple responsibilities.

[29] NPC, ACEVO & KPMG (2004), 'What is full cost recovery' – see: http://www.thinknpc.org/publications/full-cost-recovery-2/

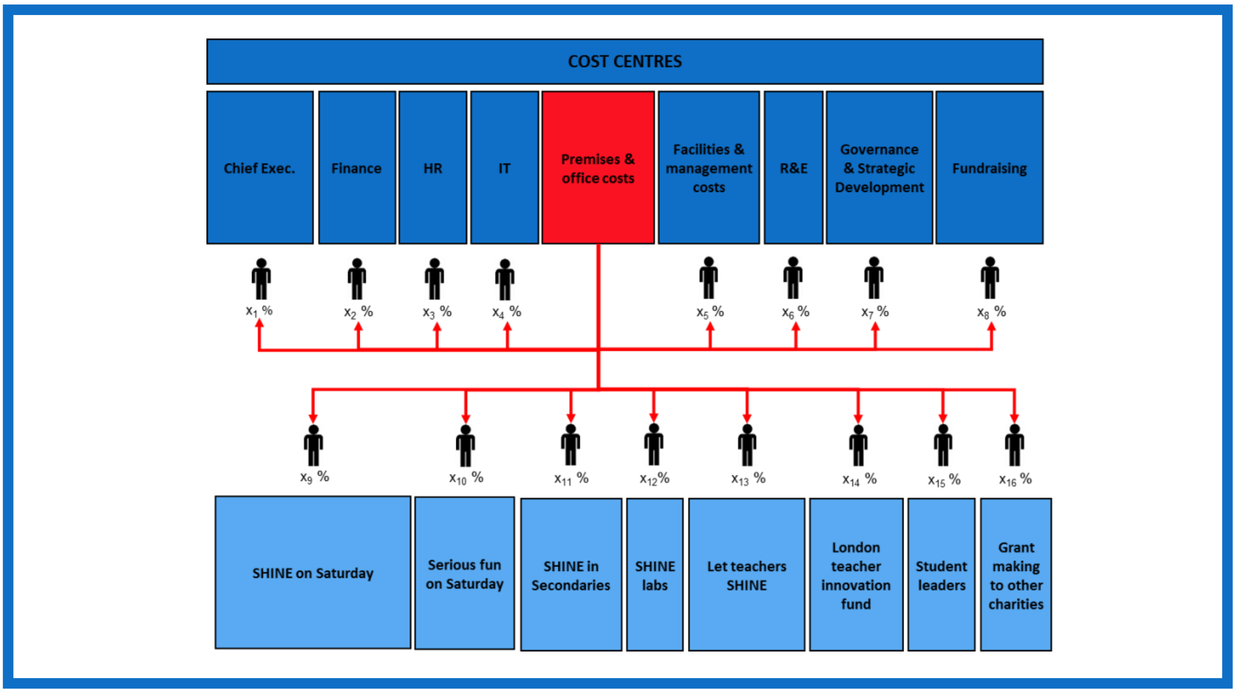

Figure 7: Allocation of indirect costs based on FAC approach

Within the cost allocation template provided by ACEVO, which charities can use for calculating their fully allocated costs, the following cost drivers are used for allocating the costs associated with each cost centre defined within the template. This could serve as a good basis for allocating SHINE's own costs.

COST CENTRE	COST DRIVER	DESCRIPTION
Facilities & office management	Headcount	% of HQ staff working on each activity
Premises & office costs	Headcount	% of HQ staff working on each activity
Research & evaluation	Outputs	% of outputs used by other departments
Chief Executive	Time	% of time spent on different activities by Chief Exec & direct
Finance	Time	% of time spent on different activities by finance team
HR	Time	% of time spent on different activities by HR team
IT	Time	% of time spent on different activities by IT team
Fundraising	Expenditure	Allocated pro-rata based on expenditure already allocated to different activities
Governance & strategic development	Expenditure	Allocated pro-rata based on expenditure already allocated to different activities

Table 3: Recommended cost drivers by cost centre - ACEVO, NPC & KPMG FAC guidance

DATA

SHINE has an internal accounting system from which simple management accounts are produced. Using historic management accounts, it should be possible to classify the direct and indirect running costs of SHINE (excluding the grants paid to schools – this is dealt with below) for specific years. Suitable drivers should then be determined, based on data available and the guidance shown in Table 3 above.

From reviewing a recent set of management accounts, the most material costs could be identified – these are listed below, along with a brief discussion of how they might be dealt with under the guidance. Note that in most cases, a finer breakdown of these costs would be required first, along with input from the head office team to create the cost driver data.

- **Staff costs** – each salary should be assigned to the relevant cost centre & then allocated to services from there.
- **Marketing costs** – the relevant cost centre depends on the purpose of the marketing materials, for example, some may be for fundraising purposes (so attributable to the "Fundraising" cost centre) or some may be for raising awareness of a particular programme with potential beneficiaries (so a direct cost attributable to that programme).
- **Technology** – this is likely to form part of the "IT" cost centre and therefore will be allocated across the other cost centres according to the proportion of time spent in supporting those activities by the IT team. To the extent that they also directly support any of the services then a proportion of these costs would also be allocated to each of these programmes.
- **Audit fees –** as an annual financial audit is a statutory requirement for all charities, this cost should be attributed to the "Governance & strategic development" cost centre and then consequently would be allocated on a pro-rata basis after all other costs have been allocated.
- **Rent, rates & electricity** – this cost forms part of "Premises & office costs" and therefore will be allocated across other cost centres / services according to the headcount of staff at the head office who work in these various areas.

Once the total costs associated with SHINE on Saturday have been identified, the next step would be simply to exclude those costs associated with the schools in Manchester running the Hallé SHINE programme. This could be done based on an estimate of the time spent by staff at head office on that specific programme. The remaining costs would reflect the total running costs related to the London based SHINE on Saturday programme in a given year, which would be an appropriate basis of comparison for the benefits generated by SHINE on Saturday for all of the pupils who received it in that same year.

SCHOOLS / TEACHERS COSTS

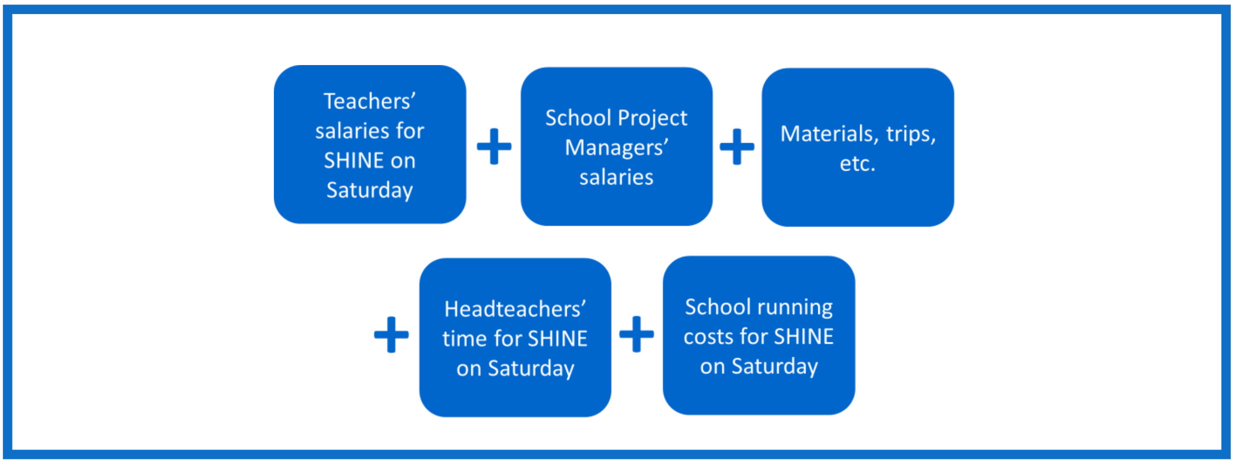

Figure 8: Total costs generated by schools & teachers

The primary schools which run SHINE on Saturday face various explicit and more implicit costs from running the programme which must be captured in the total economic costs. SHINE provides a grant to the schools to cover at least a proportion of their explicit running costs (including paying teachers who work on Saturdays to teach the pupils, the salary of the school project manager & any materials, trips, etc. which form part of the Saturday school curriculum). As the programme continues, schools tend to take on a larger share of the costs themselves, or other funders may also contribute.

Headteachers are also involved in the running of the SHINE on Saturday programme but the cost of this is implicit as a headteacher won't earn any more for offering SHINE on Saturday. Therefore a share of a headteacher's annual salary should be taken into account, to represent the time spent on the set-up phase & also time spent on ongoing management of the school project manager.

Finally, schools have to open specially on Saturday to run the SHINE on Saturday programme. Therefore there is a further implicit cost in terms of additional premises costs (e.g. heating, utilities, cleaning, etc.). Again, an estimate should be made of the share of total school running costs to reflect this.

DATA

SHINE's own records of actual grant payments, includes budgeted cost breakdowns produced by each of the schools. Regardless of the size of the grant paid by SHINE, these budgets can be used as an estimate of the explicit running costs – this includes teachers' Saturday pay, the project manager's salary & the cost of any materials or trips required by the curriculum. Historically, SHINE would normally fully fund the schools' costs for the first few years and subsequently, their contribution would decline. SHINE have confirmed that it is common for the schools to try to make cost savings, as the share of the total costs of SHINE on Saturday that they have to cover increases. This should be evident in their budgeted figures. Therefore, this element of the cost base can be determined as follows[30]:

$$\sum_{S,Y} Budgeted\ running\ costs\ of\ SHINE\ on\ Saturday_{S,Y}$$

DfE provide data which can be used to estimate the remaining elements of the schools costs. Firstly, they provide data on the range of headteachers' salaries for different geographic locations[31]. Using the average salary for an inner London school and an estimate of the proportion of working time per year spent on planning for and implementing SHINE on Saturday by headteachers, an estimate of the costs can be determined. The estimate of the proportion of a headteacher's working time can be assessed by surveying a sample of headteachers at the primary schools used in the analysis[32].

$$\begin{gathered} Average\ annual\ headteacher's\ salary \\ \times\ \%\ of\ working\ days\ normally\ spent\ on\ work\ related\ to\ SHINE\ on\ Saturday \\ \times\ No.\ of\ schools \times No.\ of\ years \end{gathered}$$

Secondly, DfE collect data on schools costs, including the costs associated with running their premises and energy use. These costs are presented for each individual school on a per pupil basis within the "workforce and finance" section of the school performance tables[33]. Therefore, using the numbers of pupils at each school, which is also available from the school performance tables, the annual cost of running these schools can be assessed. Spreading the costs over the 190 days in an academic year we can then determine the daily cost and hence the cost of running SHINE on Saturday over 30 Saturdays (i.e. one full SHINE on Saturday programme).

$$\sum_{S,Y} (Premises\ costs\ per\ pupil_{S,Y} + Energy\ costs\ per\ pupil_{S,Y}) \times\ No.of\ pupils_{S,Y} \times \frac{30}{190}$$

30 Where S = school; Y = year.

31 http://www.education.gov.uk/get-into-teaching/about-teaching/salary/pay-and-benefits.aspx

32 Note that there may be differences for schools who are running SHINE on Saturday for the first time versus schools where it is well established.

33 Results for individual schools can be looked up and compared, using this website: https://www.compare-school-performance.service.gov.uk/

PUPILS & PARENTS COSTS

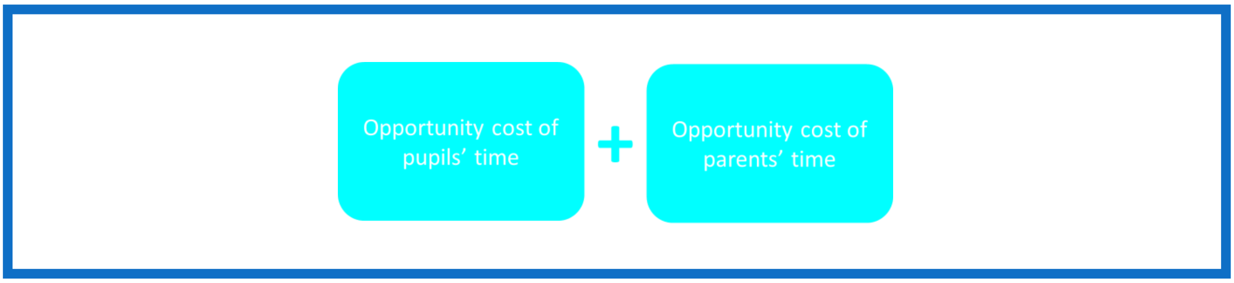

Figure 9: Total costs generated by pupils and parents

There is an implicit cost for the pupils attending the SHINE on Saturday programme and also their parents, assuming that they have to transport their children to and from the school. These can be measured in terms of the opportunity cost – i.e. the benefits from the next best foregone activity.

The children attending SHINE on Saturday are giving up leisure time to attend the SHINE on Saturday programme. However, it is very difficult to measure the opportunity cost of this, particularly as the anecdotal feedback suggests that the children get a lot of enjoyment out of the programme and don't consider it to be comparable to weekday school. In this case, it would be difficult to argue that they were truly foregoing leisure. Generally, the opportunity cost of time is measured based on wages foregone which is obviously irrelevant to children below the legal working age. Based on the literature, it would appear that normally therefore, in such situations, the cost of time foregone for children in education is deemed to be nil[34]. There is a risk that we could be understating the costs slightly, but this seems like a reasonable approach.

The opportunity costs for the parents of those children attending the programme is rather more straightforward to determine as the opportunity cost of time is more commonly measured for adults, particularly in a public policy setting. In addition, the assumption is that it is just the time spent travelling to and from the school which is "lost" by the parents.

DATA

The Department for Transport (DfT) produce figures on the value of non-working time both for commuting and other purposes in terms of cost per hour. These values are publically available in the DfT's "TAG (Transport analysis guidance) databook" which sets out the expected valuations in different years[35]. By surveying a sample of parents or by using details of the catchment area of the school, we can estimate the average duration of the journey from a pupil's home to their school. Therefore we can calculate:

$$\sum_{S} Value\ of\ non-working\ time\ (£/hour)\ \times\ Average\ duration\ of\ journey_{S}\ )\ \times\ 2\ \times\ No.of\ pupils_{S}\ \times\ No.of\ years_{S}$$

Therefore, based on the above proposed analysis, for a specific year or years, it should be possible to sum together the estimates for the share of SHINE's head office's running costs associated with SHINE on Saturday in London; total school & teacher costs attributable to operating SHINE on Saturday; and the opportunity costs associated with the time invested by parents in ensuring that their children attend SHINE on Saturday. This can then be set against the associated benefits discussed above.

[34] Committee on National Statistics, Division of Behavioral and Social Sciences and Education, National Research Council (2004), Chapter 5, 'Beyond the Market: Designing Nonmarket Accounts for the United States', Panel to Study the Design of Nonmarket Accounts

[35] See sheet A1.3.2 in the spreadsheet "TAG_data_book_autumn_2015_final_v1.4b" which is available here: https://www.gov.uk/government/publications/webtag-tag-data-book-december-2015

NET ECONOMIC IMPACT

In this section we briefly discuss how the inputs calculated in terms of economic and social benefits and costs should be combined to determine the net economic impact.

As discussed in the "Approach" section, the incremental benefits and costs attributable to the intervention should be adjusted for the "time value of money". In other words, annual net benefits (or costs) should be discounted at an appropriate rate, such as that proposed by the Treasury, which is currently 3.5%. Given the 7-9 years it takes a primary school pupil in years 4, 5, or 6 to reach KS4 and hence the first point at which they could enter the labour market and start to earn, there is definitely an important element of timing to be taken into account. In other words, the costs will be incurred 7 – 9 years before the benefits begin to be generated. In order to calculate the NPV, the reference point against which to measure this should be selected. If historic data is used, depending on how long ago the pupils received SHINE on Saturday, it may be appropriate to adjust all flows to the current time, in which case historic costs would have to be "compounded" rather than discounted (or increased in value rather than reduced). Future benefits which have yet to be realized would still require discounting to the present time however.

Due to data limitations and the time delay that tends to exist between an intervention and the intended economic impact, it can be difficult to generate reliable point estimates for the economic benefits. In addition, the costs should incorporate the total social costs incurred in providing the intervention and again obtaining robust estimates of values to be placed on non-pecuniary costs can also be challenging. For this reason, it is very important to perform sensitivity testing around any significant assumptions to determine the scale of the effect of an error in these assumptions (separately and in conjunction). Those assumptions which are likely to be least robust should be adjusted within a range of potential values and different scenarios should then be run using different combinations of these varying assumptions. This will generate a range of results which may provide a more appropriate method of presentation. For example, as discussed, in their paper on the link between the number of good GCSEs achieved and the present value of lifetime earnings, DfE present a low, central and high estimate of the marginal returns generated. This is one example of an input which could be varied as part of the sensitivity analysis.

Another useful consideration is the likelihood of a "break-even" outcome – i.e. where benefits are equal to costs. In other words, even if it is difficult to reliably predict the actual benefits, can we take a view on how likely the outcome is where the NPV of the benefits just covers the NPV of the costs? This is the minimum requirement for the intervention to make economic sense. Providing there is a reasonable margin between the lowest estimate of the present value of the benefits and the present value of the costs, then this provides a good basis for suggesting that this is a successful intervention.

NEXT STEPS

In order to run a full economic impact analysis, the next steps required are as follows.

Firstly, the results of the pilot of the SDQ should be reviewed in order to understand if there is more direct evidence that the intervention improves non-cognitive skills. This will make it much more likely that SHINE on Saturday has a lasting impact on pupils' academic performance and later labour market prospects.

In relation to the benefits element of the analysis, the next step will be to decide whether to:

1. use existing historic data combined with the conditional KS4 probabilities set out in Appendix 5 of this report; or
2. collect new data and subsequently obtain these pupils' actual KS4 results.

Under the first option, it would be important to check the completeness of the data and seek to fill any gaps with the relevant schools as well as requesting the pupils' KS2 results. In addition, it would be useful, if possible, to obtain two new indicators – one to represent the schools' views at the time of recruiting pupils for SHINE on Saturday on the expected ability or inability of pupils to meet their academic potential. This is important to ensure consistency in the characteristics of the treatment and comparator groups. Secondly, an indicator of the level of academic potential prior to the intervention would also be helpful, as this would help to explain further the pupils' KS2 results and therefore is an important additional control variable which is currently missing.

Under the second option, the key task would be to collect consent from pupils' parents (in both the treatment and comparator groups) to subsequently obtain their children's' KS4 results and link this data to other data held. This would ensure that actual, rather than just predicted KS4 results are analysed, which should be more accurate. However, the econometric analysis could become somewhat more complex, as the aim would be to explain the academic results of these pupils several years after they attended SHINE on Saturday and isolate that part of their performance that was due to the programme. Therefore, further control variables may be necessary, such as information about the secondary school attended[36]. In addition, as already mentioned, taking such an approach would mean that the results of the study would not be available until at least 2023.

In terms of the cost side, the analysis required is much simpler. The most significant task will be to survey a sample of the headteachers and parents involved to understand more about the time they invest in SHINE on Saturday. In addition, time would have to be spent with SHINE's head office staff, to understand the associated cost data better and construct a set of cost drivers.

There are various ways in which the basic economic impact analysis could be further developed which could be of use to SHINE. In particular, the analysis could be adapted in order to understand how the effectiveness of the London based SHINE on Saturday programme could be improved. For example, by:

- identifying the impact of the intervention on specific groups (e.g. year group / gender / SEN / EAL / FSM); or
- determining what the minimum level of attendance required is to ensure that the intervention is effective; or
- determining whether repeating the intervention on two or three consecutive years affects the magnitude of the impact significantly (i.e. is there evidence of a fade-out effect?)

In order to address any of these questions, however, much larger sample sizes would be required.

[36] Note that we assume that attending SHINE on Saturday would have no effect on the secondary school attended and there is therefore no issue of endogeneity.

APPENDIX 1

BIBLIOGRAPHY

- Allen R., Bibby D. & Parameshwaran M. (2015), 'Missing talent', Research brief, Edition 5, The Sutton Trust
- Blank L., Baxter S., Messina J., Fairbrother H., Goyder E. & Chilcott J. (2012), 'Summary review of the factors relating to risk of children experiencing social and emotional difficulties & cognitive difficulties', School of Health and Related Research (ScHARR), University of Sheffield on behalf of the National Institute for Clinical Excellence
- Carneiro P., Crawford C., Goodman, A. (2006), 'Which skills matter', Centre for the Economics of Education, London School of Economics
- Carneiro P., Crawford C. & Goodman A. (2007), 'The impact of early cognitive and non-cognitive skills on later outcomes', Centre for the Economics of Education, London School of Economics
- Crawford C., Macmillan L. & Vignoles A. (2014), 'Progress made by high attaining children from disadvantaged backgrounds', Centre for Analysis of Youth Transitions, Social mobility commission
- Crawford C. & Cribb J. (2013), 'Reading & maths skills at age 10 and earnings in later life: a brief analysis using the British cohort study', Centre for Analysis of Youth Transitions (partnership between Institute of Education, Institute for Fiscal Studies, National Centre for Social Research)
- Crawford C, Greaves E (2013), 'Comparison of socio-economic indicators: their relationship to economic disadvantage and relevance to Teach First', IFS Report R79, Institute for Fiscal Studies
- De Coulon A., Marcenaro-Guiterrez O. & Vignoles A. (2007), 'The value of basic skills in the British labour market', Centre for the Economics of Education, London School of Economics
- Dolton P. & Vignoles A. (2000), 'The return on post-compulsory school mathematics study', Economica, 69 (273). pp. 113-142
- Hayward H., Hunt E., Lord A. (2014), 'The economic value of key intermediate qualifications: estimating the returns and lifetime productivity gains to GCSEs, A levels and apprenticeships', Department for Education
- Education Endowment Foundation, 'Consent and the data protection act: Advice for evaluators', https://v1.educationendowmentfoundation.org.uk/uploads/pdf/EEF_guidance_for_evaluators_on_consent_and_the_Data_Protection_Act_FINAL.pdf
- Feinstein L. (2000), 'The relative importance of academic, psychological and behavioural attributes developed in childhood', Centre for economic performance, London School of Economics
- Heckman J. (2008), 'Schools, skills & synapses', Economic Inquiry, Western Economic Association International, vol. 46(3), pages 289-324, 07
- Heckman J. (2008), 'The case for investing in disadvantaged young children', Ifo Institute - Leibniz Institute for Economic Research at the University of Munich, vol. 6(2), pages 3-8, 07
- Heckman J. & Kautz T. (2012), 'Hard evidence on soft skills', Labour Economics 19(4), 451-464
- Heckman J., Pinto R., & Savelyev P. (2013), 'Understanding the mechanisms through which an influential early childhood program boosted adult outcomes, American Economic Review, 103(6), 2052-2086
- Heckman J. (2013), 'Fostering and measuring skills: interventions that improve character & cognition', In The Myth of Achievement Tests: The GED and the Role of Character in American Life, edited by J. Heckman, J.E. Humphries and T. Kautz. Chicago, IL: University of Chicago Press. pp. 341-430. (2014)
- Jenkins A., Greenwood C. & Vignoles A. (2007), 'The returns to qualifications in England: updating the evidence base on level 2 & 3 vocational qualifications', CEEDP, 89. Centre for the Economics of Education, London School of Economics
- London Economics (2011), 'The returns to higher education qualifications', BIS research paper number 45
- McLanahan S. (2004), 'Diverging destinies: How children are faring under the second demographic transition', Demography 41(4)
- Morrison Gutman L. & Schoon I. (2013), 'The impact of non-cognitive skills on outcomes for young people - literature review', Education Endowment Fund & Institute of Education
- Morrison Gutman L. & Vorhaus J. (2012), 'The impact of pupil behaviour and wellbeing on educational outcomes', Childhood Wellbeing Research Centre, Institute of Education, University of London

- NPC, ACEVO & KPMG (2004), ‘What is full cost recovery’ – http://www.thinknpc.org/publications/full-cost-recovery-2/
- Rutter M. (2006), ‘Genes and Behavior: Nature-Nurture interplay explained’, Blackwell Publishers
- Siraj-Blatchford I., Mayo A., Melhuish E., Taggart B., Sammons P. & Sylva K. (2011), ‘Performing against the odds: developmental trajectories of children in the EPPSE 3-16 study’, Department for Education / Institute of Education, University of London
- Slavin R., Lake C., Davis S., Madden N.A. (2009), ‘Effective programs for struggling readers: A best-evidence synthesis’, Best Evidence Encyclopedia
- Social Mobility & Child Poverty Commission (2013), ‘State of the nation 2013: Social mobility & child poverty in Great Britain’
- The Sutton Trust (2008), ‘Wasted talent? – Attrition rates of high-achieving pupils between school & university’
- The Sutton Trust / Boston Consulting Group (2010), ‘The mobility manifesto’

APPENDIX 2

SCHOOL PROJECT MANAGERS SURVEY

During the autumn term of 2015, a survey of SPMs (School Project Managers) was undertaken, to understand, from their perspective, the way SHINE on Saturday has an impact on pupils. It was also an opportunity to explore some of the anecdotal evidence about SHINE on Saturday obtained by the charity from schools and teachers.

DETAILS OF SURVEY

- We surveyed a sample of 18 SPMs (out of a potential 23) working in those London primary schools which were currently offering SHINE on Saturday – this excluded any SPMs who were relatively new to the role.
- The questionnaire shown below was emailed by SHINE to the SPMs for completion.
- All SPMs responded, but they did not all answer every question – where this was the case, the no response rate ranged averaged 12% across all questions.
- All results from the survey presented in the main body of this report exclude any non-responses.

QUESTIONNAIRE

SHINE

Project Manager Survey
SHINE 2015

This survey has been designed to help us understand more about certain specific aspects of SHINE programmes from the perspective of the Project Managers who lead delivery each Saturday. This includes:

- the recruitment criteria you use to select children to receive SHINE;
- the impact on children who receive the SHINE programme more than once; and
- the potential impact of SHINE on children's confidence, peer relationships, resilience, self-control and wellbeing

There are a mixture of open and closed questions, with answer boxes provided for you to explain answers or to add your own thoughts as appropriate.

SHINE may use the results of this survey to contact some individuals for follow up interviews. If you are not happy for us to do this please tick here ☐ .

Name	
Job Title (outside of SHINE)	
Name of host school	
Number of years running the project	

We would like to understand how you select pupils to attend SHINE.

1. Please tick any of the following recruitment criteria you use to **target** children for SHINE: (please note, this does not necessarily describe the group who attend)

Free School Meals		English as an Additional Language	
Special Educational Needs		Academic performance below expected levels	

Please outline any other recruitment criteria you apply in the space below.

2. Overall, approximately what proportion of students who are eligible for SHINE go on to attend? ☐ %

3. Please tick any of the following which are common characteristics of the children who **attend** SHINE:

Free School Meals		English as an Additional Language	
Special Educational Needs		Academic performance below expected levels	

Please outline any other common characteristics of children who **attend** in the space below.

4. Are there any common characteristics of children who are eligible but do not attend? Please outline below.

5. What is the maximum number of times that students have previously attended your SHINE programme? (Please circle)

Once (1 year)	Twice (2 years)	Three times (3 years)	Not sure

6. Based on your experience, and feedback from teachers and students, what do you think is the optimal number of times for a student to attend? (Please circle)

Once (1 year)	Twice (2 years)	Three times (3 years)	Not sure	It depends

Please explain your answer

7. Overall, what proportion of students would you say attend SHINE for 1 year? ____ %

8. Overall, what proportion of students would you say attend SHINE for 2 years? ____ %

9. Overall, what proportion of students would you say attend SHINE for 3 years? ____ %

Section 1: Confidence

Please circle the option which best describes your opinion regarding each of the statements below.

1. **Many of the students referred to SHINE demonstrate low levels of confidence at the start of the project.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree

2. **SHINE students on the whole show a noticeable increase in confidence by the end of the project.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree

3. **Building confidence is a key objective of the project I run.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree

4. **What (if any) aspects of your project are specifically designed to improve students' confidence?**

5. **Although we have not actively designed the course to build children's confidence, this is still a beneficial indirect outcome.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree	Not Applicable

Any other comments in this area.

Section 2: Peer relationships

Please circle the option which best describes your opinion regarding each of the statements below.

1. **Many of the students referred to SHINE experience difficulties forming and maintaining positive peer relationships with other children at the start of the project.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree

2. **Building positive peer relationships is a key objective of the project I run.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree

3. **SHINE students on the whole show a significant increase in their ability to build and maintain positive peer relationships by the end of the project.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree

4. **What (if any) aspects of your project are specifically designed to improve students' ability to build and maintain positive peer relationships?**

5. **Although we have not actively designed the course to improve peer relationships, this is still a beneficial indirect outcome.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree	Not Applicable

Any other comments in this area.

Section 3: Resilience

Please circle the option which best describes your opinion regarding each of the statements below.

1. **Many of the students referred to SHINE struggle with resilience at the start, e.g. they are scared to fail at something.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree

2. **Building resilience is a key objective of the project I run.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree

3. **SHINE students on the whole show a significant increase in their resilience by the end of the project.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree

4. **What (if any) aspects of your project are specifically designed to improve students' resilience?**

5. **Although we have not actively designed the course to improve resilience, this is still a beneficial indirect outcome.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree	Not Applicable

Any other comments in this area.

Section 4: Self-control / Behaviour

Please circle the option which best describes your opinion regarding each of the statements below.

1. **Many of the students referred to SHINE lack self-control at the start of the project e.g. they often do or say things without thinking about the consequences, or they struggle to follow instructions.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree

2. **Building self-control is a key objective of the project I run.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree

3. **SHINE students on the whole show a significant increase in their self-control by the end of the project.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree

4. **What (if any) aspects of your project are specifically designed to improve students' self-control?**

6. **Although we have not actively designed the course to improve students' self-control, this is still a beneficial indirect outcome.**

Strongly Agree	Agree	Not sure	Disagree	Strongly Disagree	Not Applicable

Any other comments in this area.

Section 5: Other social and emotional skills

Please tick any other social and emotional skills that your project aims to develop (in addition to those already covered in sections 1-4).

Communication skills		Leadership skills	
Attitudes to staff		Teamwork skills	
Attitudes to learning		Presentation skills	
Enjoyment of learning		Organisation skills	

Are there any others? If so, please outline below.

Based on the feedback of teachers, parents and students, on the whole do you think that any improvements in social and emotional skills are transferred back into weekday school? (Please circle)	Yes	No	Not sure	N/A

Please explain your answer.

Thank you for taking the time to complete this survey. The results will be analysed and circulated later in the year.

APPENDIX 3

APPLYING FOR ACCESS TO THE NATIONAL PUPIL DATABASE

The national pupil database is maintained by the Department for Education and holds a wide range of information and data about pupils at English state schools, including various pupil characteristics and details of educational performance. Data extracts of varying levels of sensitivity (based on a four tier system) are available to bodies and organisations which are conducting research and analysis.

In order to obtain access to the necessary data extracts, SHINE would need to complete the relevant application forms. This includes proving why data at that tier level is necessary for the research being performed and that the data security requirements of the DfE are met. On receipt of the data there are strict requirements regarding its use and access given the sensitive nature of some of the data which would enable the pupils to be identified. In addition, data extracts can normally only be held for a year before they have to be destroyed.

One potentially very useful item of data which could be obtained from the NPD by SHINE in the future, if they decided to pursue a full economic impact analysis, is the actual KS4 results of pupils who previously received SHINE on Saturday (and therefore have been assigned to a treatment group) and the actual KS4 results of any pupils who have been assigned to a constructed comparator group. However, due to the data protection act, there are strict rules about this process. In order to obtain KS4 results, SHINE would probably require access to "Tier 1" data, because access to the pupils' UPNs is required so that they can be cross-referred to the specific pupils who received the SHINE on Saturday programme. This requires the highest level of approval from the Data Management Advisory Panel. In addition, parental consent is required for processing this data and hence linking existing data to NPD data.

It is difficult to find definitive guidance on exactly what this means, however, it seems that an "opt-in" form of parental consent (where parents must explicitly give their consent) is required if the data requested from the NPD is personal sensitive data or the data one wishes to link the NPD data to includes personal sensitive data. However, only an "opt-out" form of consent (where consent is presumed given, unless parents respond otherwise) is required if the data requested from the NPD is just personal data or the data one wishes to link the NPD data to includes personal data. Based on documented advice about consent and the data protection act[37] provided by The Education Endowment Foundation (EEF) to evaluators working on evaluations that they fund, sensitive personal data, in the context of the NPD, would include the following variables:

- The racial or ethnic origin of the data subject – this would include:
    - Ethnicity
    - Ethnic Group
    - Ethnic Group Minor
    - Ethnic Group Major
    - Ethnic Source
    - First Language
    - Language
    - Language Group
    - Language Group Minor
    - Language Group Major
- His / her physical or mental health or condition – this would include:
    - SEN provision
    - SEN provision Major
    - Primary SEN type

[37] Education Endowment Foundation, 'Consent and the data protection act: Advice for evaluators', see - https://v1.educationendowmentfoundation.org.uk/uploads/pdf/EEF_guidance_for_evaluators_on_consent_and_the_Data_Protection_Act_FINAL.pdf

- Secondary SEN type
- Special Provision Indicator
- SEN Unit Indicator
- Resourced Provision Indicator
- Disability

They explain that the DfE cannot give advice on any data which isn't in the NPD, but it is possible that sensitive items which are more aggregated (e.g. such as a binary indicator of SEN status or ethnicity) might be considered non-sensitive due to the more aggregated nature of the variable.

Personal data is considered to be any data which relates to a living individual who could be identified either from that data or from data and other information which is in the possession of or likely to come into the possession of the data controller. However, to the extent that such data is anonymized, it would no longer be considered "personal data".

Following direct discussion with a member of the NPD team at the DfE, below is some guidance which the NPD team was making available to those wishing to link existing pupil data to data from the NPD.

GUIDANCE FROM THE NPD TEAM AT DFE

If we are sharing individual pupil information with a third party researcher who is not a government department or statutory body there are limited conditions that they may rely on for processing personal and sensitive personal data. In addition, irrespective of what they've collected (personal or sensitive personal) and what conditions for processing they have satisfied for this collection, they will need to satisfy the relevant conditions for processing of the Data Protection Act 1998 (the DPA) when linking the data they have collected to NPD (personal or sensitive personal). Where linking is taking place to personal data in NPD at least one of the conditions in Schedule 2 must be met. Additionally, where linking is taking place to sensitive personal data in NPD at least one of the conditions in Schedule 3 must be met. With regards to NPD, sensitive personal data is personal data consisting of information as to the racial or ethnic origin of the data subject or his physical or mental health or condition.

Further guidance around this can be found on the ICO website:
http://ico.org.uk/for_organisations/data_protection/the_guide/conditions_for_processing

In your request, you have confirmed that you are proposing to match the NPD data you have requested with other data that you hold. In order to do so, we need to get confirmation from yourself that any data you hold will be processed fairly **and** lawfully (see reference to DPA below) and you'll also need to specify which condition(s) for processing (see link above for further details) you will be relying on with regards to linking the data you hold to NPD data.

Fair processing is the first principle enshrined in the Data Protection Act (DPA) which says:

Schedule 1, Part 1 (1): Personal data shall be processed fairly and lawfully and, in particular, shall not be processed unless:

a) At least one of the conditions in Schedule 2 is met, and

b) In the case of **sensitive personal data** at least one of the conditions of Schedule 3 is also met.

GUIDANCE FROM THE NPD TEAM AT DFE (CONTD.)

If you believe you have a condition for processing under Schedule 3 of the DPA (e.g. *the data subject has given his explicit consent to the processing of the personal data*) for linking the data you hold to NPD data (i.e. you've had opt-in consent to link to NPD data held by the Department for Education) then it is likely we will be able to share sensitive personal data with you. However, if you believe you only have a condition for processing under Schedule 2 of the DPA (e.g. *the data subject has given his consent to the processing* or *the processing is necessary for the purposes of legitimate interests…*) for linking the data you hold to NPD data (i.e. you've had opt-out consent to link to NPD data held by the Department for Education) then it is likely we will be only able to share personal data with you and we'd need to remove any **sensitive personal data** from your request. Ultimately, it is your responsibility to ensure that all data collected and used as part of your research is processed fairly and lawfully, but by confirming that you have the relevant conditions for processing when linking your data to NPD data, we'll be able to share NPD data with you. Therefore you'll need to confirm that any data you hold will be processed fairly and lawfully (see reference to DPA below) **and** you'll also need to specify which condition(s) for processing (confirming whether these are under Schedule 2 or 3) you will be relying on with regards to linking the data you hold to NPD data.

RESOURCES – NPD ACCESS

If SHINE decide that they do wish to apply for access to data from the NPD, then the following resources may be of use.

- NPD User guide:
  https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/472700/NPD_user_guide.pdf
  This includes details on how to apply and links to the relevant application forms.
- Additional guidance from EEF on completing the NPD application form:
  http://educationendowmentfoundation.org.uk/uploads/pdf/NPD_Data_Request_Application_Form_-_with_notes_for_EEF_evaluators_-_May_2014.pdf
- Additional guidance from EEF on completing the NPD information security questionnaire:
  http://educationendowmentfoundation.org.uk/uploads/pdf/NPD_Information_Security_Questionnaire_-_with_notes_for_EEF_evaluators_-_May_2014.pdf

APPENDIX 4

ECONOMETRIC ANALYSIS - POWER CALCULATION

This appendix sets out the technical details of the econometric analysis which has been performed to provide inputs for the power calculation. In addition, the details of the power calculation itself are also provided. The power calculation is required to determine the minimum sample necessary for performing a full economic impact analysis.

This appendix is structured as follows. Firstly, the sample used is described and the variables incorporated in our analysis are defined. Initially, a comparison of the unconditional means of the KS2 results of the treatment group and comparator group is drawn. The preferred econometric model for measuring the treatment effect and the associated results are then presented, along with the power calculation results, which the predicted treatment effect feeds into. Some alternative model specifications are also set out, to show how the preferred model was selected. Finally, we consider some additional issues regarding potential heterogeneity in the treatment effect, based firstly, on the level of attendance and secondly, on whether or not the programme is repeated.

DESCRIPTION OF SAMPLE COMPOSITION / STRUCTURE

The sample is taken from a single London based primary school which has been offering SHINE on Saturday since 2005/06.

TREATMENT GROUP

Pupils are eligible for SHINE if they display some combination of the disadvantage indicators FSM (Free school meals), SEN (Special educational needs) and/or EAL (English as an additional language). In addition, as confirmed in the SPM survey, this school also considers whether the pupil's academic performance is falling below expectations. However, unfortunately, this school does not retain any data which would indicate whether a pupil was considered to be achieving at, above or below their academic potential, at the time of recruitment onto the SHINE on Saturday programme. The pupils receive SHINE on Saturday during year 4, 5 or 6 and may receive the intervention more than once (although not necessarily during consecutive years). The sample draws on multiple cohorts of data. The earliest cohort is children who took KS2 (i.e. were in year 6) in 2005-06 and the latest cohort is children who took KS2 in 2013-14.

We analyse 148 pupils who have participated in SHINE during this period. We define the treatment group as those who have participated in SHINE and who have attended at least 75% of sessions. Low attendees were excluded from the analysis entirely.

We started with a total sample of 600 pupil records and then excluded the following, to leave the final sample of 148 unique relevant pupils:

- Pupils who had not yet sat KS2 when the data was collected – i.e.
  - Year 4 & 5 pupils from the 2013/14 data
  - Year 4 pupils from the 2012/13 data
- Pupils who only displayed the EAL disadvantage indicator[38]
- Pupils whose attendance fell below 75% (see above)
- Pupils who had left the school before sitting their KS2 exams or for some other reason did not take KS2[39]

[38] According to the headteacher, these pupils tended not to have the same, more entrenched issues of disadvantage and poor academic results and tended to improve markedly academically as their English improved. Therefore, due to the risk of skewing the results, these pupils were excluded from the sample (i.e. the reason for sending these pupils on the programme was quite different).

[39] Note that the school experienced a relatively high level of turnover of pupils during the academic year

- Duplicated pupil records (i.e. pupils who received SHINE on Saturday in more than one year and therefore were included across multiple years)[40]

COMPARATOR GROUP

A comparator group has been constructed by identifying all other children from this school who would have been eligible for SHINE, based on their FSM, SEN & EAL status. Note that it was not possible to retrospectively confirm if their academic performance was falling below expectations as well.

We started with a total sample of 357 pupil records (i.e. pupils who had not received SHINE on Saturday) and then excluded the following, to leave the final sample of 109 pupils:

- Pupils who had either no disadvantage indicators or only EAL
- Pupils who left the school before sitting their KS2 exams or for some other reason did not take KS2

SAMPLE STRUCTURE

The sample structure is set out below. Figure 10 shows the distribution of pupils in both the treatment group and comparator group, based on their "cohort year" – i.e. the year in which they sat KS2 (school year 6).

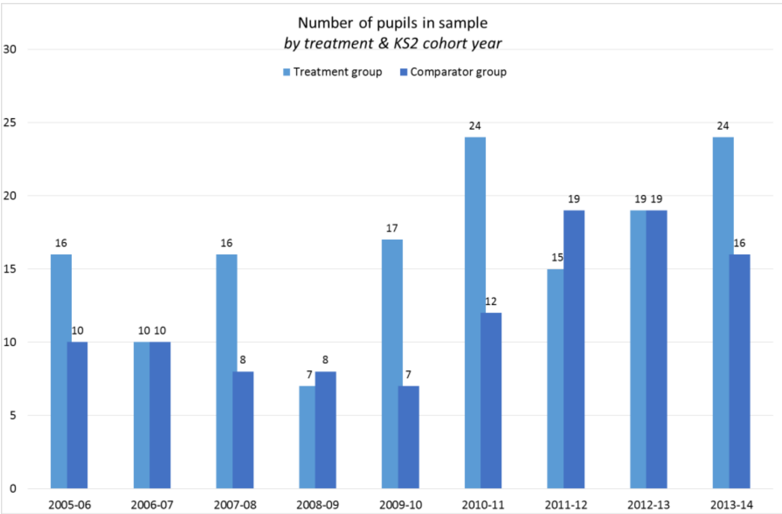

Figure 10: Distribution of sample by cohort

Figure 11 shows how prevalent the repetition of SHINE on Saturday was amongst pupils at this school – 58% attended SHINE on Saturday at least two or three times and the remaining 42% attended only once.

[40] Where pupils attended more than once, their attendance was calculated as an average across every year they attended.

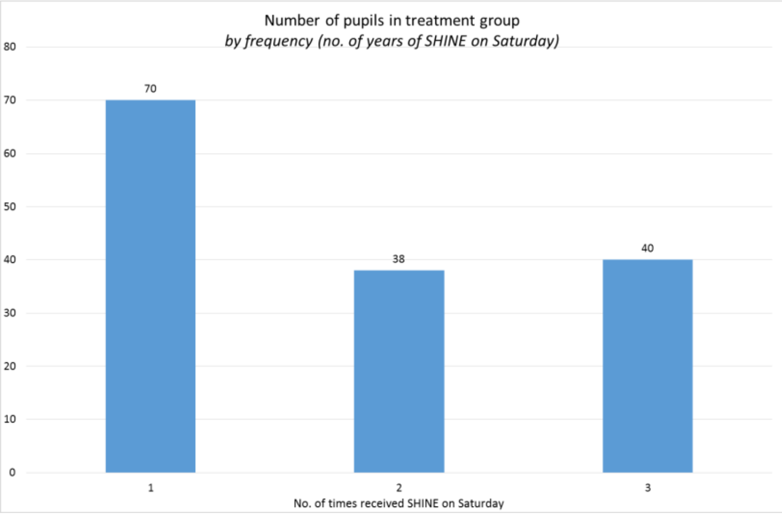

Figure 11: Distribution of treatment group sample by frequency of intervention

VARIABLES

The outcome variable of interest is the pupil's average KS2 score ("KS2avg"). In year 6, pupils take their KS2 exams and are awarded separate KS2 scores for English, Maths and Science. Although KS2 "sub-levels" would have been more useful (for identifying smaller differences in results between the treatment and comparator groups), unfortunately, the school had not retained a complete set of this data. Therefore we used KS2 "levels" data instead. KS2 levels are expressed in integers and range from 2 to 6. We focus on the average KS2 score, which is calculated as the simple average (rounded) of the individual subject KS2 scores.

We also control for various pupil characteristics using a number of dummy variables[41]. The following variables have been used:

- **Treatment effect** - This is a dummy variable equal to 1 if the pupil is "treated" – i.e. receives SHINE on Saturday. The base case is therefore the comparator group.
- **Gender -** This is a dummy variable equal to 1 if the pupil is female and therefore the base case is male.
- **Month of birth** - We use dummies for pupils born in February, March, etc. The base case is pupils born in January.
- **Season of birth** – We use dummies for spring, summer and autumn births. The base case is pupils born in winter.
- **FSM status –** This is a dummy variable equal to 1 if the pupil has free school meals status.
- **SEN status –** This is a dummy variable equal to 1 if the pupil has special educational needs status.
- **EAL status –** This is a dummy variable equal to 1 if the pupil has English as a second language status.
- **Number of years in which pupil attended SHINE on Saturday -** The dummy shine_2times=1 identifies those who have attended SHINE on Saturday twice and the dummy shine_3times=1 identifies those who have taken attended SHINE on Saturday three times. The base case is not having attended SHINE on Saturday at all (i.e. the comparator group).
- **Cohort dummy identifying year in which pupil took KS2** – We have tested both individual year dummies (base case = 2005-06) and a single dummy to distinguish 2005-06 and 2006-07 cohorts from later cohorts (base case is later cohort).

[41] A dummy variable is binary and is therefore either "on" (i.e. has a value of 1) or "off" (i.e. has a value of 0) for each observation in the data set. Therefore, for example, the gender dummy variable measures the effect of being female relative to being male, given all other explanatory characteristics included in the regression.

- **School years in which attended SHINE on Saturday** – The dummy year4=1 applies to those who attended SHINE on Saturday in year 4. The dummy year5=1 applies to those who attended SHINE on Saturday in year 5. Therefore, the base case is a pupil who attended the SHINE on Saturday programme in year 6.
- **Proportion of SHINE sessions attended** – Although this is available as a linear variable, we use the dummy "low_attend" to identify those who participated in shine, but with an attendance rate less than 80%, 90% or 95%, depending on the case (we experiment with different definitions). The base case is either a pupil that did not participate in SHINE on Saturday at all (comparator group), or a pupil in the treatment group who's attendance was above the attendance threshold level.

COMPARISON OF UNCONDITIONAL MEANS

We first compare means in the treatment and control groups. In the comparator group, the average KS2 score is 3.75, whereas in the treatment group the average is 3.752+ 0.281 = 4.03. The 0.28 points uplift in KS2 scores which is associated with treatment (see co-efficient on the treatment effect variable) is statistically significant at the 99% level – see Table 4 below.

VARIABLES	(1) KS2avg
Treatment effect	0.281*** (0.0905)
Constant	3.752*** (0.0687)
Observations	257
R-squared	0.037

Table 4: Treatment effect with no controls[42]

MAIN SPECIFICATION

The unconditional means above do not account for the effect that observable characteristics such as gender, FSM, EAL, SEN, age and cohort have on KS2 outcomes. When these variables are included in the regression (i.e. we control for the fact that there are other factors which could explain variation in KS2 results) we continue to see a positive effect of treatment (attendance on the SHINE on Saturday programme) on KS2 performance[43].

Given the limited sample size of 257, it may not be appropriate to control for every possible pupil characteristic within the same regression. This is because some characteristics will be correlated with each other or will (either alone or in combination) relate to small groups, in which case the estimate of the treatment effect will become imprecise. We have tested a number of specifications including different selections of control variables. These are shown in Table 6.

Our preferred specification is shown in Table 5 below and follows extensive model testing. This model includes the following dummy variables: treatment effect, gender, FSM, EAL, SEN, season of birth and cohort (to distinguish 2005-06 and 2006-07 cohorts from later cohorts). The base case would be a non-treated male without FSM, EAL or SEN, born in winter, and in a cohort after 2006-07. An individual with these characteristics would have a predicted KS2avg of 4.50 (the constant). The KS2avg can be predicted for other groups by adding in the relevant dummy variables. For example, a non-treated female with FSM, in the 2005-06 or 2006-07 cohort, born in spring would have a predicted KS2avg of 4.08 (= 4.50 - 0.12 - 0.29 + 0.22 - 0.23). However, the overall treatment effect across all the pupils (i.e. the differential in average KS2 results between the treatment and comparator groups,

[42] *** p<0.01, ** p<0.05, * p<0.1; standard errors in parentheses

[43] As the comparator group was selected ex-post and has not been perfectly matched with the treatment group, they may differ in terms of observable characteristics. As these may in turn affect KS2 performance, these compositional differences should be controlled for.

which is due to the intervention) can be seen from the coefficient on the treatment effect. This is +0.22, an effect that is significant at the 95% level which should be interpreted as 0.22 of one KS2 level. The confidence interval around this sample mean is reasonably wide at 0.04 – 0.39.

VARIABLES	(2) KS2avg
Treatment effect	0.216** (0.0892)
Gender	-0.118 (0.0880)
FSM	-0.287** (0.125)
EAL	-0.0569 (0.0930)
SEN	-0.441*** (0.101)
KS2 in 05/06 or 06/07	0.217* (0.116)
Born in spring	-0.229* (0.119)
Born in summer	-0.391*** (0.126)
Born in autumn	-0.197* (0.119)
Constant	4.499*** (0.188)
Observations	257
R-squared	0.143

Table 5: Regression results- preferred model specification[44]

Comparing this treatment effect of +0.22 with the +0.28 difference between unconditional means, would suggest that nearly one quarter of the +0.28 is due to differences in observable characteristics between the two groups[45]. The other three-quarters are due either to the effect of treatment, or to unobservable heterogeneity. For example, we do not observe the KS1 performance of the pupils or other measures of their ability prior to the intervention. As assignment into the treatment is not random, there is scope for these characteristics to vary between groups. It is not obvious in which direction any such effects might run – although we know that teachers look to recruit pupils at risk of failing to meet their academic potential, it is not clear what the academic status of the pupils captured in the comparator group are. Some of these may be pupils who were offered the opportunity to attend SHINE on Saturday, due to their academic performance, but chose not to take up the opportunity. Some may have fitted the criteria of failing to meet their academic potential, but there was not sufficient capacity available to offer them a place. While some other pupils may have met the disadvantage criteria but have been performing at their expected level of academic performance and hence were never considered part of the target group. In addition, it

44 *** p<0.01, ** p<0.05, * p<0.1; standard errors in parentheses

45 i.e. 0.28 – 0.22 = 0.6 and 0.28 / 4 = 0.7

is possible that recruitment policies may have varied over time within this school and are likely to vary to some extent between schools.

Most of the coefficients have plausible signs and impacts. FSM and SEN both have negative impacts on KS2avg, whereas EAL has only a small impact. The cohort dummy is significant at the 90% level. The season of birth dummies are broadly in line with the older children performing better (although winter births appear to perform better than autumn in this dataset). The female dummy is negative, which is at odds with other published work on pupil attainment. This could be a quirk in the data for this particular school, which we would not expect to see if the study were generalised to more schools.

SAMPLE SIZE CALCULATION

Firstly, it should be noted that the estimated treatment effect generated by the preferred model specification above is statistically significant at the 95% level of confidence and nearly significant at the 99% level (p-value = 0.016). This suggests that the sample size employed for this analysis, given the magnitude of the treatment effect estimated here, is close to the minimum number of observations needed to detect a significant effect.

However, in order to be more precise, we have also performed a power calculation to backward induce the sample size that would ensure a statistically robust result. The whole reason for analysing the treatment effect in one school was to generate an estimate of the true treatment effect size and hence be able to determine the required sample size if a broader analysis was performed. See below for an explanation of the derivation of the formula used.

FORMULA FOR POWER CALCULATION

Suppose that:

y_i = a measure of attainment for each child i

n = the number of children who receive the intervention (i.e. the sample size)

w_i = the treatment effect (i.e. a dummy which is 1 if the pupil is in the treatment group)

$\hat{\beta}$ = the estimated treatment effect size

x_i = other determinants of y_i (e.g. FSM status, gender, etc.)

u_i = the regression errors

Then the standard treatment effect model (as we have used) can be defined as follows:

$$y_i = \alpha + \beta w_i + \gamma x_i + u_i \quad \text{(A)}$$

Due to the complexity of determining the covariances between y_i & x_i, it is helpful to also run the following regressions:

$$y_i = \alpha_y + \beta_y x_i + ey_i \quad \text{(B)}$$

$$w_i = a_w + b_w x_i + ew_i \quad \text{(C)}$$

Then using the residuals from equations (B) & (C), it is possible to identify both $\hat{\beta}$ and its standard error, by running equation (D):

$$ey_i = \beta ew_i + u_i \quad \text{(D)}$$

This in turn allows the t-statistic (τ) to be determined and hence the minimum sample size, which is a function of this:

$$\tau = \frac{\hat{\beta}}{\sqrt{\sigma_u^2 / n\sigma_{ew}^2}}$$

$$\therefore n = \frac{\sigma_u^2 . \tau^2}{\sigma_{ew}^2 . \hat{\beta}^2}$$

Where:

σ_u^2 = the variance of the errors from the regression used to estimate the treatment effect (from (A))

σ_{ew}^2 = the variance of the part of the treatment effect indicator w_i unexplained by the other variables (from (C))

Consequently, the inputs required to perform this calculation are:

1. An "estimate" of the impact of the intervention on academic results relative to the comparator group
2. An "estimate" of the variance[46] of the errors from the regression used to estimate the treatment effect
3. An "estimate" of the variance of the errors from the regression used to estimate the role of the other determinants of y_i in explaining the treatment effect.

[46] The variance is the square of the standard error, which in turn measures the deviation between sample means and the population mean.

This formula therefore indicates the number of observations which would need to be taken into account, to ensure that with some specified level of certainty (which determines the value of τ), we can determine the incremental impact of the intervention on the future academic performance of the pupils who took part.

Minimum sample size

The full results of regressions (B), (C) & (D) are presented in the Annex to this appendix. The inputs obtained from these regressions, which are required to estimate the required sample size are as follows:

$$\hat{\beta} = 0.216$$

$$\sigma_u^2 = 0.456$$

$$\sigma_{ew}^2 = 0.240$$

$$\tau = 1.96 \text{ (for 95\% confidence)}$$

Figure 12: Inputs for power calculation

Plugging these values into the expression for n above, given the desired level of τ (in turn, determined by the level of statistical significance the researcher wishes to achieve) it is possible to determine the minimum required sample size *n.* To achieve 95% significance, the sample size required would be **157**.

Optimal sample size

Given the importance of this part of the analysis to informing the next steps, it seems appropriate to perform some sensitivity analysis around the sample size. As already mentioned, the 95% confidence interval around the mean effect size is reasonably broad. If the actual effect size for all pupils was at the lower bound of the confidence interval (0.04), what size would the sample have to be to pick up this impact? The inputs to the power calculation would remain the same – see Figure 12, although $\hat{\beta}$ would now be reduced to 0.04. For 95% significance, using the same power calculation formula, the required sample size now increases to 4,311. While this is an extreme result and unlikely to be wholly necessary, it does highlight the high sensitivity of the results of this power calculation.

In general, the larger the dataset the better, as accuracy should improve, but there are definite costs for SHINE in terms of time and effort associated with collecting ever more data. On balance therefore, it would therefore seem appropriate to aim for somewhere between the extreme point referred to above and the minimum. Consequently, a reasonable aim would be a sample of approx. **2,000 – 2,200** with a roughly even split between treated pupils and comparator pupils. Another reason a much larger sample size would be beneficial is that the role of repeating the intervention can then be revisited. As discussed in the section below, it was not possible to draw any robust conclusions on this using the data for the pilot study, as the dataset was too small.

One final important consideration of any further econometric analysis is the need to include multiple schools - as discussed, the results of this pilot could include some specific "school effect" which is not being captured. Therefore, it would seem sensible to obtain the target sample from across approx. 10-12 schools. This ensures a good level of variation, but means that should it be necessary to consider the results at a school level, these samples should be sufficient to generate robust results (i.e. at least 160 observations per school).

FURTHER ECONOMETRIC ANALYSIS

RESULTS OF ALTERNATIVE MODEL SPECIFICATIONS

We have run a number of alternative specifications including or excluding a selection of different control variables. As shown in Table 6 below, the estimated size of the treatment effect is affected mainly by the inclusion of demographic and socioeconomic variables (gender, FSM status, SEN status, EAL status), which brings the treatment effect down from 0.281 (see Table 5 above) to 0.231 (see model (1) in Table 6 below), and of cohort and month of birth dummies, which further reduces the estimated treatment effect to around 0.21 depending on the exact specification (models (3) to (6) below). The main specification has been chosen over these alternative models, as the cohort dummies by year and the individual month of birth dummies are largely insignificant in this form, but are significant when used in a more aggregated form (month of birth aggregated by season, cohort dummy aggregated according to whether KS2 was sat in 05/06 & 06/07 or later in 07/08 – 13/14.

VARIABLES	(3) KS2avg	(4) KS2avg	(5) KS2avg	(6) KS2avg
Treatment effect	0.231** (0.0902)	0.246*** (0.0913)	0.228** (0.0897)	0.213** (0.0918)
Gender	-0.109 (0.0886)	-0.116 (0.0898)	-0.130 (0.0886)	-0.131 (0.0905)
FSM	-0.270** (0.126)	-0.286** (0.129)	-0.298** (0.126)	-0.298** (0.133)
EAL	-0.0478 (0.0925)	-0.0871 (0.0944)	-0.0393 (0.0921)	-0.0547 (0.0963)
SEN	-0.391*** (0.102)	-0.471*** (0.106)	-0.428*** (0.103)	-0.436*** (0.105)
KS2 in 05/06 or 06/07			0.235** (0.116)	0.200* (0.120)
Born in February				0.0771 (0.245)
Born in March				-0.206 (0.206)
Born in April				-0.275 (0.199)
Born in May				-0.178 (0.212)
Born in June				-0.344 (0.214)
Born in July				-0.194 (0.222)
Born in August				-0.698*** (0.240)
Born in September				-0.205 (0.206)
Born in October				-0.118 (0.228)
Born in November				-0.231 (0.206)
Born in December				-0.0153 (0.188)

KS2 in 06/07		-0.263 (0.215)		
KS2 in 07/08		-0.401** (0.202)		
KS2 in 08/09		-0.244 (0.232)		
KS2 in 09/10		-0.438** (0.202)		
KS2 in 10/11		-0.568*** (0.187)		
KS2 in 11/12		-0.218 (0.187)		
KS2 in 12/13		-0.287 (0.184)		
KS2 in 13/14		-0.285 (0.180)		
Constant	4.291*** (0.176)	4.674*** (0.240)	4.297*** (0.175)	4.509*** (0.250)
Observations	257	257	257	257
R-squared	0.094	0.134	0.108	0.160

Table 6: Alternative model specifications with homogeneous treatment effect[47]

HETEROGENEITY IN THE TREATMENT EFFECT

Effect of attendance rate

Table 7 below investigates variation in the effect of attending the SHINE on Saturday programme depending on the frequency of attendance. In each specification, the effect of treatment is estimated separately for two cases: 'high-frequency' attendees (coefficient on the 'treated' variable), and 'low-frequency' attendees (coefficient on 'treated' + coefficient on 'low_attend', a dummy for attendance below a certain threshold). The threshold for low attendance varies across the three specifications. Model (9) suggests that the average effect across all attendees seen in the tables above is driven by pupils who attended nearly all sessions. For these pupils, the effect of treatment is nearly double the size of the average effect; for other pupils (with lower SHINE attendance), the effect is much smaller. This finding, however, needs to be interpreted with caution given the relatively small sample of very high frequency pupils (i.e. 47 pupils had an attendance rate > 95%).

VARIABLES	(7) Low attendance (<80%) KS2avg	(8) Low attendance (<90%) KS2avg	(9) Low attendance (<95%) KS2avg
Treatment effect	0.216** (0.0910)	0.260** (0.103)	0.409*** (0.122)
Gender	-0.118 (0.0881)	-0.120 (0.0880)	-0.1000 (0.0875)

[47] *** p<0.01, ** p<0.05, * p<0.1; standard errors in parentheses

FSM	-0.287** (0.126)	-0.282** (0.126)	-0.282** (0.124)
EAL	-0.0569 (0.0932)	-0.0659 (0.0936)	-0.0826 (0.0929)
SEN	-0.441*** (0.102)	-0.436*** (0.102)	-0.440*** (0.101)
Born in spring	-0.229* (0.119)	-0.228* (0.119)	-0.223* (0.118)
Born in summer	-0.391*** (0.126)	-0.387*** (0.126)	-0.365*** (0.126)
Born in autumn	-0.197 (0.120)	-0.200* (0.119)	-0.184 (0.118)
Attendance rate less than specified threshold	0.00247 (0.208)	-0.102 (0.115)	-0.284** (0.122)
KS2 in 05/06 or 06/07	0.217* (0.116)	0.213* (0.116)	0.207* (0.115)
Constant	4.499*** (0.189)	4.499*** (0.188)	4.494*** (0.187)
Observations	257	257	257
R-squared	0.143	0.146	0.162

Table 7: Results by attendance level[48]

Effect of repeating the programme

We also investigated whether the effect of treatment varies depending on how many years the programme was attended for and also considered which school years the pupil was in when they attended the SHINE on Saturday programme. Results for model (10) in Table 8 below suggest that the effect of treatment may be strongest for pupils who attended SHINE sessions in year 6, regardless of the number of times the programme was attended (the base case within the treatment group). However, two issues need to be taken into account:

- The year when SHINE on Saturday was attended is not independent of the number of years in which the pupil attended SHINE on Saturday.
- The size of the sample at hand limits the number of interactions that can be investigated with confidence.

Model (12) below separates out the effect of treatment according to the school year in which the intervention was received and the number of years of treatment. The signs and magnitudes of the coefficients would suggest that, as considered above, the effect of treatment is strongest with attendance in year 6, lower for previous years, and repeat attendance has a limited effect. However, given the limited sample size, the standard errors of these coefficients are large relative to their estimated magnitude. Drawing robust conclusions on the heterogeneity of treatment effects by school year and frequency of treatment would definitely require further research, using a larger sample size and therefore we can’t conclude on the role of repeating SHINE on Saturday in determining any improvement in K2 results.

[48] *** p<0.01, ** p<0.05, * p<0.1; standard errors in parentheses

VARIABLES	(10) KS2avg	(11) KS2avg	(12) KS2avg
Treatment effect	0.332*** (0.125)	0.310*** (0.108)	0.326** (0.131)
Received SHINE in yr 4	-0.0705 (0.125)		-0.0227 (0.164)
Received SHINE in yr 5	-0.140 (0.120)		-0.0429 (0.190)
Received SHINE x2		-0.183 (0.0141)	-0.152 (0.195)
Received SHINE x3		-0.184 (0.0141)	-0.136 (0.257)
Controls	Constant, Gender, FSM, SEN, EAL, Single cohort dummy, Season of birth	Constant, Gender, FSM, SEN, EAL, Single cohort dummy, Season of birth	Constant, Gender, FSM, SEN, EAL, Single cohort dummy, Season of birth
Observations	257	257	257
R-squared	0.154	0.152	0.155

Table 8: Results by timing & frequency of programme[49]

[49] *** p<0.01, ** p<0.05, * p<0.1; standard errors in parentheses

Source	SS	df	MS			
Model	16.728676	8	2.091085			
Residual	119.38805	248	0.481403			
Total	136.11673	256	0.531706			

Number of obs	257
F(8, 248)	4.34
Prob > F	0.0001
R-squared	0.1229
Adj R-squared	0.0946
Root MSE	0.69383

KS2avg	Coef.	Std. Err.	t	P>|t|	[95% Conf. interval]	
Gender	-.1128793	0.088781	-1.27	0.205	-0.2877406	0.061982
FSM	-.284898	0.126699	-2.25	0.025	-0.5344404	-0.0353556
EAL	-.0549279	0.093917	-0.58	0.559	-0.2399047	0.1300488
SEN	-.4816207	0.10099	-4.77	0	-0.6805291	-0.2827124
KS2 early	.2183895	0.116976	1.87	0.063	-0.012004	0.4487831
Born spring	-.2581572	0.11941	-2.16	0.032	-0.4933436	-0.0229709
Born summer	-.3970186	0.127372	-3.12	0.002	-0.6478882	-0.146149
Born autumn	-.2261594	0.119647	-1.89	0.06	-0.4618124	0.0094935
Constant	4.655578	0.178705	26.05	0	4.303605	5.00755

Figure 13: Results of regression B

Source	SS	df	MS
Model	3.36356927	8	0.420446
Residual	59.4068588	248	0.239544
Total	62.770428	256	0.245197

Number of obs	257
F(8, 248)	1.76
Prob > F	0.0865
R-squared	0.0536
Adj R-squared	0.0231
Root MSE	0.48943

Treatment effect	Coef.	Std. Err.	t	P>|t|	[95% Conf. interval]	
Gender	0.0250706	0.062627	0.4	0.689	-0.0982773	0.1484184
FSM	.0083365	0.089374	0.09	0.926	-0.1676917	0.1843647
EAL	.0091976	0.06625	0.14	0.89	-0.1212857	0.139681
SEN	-.1866841	0.071239	-2.62	0.009	-0.3269948	-0.0463734
KS2 early	.0044559	0.082516	0.05	0.957	-0.1580647	0.1669764
Born spring	-.133672	0.084232	-1.59	0.114	-0.2995734	0.0322293
Born summer	-.0296327	0.089849	-0.33	0.742	-0.2065971	0.1473317
Born autumn	-.1348833	0.084399	-1.6	0.111	-0.3011138	0.0313472
Constant	4.655578	0.178705	26.05	0	4.303605	5.00755

Figure 14: Results of regression C

Source	SS	df	MS
Model	2.76450817	1	2.764508
Residual	116.623547	256	0.455561
Total	119.388055	257	0.464545

Number of obs	257
F(8, 248)	6.07
Prob > F	0.0144
R-squared	0.0232
Adj R-squared	0.0193
Root MSE	0.67495

Errors - KS2avg	Coef.	Std. Err.	t	P>|t|	[95% Conf. interval]	
Errors - treatment effect	0.2157201	0.08757	2.46	0.014	0.043271	0.3881692

Figure 15: Results of regression D

APPENDIX 5

CONDITIONAL KS4 PROBABILITIES & EXPECTED KS4 RESULTS

CONDITIONAL KS4 PROBABILITIES

The table below sets out the conditional probabilities of different KS4 results for primary school pupils in London, conditional on their KS2 results and certain pupil characteristics (i.e. gender and indicators of disadvantage). KS4 results are measured as ranges of numbers of GCSEs at A*- C, to coincide with the way GCSE results are measured in the DfE paper referred to in Appendix 6. These probabilities were calculated using annual data provided by FFT which related to all pupils in London with these various combinations of characteristics and KS2 results who sat their GCSEs during the period 2011 – 2015. Note that the original raw data from FFT was adjusted to replace any supressed "small number" with an estimate of 3 (i.e. a "small number" is less than 5, so as 3 is the mid-point between 0 and 5, this was used as a proxy for the missing data).

2011:

KS4 year	KS2 year	Av. KS2	Group	Conditional probabilities (%)															
				Female / FSM only	Male / FSM only	Female / FSM + SEN	Male / FSM + SEN	Female / FSM + SEN + EAL	Male / FSM + SEN + EAL	Female / SEN only	Male / SEN only	Female / SEN + EAL	Male / SEN + EAL	Female / FSM + EAL	Male / FSM + EAL	Female / EAL only	Male / EAL only
2010/11	2005/06	<2	gcse0 - pupils getting 0 A*-C	78%	76%	79%	88%	71%	90%	78%	76%	70%	81%	70%	77%	64%	72%
		<2	gcse2 - pupils getting 1-2 A*-C	5%	16%	7%	6%	10%	10%	9%	18%	11%	14%	11%	17%	19%	17%
		<2	gcse4 - pupils getting 3-4 A*-C	5%	0%	7%	0%	10%	0%	4%	3%	6%	5%	7%	0%	6%	3%
		<2	gcse7 - pupils getting 5-7 A*-C	5%	4%	7%	6%	10%	0%	4%	3%	6%	0%	7%	0%	4%	5%
		<2	gcse8 - pupils getting 8+ A*-C	5%	4%	0%	0%	0%	0%	4%	0%	6%	0%	7%	6%	7%	3%
		2	gcse0 - pupils getting 0 A*-C	62%	72%	63%	74%	52%	64%	67%	73%	53%	64%	45%	64%	46%	62%
		2	gcse2 - pupils getting 1-2 A*-C	25%	19%	24%	17%	34%	23%	23%	18%	27%	23%	26%	25%	24%	25%
		2	gcse4 - pupils getting 3-4 A*-C	8%	6%	6%	5%	6%	8%	6%	6%	7%	7%	7%	9%	9%	8%
		2	gcse7 - pupils getting 5-7 A*-C	5%	3%	1%	3%	6%	2%	1%	2%	7%	3%	6%	1%	8%	4%
		2	gcse8 - pupils getting 8+ A*-C	1%	1%	6%	2%	2%	3%	4%	1%	6%	2%	15%	1%	14%	1%
		3c	gcse0 - pupils getting 0 A*-C	49%	63%	57%	68%	49%	54%	55%	64%	45%	51%	38%	49%	34%	46%
		3c	gcse2 - pupils getting 1-2 A*-C	19%	16%	19%	15%	21%	22%	21%	17%	24%	22%	19%	22%	22%	21%
		3c	gcse4 - pupils getting 3-4 A*-C	10%	9%	9%	8%	8%	10%	10%	7%	10%	10%	10%	12%	12%	12%
		3c	gcse7 - pupils getting 5-7 A*-C	9%	5%	7%	4%	8%	6%	6%	5%	9%	7%	11%	7%	12%	8%
		3c	gcse8 - pupils getting 8+ A*-C	13%	7%	8%	5%	14%	8%	8%	6%	12%	10%	21%	11%	21%	14%
		3b	gcse0 - pupils getting 0 A*-C	36%	47%	43%	49%	25%	37%	41%	49%	24%	39%	20%	35%	19%	34%
		3b	gcse2 - pupils getting 1-2 A*-C	17%	24%	19%	23%	22%	29%	19%	22%	19%	26%	20%	27%	18%	25%
		3b	gcse4 - pupils getting 3-4 A*-C	15%	12%	14%	12%	16%	14%	14%	11%	15%	13%	16%	15%	15%	14%
		3b	gcse7 - pupils getting 5-7 A*-C	10%	6%	8%	6%	11%	7%	9%	7%	13%	9%	12%	8%	13%	9%
		3b	gcse8 - pupils getting 8+ A*-C	22%	11%	16%	10%	26%	12%	17%	11%	28%	14%	32%	15%	35%	18%
		3a	gcse0 - pupils getting 0 A*-C	28%	36%	34%	44%	16%	26%	29%	39%	15%	23%	13%	21%	12%	20%
		3a	gcse2 - pupils getting 1-2 A*-C	15%	18%	16%	19%	16%	18%	16%	19%	14%	19%	14%	18%	12%	18%
		3a	gcse4 - pupils getting 3-4 A*-C	14%	13%	13%	12%	15%	16%	13%	13%	14%	15%	14%	15%	13%	14%
		3a	gcse7 - pupils getting 5-7 A*-C	10%	10%	9%	9%	12%	12%	10%	9%	10%	12%	11%	12%	10%	12%
		3a	gcse8 - pupils getting 8+ A*-C	33%	23%	27%	17%	41%	28%	32%	21%	47%	31%	48%	34%	52%	36%
		4c	gcse0 - pupils getting 0 A*-C	17%	26%	24%	34%	10%	22%	20%	26%	9%	18%	7%	15%	6%	12%
		4c	gcse2 - pupils getting 1-2 A*-C	12%	15%	15%	17%	11%	16%	13%	17%	9%	13%	9%	13%	7%	12%
		4c	gcse4 - pupils getting 3-4 A*-C	11%	13%	12%	12%	11%	14%	11%	13%	9%	14%	9%	13%	8%	13%
		4c	gcse7 - pupils getting 5-7 A*-C	11%	10%	10%	10%	12%	9%	11%	10%	12%	10%	12%	11%	11%	12%
		4c	gcse8 - pupils getting 8+ A*-C	49%	36%	40%	27%	56%	39%	45%	34%	61%	44%	64%	48%	69%	52%
		4b	gcse0 - pupils getting 0 A*-C	10%	15%	15%	21%	2%	13%	11%	15%	5%	9%	3%	8%	2%	6%
		4b	gcse2 - pupils getting 1-2 A*-C	7%	11%	10%	13%	7%	9%	9%	11%	6%	8%	4%	7%	3%	6%
		4b	gcse4 - pupils getting 3-4 A*-C	9%	10%	12%	11%	9%	8%	10%	12%	6%	9%	5%	7%	4%	7%
		4b	gcse7 - pupils getting 5-7 A*-C	9%	10%	8%	9%	9%	8%	10%	9%	9%	8%	8%	8%	7%	7%
		4b	gcse8 - pupils getting 8+ A*-C	66%	55%	54%	45%	73%	62%	60%	52%	74%	66%	80%	69%	83%	73%
		4a	gcse0 - pupils getting 0 A*-C	4%	6%	10%	12%	5%	2%	1%	9%	3%	5%	1%	0%	1%	2%
		4a	gcse2 - pupils getting 1-2 A*-C	4%	6%	8%	11%	5%	6%	6%	7%	3%	5%	2%	3%	1%	3%
		4a	gcse4 - pupils getting 3-4 A*-C	5%	6%	8%	6%	5%	2%	5%	7%	2%	4%	2%	4%	2%	3%
		4a	gcse7 - pupils getting 5-7 A*-C	6%	7%	8%	10%	5%	8%	7%	8%	2%	7%	3%	5%	2%	5%
		4a	gcse8 - pupils getting 8+ A*-C	82%	75%	66%	61%	80%	81%	82%	69%	90%	78%	92%	87%	94%	88%
		5c	gcse0 - pupils getting 0 A*-C	0%	0%	0%	11%	0%	8%	0%	4%	0%	4%	1%	1%	0%	1%
		5c	gcse2 - pupils getting 1-2 A*-C	1%	2%	8%	3%	0%	8%	2%	2%	0%	2%	1%	1%	0%	0%
		5c	gcse4 - pupils getting 3-4 A*-C	2%	3%	8%	7%	0%	8%	2%	4%	0%	2%	1%	1%	0%	1%
		5c	gcse7 - pupils getting 5-7 A*-C	2%	5%	0%	3%	0%	8%	0%	4%	0%	4%	1%	3%	0%	2%
		5c	gcse8 - pupils getting 8+ A*-C	95%	90%	84%	76%	100%	69%	97%	86%	100%	87%	98%	94%	99%	97%
		5b	gcse0 - pupils getting 0 A*-C	1%	1%	0%	0%	0%	0%	5%	2%	0%	0%	1%	0%	0%	0%
		5b	gcse2 - pupils getting 1-2 A*-C	1%	0%	0%	0%	0%	0%	5%	3%	0%	8%	1%	0%	0%	0%
		5b	gcse4 - pupils getting 3-4 A*-C	1%	1%	0%	0%	0%	0%	0%	2%	0%	0%	1%	0%	0%	0%
		5b	gcse7 - pupils getting 5-7 A*-C	1%	1%	0%	0%	0%	0%	0%	2%	0%	0%	1%	1%	0%	1%
		5b	gcse8 - pupils getting 8+ A*-C	97%	98%	100%	100%	100%	100%	89%	91%	100%	92%	95%	99%	99%	99%
		5a	gcse0 - pupils getting 0 A*-C	0%	0%	0%	0%	0%	0%	0%	18%	0%	0%	0%	0%	0%	0%
		5a	gcse2 - pupils getting 1-2 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse4 - pupils getting 3-4 A*-C	0%	13%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%
		5a	gcse7 - pupils getting 5-7 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse8 - pupils getting 8+ A*-C	100%	87%	100%	0%	0%	0%	100%	82%	0%	100%	100%	100%	98%	100%

Figure 16: Conditional probabilities of KS4 results by KS2 result & pupil characteristics, in London schools in 2011

2012:

KS4 year	KS2 year	Av. KS2	Group	Conditional probabilities (%)													
				Female / FSM only	Male / FSM only	Female / FSM + SEN	Male / FSM + SEN	Female / FSM + SEN + EAL	Male / FSM + SEN + EAL	Female / SEN only	Male / SEN only	Female / SEN + EAL	Male / SEN + EAL	Female / FSM + EAL	Male / FSM + EAL	Female / EAL only	Male / EAL only
2011/12	2006/07	<2	gcse0 - pupils getting 0 A*-C	81%	76%	85%	69%	80%	76%	85%	85%	78%	73%	77%	76%	64%	66%
		<2	gcse2 - pupils getting 1-2 A*-C	5%	5%	7%	13%	10%	8%	7%	3%	10%	13%	6%	6%	15%	14%
		<2	gcse4 - pupils getting 3-4 A*-C	5%	8%	0%	6%	0%	8%	0%	5%	0%	5%	6%	6%	3%	3%
		<2	gcse7 - pupils getting 5-7 A*-C	5%	5%	7%	6%	10%	0%	4%	3%	6%	5%	6%	6%	3%	3%
		<2	gcse8 - pupils getting 8+ A*-C	5%	5%	0%	6%	0%	8%	4%	3%	6%	5%	6%	6%	15%	13%
		2	gcse0 - pupils getting 0 A*-C	64%	64%	67%	69%	58%	60%	67%	70%	57%	64%	50%	55%	44%	54%
		2	gcse2 - pupils getting 1-2 A*-C	19%	21%	19%	22%	22%	26%	17%	19%	21%	22%	20%	24%	21%	23%
		2	gcse4 - pupils getting 3-4 A*-C	10%	6%	9%	5%	13%	8%	9%	5%	15%	7%	13%	8%	15%	9%
		2	gcse7 - pupils getting 5-7 A*-C	5%	3%	1%	2%	2%	4%	3%	3%	1%	3%	7%	4%	6%	4%
		2	gcse8 - pupils getting 8+ A*-C	1%	6%	4%	2%	6%	2%	4%	2%	6%	3%	10%	8%	14%	9%
		3c	gcse0 - pupils getting 0 A*-C	46%	64%	50%	66%	35%	55%	50%	64%	35%	54%	28%	48%	26%	47%
		3c	gcse2 - pupils getting 1-2 A*-C	22%	20%	23%	16%	26%	19%	23%	18%	28%	19%	24%	23%	26%	21%
		3c	gcse4 - pupils getting 3-4 A*-C	15%	11%	13%	10%	16%	15%	13%	9%	15%	13%	19%	15%	15%	15%
		3c	gcse7 - pupils getting 5-7 A*-C	5%	4%	5%	4%	9%	5%	5%	5%	9%	6%	8%	6%	11%	6%
		3c	gcse8 - pupils getting 8+ A*-C	13%	1%	8%	4%	14%	6%	9%	5%	13%	7%	21%	8%	22%	11%
		3b	gcse0 - pupils getting 0 A*-C	40%	49%	43%	52%	27%	40%	42%	52%	27%	39%	25%	37%	23%	33%
		3b	gcse2 - pupils getting 1-2 A*-C	18%	23%	19%	23%	18%	27%	20%	22%	20%	25%	17%	27%	19%	24%
		3b	gcse4 - pupils getting 3-4 A*-C	14%	10%	14%	10%	18%	12%	14%	10%	19%	13%	17%	12%	16%	13%
		3b	gcse7 - pupils getting 5-7 A*-C	10%	7%	9%	6%	13%	7%	9%	7%	11%	10%	13%	8%	12%	9%
		3b	gcse8 - pupils getting 8+ A*-C	19%	11%	14%	10%	24%	13%	15%	10%	23%	14%	28%	15%	31%	21%
		3a	gcse0 - pupils getting 0 A*-C	25%	35%	30%	41%	17%	28%	28%	40%	17%	26%	14%	24%	13%	21%
		3a	gcse2 - pupils getting 1-2 A*-C	16%	21%	17%	19%	16%	19%	17%	19%	17%	20%	14%	20%	13%	19%
		3a	gcse4 - pupils getting 3-4 A*-C	13%	15%	14%	14%	17%	17%	14%	14%	14%	16%	14%	16%	13%	15%
		3a	gcse7 - pupils getting 5-7 A*-C	12%	10%	12%	10%	13%	13%	11%	9%	14%	12%	13%	11%	13%	11%
		3a	gcse8 - pupils getting 8+ A*-C	33%	21%	27%	16%	36%	23%	29%	19%	39%	26%	45%	28%	49%	33%
		4c	gcse0 - pupils getting 0 A*-C	15%	24%	19%	31%	11%	19%	17%	26%	11%	15%	7%	14%	5%	11%
		4c	gcse2 - pupils getting 1-2 A*-C	9%	16%	12%	18%	8%	15%	12%	17%	8%	15%	6%	12%	6%	12%
		4c	gcse4 - pupils getting 3-4 A*-C	12%	13%	14%	13%	12%	17%	12%	13%	10%	15%	10%	14%	8%	12%
		4c	gcse7 - pupils getting 5-7 A*-C	10%	11%	10%	12%	8%	13%	12%	13%	8%	13%	9%	11%	9%	11%
		4c	gcse8 - pupils getting 8+ A*-C	54%	36%	45%	25%	61%	36%	49%	31%	64%	41%	68%	49%	72%	53%
		4b	gcse0 - pupils getting 0 A*-C	9%	11%	13%	18%	6%	9%	9%	14%	5%	8%	3%	6%	2%	5%
		4b	gcse2 - pupils getting 1-2 A*-C	7%	11%	9%	14%	8%	10%	8%	13%	6%	9%	4%	7%	3%	6%
		4b	gcse4 - pupils getting 3-4 A*-C	7%	13%	8%	16%	2%	15%	7%	13%	6%	13%	4%	11%	4%	9%
		4b	gcse7 - pupils getting 5-7 A*-C	7%	11%	10%	11%	10%	9%	10%	11%	9%	9%	6%	10%	6%	9%
		4b	gcse8 - pupils getting 8+ A*-C	71%	54%	59%	42%	73%	57%	66%	48%	75%	60%	83%	66%	85%	71%
		4a	gcse0 - pupils getting 0 A*-C	4%	5%	2%	8%	5%	2%	6%	6%	5%	1%	2%	2%	1%	1%
		4a	gcse2 - pupils getting 1-2 A*-C	3%	6%	2%	9%	3%	6%	5%	7%	4%	5%	1%	4%	1%	2%
		4a	gcse4 - pupils getting 3-4 A*-C	5%	7%	2%	11%	3%	11%	1%	8%	3%	7%	3%	6%	2%	4%
		4a	gcse7 - pupils getting 5-7 A*-C	5%	8%	12%	10%	10%	8%	7%	9%	7%	8%	4%	6%	3%	5%
		4a	gcse8 - pupils getting 8+ A*-C	84%	75%	81%	63%	78%	73%	82%	70%	81%	79%	91%	83%	94%	87%
		5c	gcse0 - pupils getting 0 A*-C	1%	2%	5%	5%	0%	4%	3%	3%	0%	2%	0%	1%	0%	0%
		5c	gcse2 - pupils getting 1-2 A*-C	2%	2%	0%	2%	0%	4%	2%	3%	0%	2%	0%	0%	0%	0%
		5c	gcse4 - pupils getting 3-4 A*-C	2%	3%	5%	6%	0%	4%	2%	3%	0%	3%	0%	2%	0%	1%
		5c	gcse7 - pupils getting 5-7 A*-C	2%	4%	0%	6%	0%	4%	2%	5%	0%	5%	0%	3%	1%	2%
		5c	gcse8 - pupils getting 8+ A*-C	94%	89%	89%	81%	100%	83%	92%	86%	100%	89%	98%	94%	99%	97%
		5b	gcse0 - pupils getting 0 A*-C	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%
		5b	gcse2 - pupils getting 1-2 A*-C	1%	1%	0%	7%	0%	15%	5%	2%	0%	6%	0%	1%	0%	0%
		5b	gcse4 - pupils getting 3-4 A*-C	1%	2%	0%	7%	0%	15%	0%	3%	0%	6%	0%	1%	0%	0%
		5b	gcse7 - pupils getting 5-7 A*-C	1%	1%	0%	7%	0%	15%	0%	2%	0%	6%	0%	1%	0%	0%
		5b	gcse8 - pupils getting 8+ A*-C	97%	97%	100%	79%	100%	55%	95%	94%	100%	83%	99%	96%	100%	99%
		5a	gcse0 - pupils getting 0 A*-C	9%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse2 - pupils getting 1-2 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse4 - pupils getting 3-4 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse7 - pupils getting 5-7 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%
		5a	gcse8 - pupils getting 8+ A*-C	91%	100%	0%	100%	0%	0%	100%	100%	100%	100%	100%	100%	98%	100%

Figure 17: Conditional probabilities of KS4 results by KS2 result & pupil characteristics, in London schools in 2012

2013:

KS4 year	KS2 year	Av. KS2	Group	Conditional probabilities (%) Female / FSM only	Male / FSM only	Female / FSM + SEN	Male / FSM + SEN	Female / FSM + SEN + EAL	Male / FSM + SEN + EAL	Female / SEN only	Male / SEN only	Female / SEN + EAL	Male / SEN + EAL	Female / FSM + EAL	Male / FSM + EAL	Female / EAL only	Male / EAL only
2012/13	2007/08	<2	gcse0 - pupils getting 0 A*-C	82%	82%	84%	84%	72%	77%	76%	88%	78%	78%	78%	76%	58%	66%
		<2	gcse2 - pupils getting 1-2 A*-C	6%	6%	8%	8%	19%	12%	15%	4%	7%	7%	8%	8%	20%	24%
		<2	gcse4 - pupils getting 3-4 A*-C	6%	0%	8%	0%	9%	0%	5%	4%	7%	7%	8%	0%	10%	3%
		<2	gcse7 - pupils getting 5-7 A*-C	0%	6%	0%	8%	0%	12%	0%	4%	0%	7%	0%	8%	3%	3%
		<2	gcse8 - pupils getting 8+ A*-C	6%	6%	0%	0%	0%	0%	5%	0%	7%	0%	8%	8%	8%	3%
		2	gcse0 - pupils getting 0 A*-C	71%	72%	71%	78%	52%	71%	68%	75%	56%	69%	51%	60%	44%	54%
		2	gcse2 - pupils getting 1-2 A*-C	17%	18%	15%	15%	24%	20%	20%	16%	26%	21%	22%	22%	23%	23%
		2	gcse4 - pupils getting 3-4 A*-C	10%	1%	9%	3%	14%	2%	8%	5%	11%	6%	13%	5%	12%	8%
		2	gcse7 - pupils getting 5-7 A*-C	1%	4%	1%	1%	4%	3%	4%	2%	5%	2%	6%	5%	7%	4%
		2	gcse8 - pupils getting 8+ A*-C	1%	6%	4%	2%	7%	4%	1%	2%	1%	1%	8%	8%	14%	11%
		3c	gcse0 - pupils getting 0 A*-C	49%	56%	50%	60%	34%	52%	54%	62%	38%	51%	36%	49%	33%	54%
		3c	gcse2 - pupils getting 1-2 A*-C	20%	21%	20%	19%	25%	22%	21%	19%	23%	23%	22%	23%	22%	25%
		3c	gcse4 - pupils getting 3-4 A*-C	13%	10%	14%	10%	15%	15%	11%	10%	15%	13%	14%	14%	14%	15%
		3c	gcse7 - pupils getting 5-7 A*-C	8%	4%	7%	3%	10%	2%	7%	3%	10%	3%	11%	2%	11%	5%
		3c	gcse8 - pupils getting 8+ A*-C	10%	9%	8%	7%	17%	9%	7%	7%	14%	9%	17%	12%	20%	1%
		3b	gcse0 - pupils getting 0 A*-C	41%	51%	46%	55%	29%	40%	42%	52%	29%	37%	25%	36%	23%	32%
		3b	gcse2 - pupils getting 1-2 A*-C	18%	23%	20%	23%	22%	27%	21%	23%	21%	28%	20%	27%	18%	26%
		3b	gcse4 - pupils getting 3-4 A*-C	13%	11%	13%	10%	16%	13%	13%	11%	17%	16%	14%	15%	16%	16%
		3b	gcse7 - pupils getting 5-7 A*-C	10%	5%	8%	3%	11%	5%	9%	5%	12%	6%	13%	8%	14%	8%
		3b	gcse8 - pupils getting 8+ A*-C	18%	10%	13%	9%	22%	14%	15%	9%	21%	13%	28%	15%	28%	17%
		3a	gcse0 - pupils getting 0 A*-C	25%	37%	31%	42%	21%	31%	27%	38%	17%	29%	16%	28%	13%	23%
		3a	gcse2 - pupils getting 1-2 A*-C	17%	20%	18%	21%	14%	19%	18%	22%	13%	19%	12%	18%	11%	17%
		3a	gcse4 - pupils getting 3-4 A*-C	14%	13%	14%	12%	15%	16%	15%	13%	14%	17%	15%	16%	14%	16%
		3a	gcse7 - pupils getting 5-7 A*-C	11%	10%	11%	11%	12%	13%	11%	10%	11%	11%	12%	12%	11%	11%
		3a	gcse8 - pupils getting 8+ A*-C	33%	19%	27%	14%	38%	21%	30%	17%	45%	24%	45%	27%	51%	32%
		4c	gcse0 - pupils getting 0 A*-C	14%	22%	19%	28%	9%	17%	16%	24%	8%	14%	6%	13%	6%	10%
		4c	gcse2 - pupils getting 1-2 A*-C	13%	16%	14%	18%	9%	17%	13%	18%	10%	16%	8%	15%	7%	14%
		4c	gcse4 - pupils getting 3-4 A*-C	11%	14%	12%	13%	10%	13%	11%	13%	11%	14%	9%	14%	9%	12%
		4c	gcse7 - pupils getting 5-7 A*-C	11%	11%	11%	10%	12%	13%	12%	12%	9%	12%	10%	11%	8%	11%
		4c	gcse8 - pupils getting 8+ A*-C	52%	36%	44%	31%	60%	41%	48%	33%	62%	45%	67%	47%	70%	52%
		4b	gcse0 - pupils getting 0 A*-C	8%	12%	15%	18%	2%	10%	11%	14%	6%	7%	3%	6%	2%	4%
		4b	gcse2 - pupils getting 1-2 A*-C	7%	11%	8%	13%	6%	10%	7%	11%	6%	8%	4%	8%	3%	7%
		4b	gcse4 - pupils getting 3-4 A*-C	8%	11%	11%	11%	8%	11%	9%	11%	6%	11%	5%	9%	4%	8%
		4b	gcse7 - pupils getting 5-7 A*-C	9%	11%	10%	12%	10%	10%	9%	10%	8%	9%	8%	8%	6%	8%
		4b	gcse8 - pupils getting 8+ A*-C	67%	55%	56%	46%	74%	59%	63%	54%	73%	65%	80%	69%	84%	74%
		4a	gcse0 - pupils getting 0 A*-C	3%	5%	1%	8%	3%	2%	5%	6%	2%	3%	0%	0%	0%	2%
		4a	gcse2 - pupils getting 1-2 A*-C	3%	6%	9%	9%	3%	2%	5%	7%	4%	6%	2%	3%	1%	3%
		4a	gcse4 - pupils getting 3-4 A*-C	3%	7%	7%	10%	3%	9%	5%	8%	1%	7%	0%	5%	1%	3%
		4a	gcse7 - pupils getting 5-7 A*-C	5%	7%	5%	10%	5%	9%	4%	7%	1%	6%	4%	6%	2%	4%
		4a	gcse8 - pupils getting 8+ A*-C	85%	75%	77%	64%	86%	79%	82%	72%	92%	80%	94%	85%	95%	88%
		5c	gcse0 - pupils getting 0 A*-C	1%	1%	5%	5%	0%	5%	4%	0%	0%	2%	0%	0%	0%	0%
		5c	gcse2 - pupils getting 1-2 A*-C	1%	2%	5%	3%	11%	0%	2%	3%	4%	2%	0%	0%	0%	0%
		5c	gcse4 - pupils getting 3-4 A*-C	0%	2%	5%	3%	0%	5%	1%	3%	0%	2%	0%	0%	0%	1%
		5c	gcse7 - pupils getting 5-7 A*-C	1%	4%	0%	7%	0%	5%	1%	4%	4%	2%	0%	2%	0%	1%
		5c	gcse8 - pupils getting 8+ A*-C	96%	91%	84%	83%	89%	84%	92%	91%	93%	93%	99%	97%	99%	98%
		5b	gcse0 - pupils getting 0 A*-C	0%	0%	0%	0%	0%	0%	0%	2%	0%	6%	0%	0%	0%	0%
		5b	gcse2 - pupils getting 1-2 A*-C	1%	1%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	0%	0%
		5b	gcse4 - pupils getting 3-4 A*-C	0%	1%	0%	11%	0%	0%	0%	2%	0%	0%	0%	1%	0%	0%
		5b	gcse7 - pupils getting 5-7 A*-C	1%	1%	23%	0%	0%	0%	6%	2%	0%	6%	1%	0%	0%	0%
		5b	gcse8 - pupils getting 8+ A*-C	98%	98%	77%	89%	100%	100%	94%	92%	100%	88%	99%	99%	99%	99%
		5a	gcse0 - pupils getting 0 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse2 - pupils getting 1-2 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse4 - pupils getting 3-4 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse7 - pupils getting 5-7 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse8 - pupils getting 8+ A*-C	100%	100%	0%	0%	0%	0%	0%	100%	0%	100%	100%	100%	100%	100%

Figure 18: Conditional probabilities of KS4 results by KS2 result & pupil characteristics, in London schools in 2013

2014:

KS4 year	KS2 year	Av. KS2	Group	Conditional probabilities (%)													
				Female / FSM only	Male / FSM only	Female / FSM + SEN	Male / FSM + SEN	Female / FSM + SEN + EAL	Male / FSM + SEN + EAL	Female / SEN only	Male / SEN only	Female / SEN + EAL	Male / SEN + EAL	Female / FSM + EAL	Male / FSM + EAL	Female / EAL only	Male / EAL only
2013/14	2008/09	<2	gcse0 - pupils getting 0 A*-C	64%	81%	81%	82%	73%	70%	77%	90%	76%	78%	65%	67%	65%	64%
		<2	gcse2 - pupils getting 1-2 A*-C	18%	6%	9%	9%	14%	15%	13%	3%	8%	17%	9%	17%	3%	21%
		<2	gcse4 - pupils getting 3-4 A*-C	6%	6%	0%	0%	0%	0%	5%	3%	8%	0%	9%	8%	9%	3%
		<2	gcse7 - pupils getting 5-7 A*-C	6%	6%	9%	9%	14%	15%	5%	3%	8%	6%	9%	8%	3%	8%
		<2	gcse8 - pupils getting 8+ A*-C	6%	0%	0%	0%	0%	0%	0%	0%	0%	0%	9%	0%	19%	5%
		2	gcse0 - pupils getting 0 A*-C	61%	69%	62%	73%	51%	68%	64%	69%	53%	61%	47%	63%	43%	49%
		2	gcse2 - pupils getting 1-2 A*-C	24%	17%	23%	16%	26%	18%	23%	19%	27%	23%	26%	18%	25%	21%
		2	gcse4 - pupils getting 3-4 A*-C	9%	7%	9%	5%	14%	7%	8%	6%	13%	9%	13%	9%	11%	11%
		2	gcse7 - pupils getting 5-7 A*-C	4%	6%	3%	5%	2%	5%	4%	3%	6%	3%	4%	7%	7%	7%
		2	gcse8 - pupils getting 8+ A*-C	1%	2%	4%	1%	7%	2%	1%	2%	1%	3%	10%	3%	14%	12%
		3c	gcse0 - pupils getting 0 A*-C	51%	62%	57%	66%	48%	59%	52%	59%	37%	49%	35%	56%	29%	48%
		3c	gcse2 - pupils getting 1-2 A*-C	20%	20%	21%	22%	28%	23%	24%	21%	28%	24%	24%	21%	25%	24%
		3c	gcse4 - pupils getting 3-4 A*-C	13%	8%	10%	8%	3%	10%	10%	9%	13%	10%	16%	9%	16%	9%
		3c	gcse7 - pupils getting 5-7 A*-C	6%	4%	5%	3%	8%	5%	6%	4%	8%	7%	8%	6%	7%	8%
		3c	gcse8 - pupils getting 8+ A*-C	11%	6%	8%	1%	14%	3%	8%	7%	15%	10%	16%	8%	23%	11%
		3b	gcse0 - pupils getting 0 A*-C	34%	46%	36%	50%	28%	39%	37%	49%	28%	38%	26%	37%	22%	32%
		3b	gcse2 - pupils getting 1-2 A*-C	20%	23%	20%	24%	15%	26%	20%	24%	16%	26%	16%	25%	17%	23%
		3b	gcse4 - pupils getting 3-4 A*-C	16%	10%	17%	9%	20%	11%	14%	10%	16%	11%	19%	11%	15%	13%
		3b	gcse7 - pupils getting 5-7 A*-C	12%	9%	12%	7%	15%	8%	12%	7%	16%	9%	14%	9%	15%	9%
		3b	gcse8 - pupils getting 8+ A*-C	19%	12%	15%	10%	22%	16%	16%	10%	24%	15%	25%	19%	31%	22%
		3a	gcse0 - pupils getting 0 A*-C	22%	34%	27%	38%	14%	24%	26%	37%	15%	26%	13%	22%	11%	22%
		3a	gcse2 - pupils getting 1-2 A*-C	16%	18%	18%	19%	15%	19%	17%	18%	13%	18%	13%	17%	11%	16%
		3a	gcse4 - pupils getting 3-4 A*-C	15%	15%	14%	13%	15%	16%	15%	13%	15%	15%	14%	17%	12%	15%
		3a	gcse7 - pupils getting 5-7 A*-C	10%	11%	11%	11%	12%	14%	13%	11%	14%	14%	11%	14%	12%	13%
		3a	gcse8 - pupils getting 8+ A*-C	37%	22%	30%	18%	44%	27%	29%	21%	43%	27%	50%	30%	54%	34%
		4c	gcse0 - pupils getting 0 A*-C	14%	22%	19%	29%	10%	19%	14%	24%	7%	15%	7%	14%	5%	11%
		4c	gcse2 - pupils getting 1-2 A*-C	10%	17%	14%	19%	8%	16%	14%	16%	8%	14%	6%	15%	6%	12%
		4c	gcse4 - pupils getting 3-4 A*-C	11%	11%	12%	11%	12%	11%	11%	12%	9%	13%	8%	11%	7%	11%
		4c	gcse7 - pupils getting 5-7 A*-C	11%	11%	11%	12%	11%	14%	12%	12%	12%	12%	10%	12%	9%	12%
		4c	gcse8 - pupils getting 8+ A*-C	54%	38%	45%	30%	60%	40%	49%	36%	63%	46%	69%	49%	73%	54%
		4b	gcse0 - pupils getting 0 A*-C	8%	12%	10%	17%	1%	10%	8%	13%	3%	9%	3%	7%	2%	5%
		4b	gcse2 - pupils getting 1-2 A*-C	7%	10%	13%	13%	11%	11%	9%	11%	7%	9%	5%	7%	3%	6%
		4b	gcse4 - pupils getting 3-4 A*-C	8%	10%	11%	12%	8%	10%	9%	11%	6%	10%	6%	8%	4%	7%
		4b	gcse7 - pupils getting 5-7 A*-C	7%	10%	9%	10%	8%	10%	9%	11%	8%	10%	5%	10%	5%	9%
		4b	gcse8 - pupils getting 8+ A*-C	69%	57%	58%	48%	72%	59%	66%	54%	77%	63%	81%	69%	86%	73%
		4a	gcse0 - pupils getting 0 A*-C	3%	5%	5%	9%	3%	1%	3%	7%	1%	1%	0%	0%	1%	2%
		4a	gcse2 - pupils getting 1-2 A*-C	3%	6%	4%	11%	3%	8%	4%	7%	1%	6%	2%	4%	1%	2%
		4a	gcse4 - pupils getting 3-4 A*-C	3%	6%	2%	6%	3%	1%	4%	6%	3%	4%	2%	4%	1%	3%
		4a	gcse7 - pupils getting 5-7 A*-C	6%	7%	12%	8%	7%	8%	5%	6%	4%	6%	3%	5%	2%	4%
		4a	gcse8 - pupils getting 8+ A*-C	86%	77%	77%	67%	83%	81%	84%	74%	91%	84%	93%	87%	95%	89%
		5c	gcse0 - pupils getting 0 A*-C	0%	0%	4%	4%	0%	4%	1%	0%	3%	1%	0%	0%	0%	0%
		5c	gcse2 - pupils getting 1-2 A*-C	1%	2%	4%	2%	0%	4%	1%	4%	0%	1%	0%	1%	0%	0%
		5c	gcse4 - pupils getting 3-4 A*-C	1%	3%	4%	4%	0%	4%	1%	0%	3%	3%	0%	2%	0%	1%
		5c	gcse7 - pupils getting 5-7 A*-C	2%	3%	4%	2%	0%	4%	1%	2%	3%	3%	0%	2%	0%	1%
		5c	gcse8 - pupils getting 8+ A*-C	95%	92%	83%	88%	100%	86%	95%	93%	90%	91%	98%	94%	99%	97%
		5b	gcse0 - pupils getting 0 A*-C	1%	1%	19%	0%	0%	0%	4%	2%	0%	0%	0%	0%	0%	0%
		5b	gcse2 - pupils getting 1-2 A*-C	0%	1%	0%	0%	0%	0%	4%	2%	0%	4%	0%	0%	0%	0%
		5b	gcse4 - pupils getting 3-4 A*-C	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	1%	1%	0%	0%
		5b	gcse7 - pupils getting 5-7 A*-C	1%	1%	19%	7%	0%	12%	4%	2%	0%	4%	0%	1%	0%	0%
		5b	gcse8 - pupils getting 8+ A*-C	98%	97%	63%	93%	100%	88%	89%	95%	100%	92%	99%	98%	99%	99%
		5a	gcse0 - pupils getting 0 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse2 - pupils getting 1-2 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse4 - pupils getting 3-4 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse7 - pupils getting 5-7 A*-C	0%	0%	0%	0%	0%	0%	50%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse8 - pupils getting 8+ A*-C	100%	100%	0%	0%	0%	0%	50%	100%	100%	0%	100%	100%	100%	100%

Figure 19: Conditional probabilities of KS4 results by KS2 result & pupil characteristics, in London schools in 2014

2015:

KS4 year	KS2 year	Av. KS2	Group	Conditional probabilities (%) Female / FSM only	Male / FSM only	Female / FSM + SEN	Male / FSM + SEN	Female / FSM + SEN + EAL	Male / FSM + SEN + EAL	Female / SEN only	Male / SEN only	Female / SEN + EAL	Male / SEN + EAL	Female / FSM + EAL	Male / FSM + EAL	Female / EAL only	Male / EAL only
2014/15	2009/10	<2	gcse0 - pupils getting 0 A*-C	25%	81%	57%	79%	33%	79%	65%	90%	55%	90%	45%	70%	55%	74%
		<2	gcse2 - pupils getting 1-2 A*-C	25%	10%	21%	11%	33%	21%	16%	5%	15%	10%	25%	15%	17%	5%
		<2	gcse4 - pupils getting 3-4 A*-C	25%	10%	21%	11%	33%	0%	10%	5%	15%	0%	15%	15%	5%	5%
		<2	gcse7 - pupils getting 5-7 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	5%	5%
		<2	gcse8 - pupils getting 8+ A*-C	25%	0%	0%	0%	0%	0%	10%	0%	15%	0%	15%	0%	19%	11%
		2	gcse0 - pupils getting 0 A*-C	64%	66%	66%	70%	59%	62%	63%	71%	56%	65%	56%	56%	40%	50%
		2	gcse2 - pupils getting 1-2 A*-C	18%	20%	16%	18%	19%	23%	20%	19%	23%	21%	21%	25%	22%	23%
		2	gcse4 - pupils getting 3-4 A*-C	11%	5%	9%	5%	10%	6%	9%	6%	9%	8%	14%	6%	12%	9%
		2	gcse7 - pupils getting 5-7 A*-C	6%	4%	4%	3%	6%	5%	4%	0%	6%	4%	9%	6%	9%	6%
		2	gcse8 - pupils getting 8+ A*-C	1%	5%	4%	3%	6%	4%	4%	3%	7%	1%	2%	7%	17%	12%
		3c	gcse0 - pupils getting 0 A*-C	38%	48%	39%	50%	28%	47%	39%	50%	28%	46%	28%	46%	24%	41%
		3c	gcse2 - pupils getting 1-2 A*-C	24%	26%	26%	27%	23%	26%	26%	26%	24%	26%	21%	27%	21%	26%
		3c	gcse4 - pupils getting 3-4 A*-C	17%	10%	16%	10%	19%	12%	14%	10%	19%	9%	21%	13%	18%	11%
		3c	gcse7 - pupils getting 5-7 A*-C	7%	5%	5%	5%	10%	7%	7%	6%	11%	7%	10%	2%	11%	8%
		3c	gcse8 - pupils getting 8+ A*-C	14%	10%	14%	8%	21%	9%	13%	8%	18%	12%	20%	12%	26%	15%
		3b	gcse0 - pupils getting 0 A*-C	35%	44%	40%	47%	33%	38%	32%	42%	24%	33%	27%	35%	19%	30%
		3b	gcse2 - pupils getting 1-2 A*-C	22%	22%	23%	23%	20%	23%	20%	21%	17%	23%	19%	23%	16%	21%
		3b	gcse4 - pupils getting 3-4 A*-C	12%	13%	12%	13%	14%	14%	14%	14%	16%	15%	14%	15%	15%	15%
		3b	gcse7 - pupils getting 5-7 A*-C	9%	7%	8%	6%	8%	8%	9%	7%	10%	10%	8%	9%	9%	10%
		3b	gcse8 - pupils getting 8+ A*-C	22%	14%	17%	12%	25%	16%	24%	16%	34%	20%	33%	19%	42%	25%
		3a	gcse0 - pupils getting 0 A*-C	19%	28%	21%	32%	9%	26%	19%	28%	11%	23%	7%	22%	8%	20%
		3a	gcse2 - pupils getting 1-2 A*-C	17%	21%	18%	22%	17%	22%	16%	21%	13%	20%	16%	20%	12%	17%
		3a	gcse4 - pupils getting 3-4 A*-C	17%	13%	19%	13%	16%	14%	16%	13%	13%	13%	15%	12%	13%	13%
		3a	gcse7 - pupils getting 5-7 A*-C	12%	9%	9%	8%	8%	8%	11%	10%	11%	10%	12%	8%	11%	10%
		3a	gcse8 - pupils getting 8+ A*-C	36%	30%	34%	25%	51%	30%	39%	29%	52%	34%	51%	37%	55%	40%
		4c	gcse0 - pupils getting 0 A*-C	13%	21%	16%	23%	7%	16%	14%	19%	6%	14%	7%	15%	5%	12%
		4c	gcse2 - pupils getting 1-2 A*-C	12%	14%	14%	16%	8%	14%	13%	15%	8%	14%	6%	12%	6%	12%
		4c	gcse4 - pupils getting 3-4 A*-C	12%	13%	13%	13%	13%	12%	10%	13%	10%	13%	10%	12%	9%	12%
		4c	gcse7 - pupils getting 5-7 A*-C	10%	11%	9%	12%	8%	12%	11%	12%	10%	12%	9%	11%	9%	10%
		4c	gcse8 - pupils getting 8+ A*-C	53%	41%	48%	36%	63%	46%	52%	40%	66%	48%	68%	50%	71%	55%
		4b	gcse0 - pupils getting 0 A*-C	6%	10%	12%	14%	4%	7%	8%	11%	3%	6%	2%	5%	2%	4%
		4b	gcse2 - pupils getting 1-2 A*-C	6%	11%	9%	12%	7%	10%	8%	10%	5%	9%	4%	9%	3%	7%
		4b	gcse4 - pupils getting 3-4 A*-C	6%	10%	9%	11%	7%	10%	8%	8%	7%	8%	4%	9%	4%	6%
		4b	gcse7 - pupils getting 5-7 A*-C	8%	7%	9%	8%	8%	7%	8%	9%	7%	9%	7%	6%	5%	7%
		4b	gcse8 - pupils getting 8+ A*-C	73%	63%	62%	56%	74%	66%	68%	61%	78%	69%	83%	71%	86%	76%
		4a	gcse0 - pupils getting 0 A*-C	3%	4%	2%	6%	4%	2%	3%	5%	1%	1%	1%	1%	1%	1%
		4a	gcse2 - pupils getting 1-2 A*-C	3%	4%	9%	7%	4%	5%	4%	6%	3%	3%	1%	3%	1%	2%
		4a	gcse4 - pupils getting 3-4 A*-C	3%	5%	2%	7%	4%	5%	2%	6%	1%	5%	2%	3%	1%	3%
		4a	gcse7 - pupils getting 5-7 A*-C	5%	7%	8%	7%	4%	5%	6%	6%	1%	5%	4%	5%	2%	3%
		4a	gcse8 - pupils getting 8+ A*-C	87%	80%	78%	73%	83%	83%	84%	77%	93%	86%	93%	89%	95%	91%
		5c	gcse0 - pupils getting 0 A*-C	1%	0%	4%	3%	0%	3%	4%	0%	2%	1%	0%	0%	0%	0%
		5c	gcse2 - pupils getting 1-2 A*-C	1%	2%	4%	2%	8%	3%	2%	3%	2%	1%	0%	0%	0%	1%
		5c	gcse4 - pupils getting 3-4 A*-C	2%	3%	7%	5%	8%	8%	1%	3%	2%	4%	1%	2%	0%	1%
		5c	gcse7 - pupils getting 5-7 A*-C	2%	3%	4%	2%	8%	3%	1%	4%	2%	1%	0%	2%	1%	1%
		5c	gcse8 - pupils getting 8+ A*-C	94%	92%	82%	88%	78%	84%	92%	90%	91%	93%	98%	95%	98%	96%
		5b	gcse0 - pupils getting 0 A*-C	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%	0%	0%
		5b	gcse2 - pupils getting 1-2 A*-C	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%	0%	0%
		5b	gcse4 - pupils getting 3-4 A*-C	0%	0%	0%	5%	0%	0%	0%	1%	0%	0%	1%	1%	0%	0%
		5b	gcse7 - pupils getting 5-7 A*-C	1%	2%	11%	5%	0%	10%	3%	1%	0%	3%	1%	1%	0%	0%
		5b	gcse8 - pupils getting 8+ A*-C	98%	97%	89%	90%	100%	90%	97%	96%	100%	97%	97%	99%	99%	100%
		5a	gcse0 - pupils getting 0 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse2 - pupils getting 1-2 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%
		5a	gcse4 - pupils getting 3-4 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse7 - pupils getting 5-7 A*-C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		5a	gcse8 - pupils getting 8+ A*-C	100%	100%	0%	0%	0%	0%	100%	100%	100%	100%	100%	100%	98%	100%

Figure 20: Conditional probabilities of KS4 results by KS2 result & pupil characteristics, in London schools in 2015

EXPECTED KS4 RESULTS, CONDITIONAL ON KS2 + PUPIL CHARACTERISTICS

The table below sets out expected, probability-weighted KS4 results, conditional on KS2 results and certain pupil characteristics. These were calculated using the conditional probabilities set out above, along with the mid-point of each range of GCSE results (e.g. 1-2 GCSEs at A*-C = 1.5 & 2-3 GCSEs at A*-C = 2.5) and rounded to the nearest whole integer.

KS4 year	Av. KS2 result	Expected no. of A*- C GCSEs													
		Female / FSM only	Male / FSM only	Female / FSM + SEN	Male / FSM + SEN	Female / FSM + SEN + EAL	Male / FSM + SEN + EAL	Female / SEN only	Male / SEN only	Female / SEN + EAL	Male / SEN + EAL	Female / FSM + EAL	Male / FSM + EAL	Female / EAL only	Male / EAL only
2011	**<2**	1	1	1	0	1	0	1	1	1	0	1	1	1	1
2011	**2**	1	1	1	1	1	1	1	1	2	1	2	1	2	1
2011	**3c**	2	1	2	1	2	2	2	1	2	2	3	2	3	2
2011	**3b**	3	2	3	2	4	2	3	2	4	3	4	3	5	3
2011	**3a**	4	3	4	3	5	4	4	3	6	4	6	5	6	5
2011	**4c**	6	5	5	4	6	5	5	4	7	5	7	6	7	6
2011	**4b**	7	6	6	5	8	7	6	6	7	7	8	7	8	7
2011	**4a**	8	8	7	6	8	8	8	7	8	8	9	8	9	8
2011	**5c**	9	9	8	7	9	7	9	8	9	8	9	9	9	9
2011	**5b**	9	9	9	9	9	9	8	8	9	8	9	9	9	9
2011	**5a**	9	8	9	0	0	0	9	7	0	9	9	9	9	9
2012	**<2**	1	1	1	1	1	1	1	1	1	1	1	1	2	2
2012	**2**	1	1	1	1	1	1	1	1	1	1	2	2	2	2
2012	**3c**	2	1	2	1	3	2	2	1	3	2	3	2	4	2
2012	**3b**	3	2	3	2	4	2	3	2	4	3	4	3	4	3
2012	**3a**	4	3	4	3	5	4	4	3	5	4	6	4	6	4
2012	**4c**	6	5	5	4	6	5	6	4	7	5	7	6	7	6
2012	**4b**	7	6	6	5	7	6	7	6	8	7	8	7	8	7
2012	**4a**	8	8	8	7	8	8	8	7	8	8	9	8	9	8
2012	**5c**	9	8	8	8	9	8	9	8	9	8	9	9	9	9
2012	**5b**	9	9	9	8	9	7	9	9	9	8	9	9	9	9
2012	**5a**	8	9	0	9	0	0	9	9	9	9	9	9	9	9
2013	**<2**	1	1	0	1	1	1	1	0	1	1	1	1	2	1
2013	**2**	1	1	1	1	2	1	1	1	1	1	2	2	2	2
2013	**3c**	2	2	2	2	3	2	2	1	3	2	3	2	3	1
2013	**3b**	3	2	2	2	4	2	3	2	4	3	4	3	4	3
2013	**3a**	4	3	4	3	5	3	4	3	5	4	5	4	6	4
2013	**4c**	6	5	5	4	7	5	6	4	7	5	7	6	7	6
2013	**4b**	7	6	6	5	8	6	7	6	7	7	8	7	8	7
2013	**4a**	8	8	8	7	8	8	8	7	8	8	9	8	9	8
2013	**5c**	9	9	8	8	8	8	8	9	9	9	9	9	9	9
2013	**5b**	9	9	8	8	9	9	9	9	9	8	9	9	9	9
2013	**5a**	9	9	0	0	0	0	0	9	0	9	9	9	9	9
2014	**<2**	1	1	1	1	1	1	1	0	1	1	2	1	2	1
2014	**2**	1	1	1	1	2	1	1	1	1	1	2	1	2	2
2014	**3c**	2	1	2	1	2	1	2	2	3	2	3	2	3	2
2014	**3b**	3	2	3	2	4	3	3	2	4	3	4	3	4	3
2014	**3a**	5	3	4	3	5	4	4	3	5	4	6	4	6	5
2014	**4c**	6	5	5	4	7	5	6	5	7	6	7	6	7	6
2014	**4b**	7	6	6	6	7	6	7	6	8	7	8	7	8	7
2014	**4a**	8	8	8	7	8	8	8	7	9	8	9	8	9	8
2014	**5c**	9	9	8	8	9	8	9	9	8	9	9	9	9	9
2014	**5b**	9	9	7	9	9	9	8	9	9	9	9	9	9	9
2014	**5a**	9	9	0	0	0	0	8	9	9	0	9	9	9	9
2015	**<2**	4	0	1	1	2	0	1	0	2	0	2	1	2	2
2015	**2**	1	1	1	1	2	1	1	1	2	1	1	2	3	2
2015	**3c**	3	2	3	2	3	2	3	2	3	2	3	2	4	3
2015	**3b**	3	2	3	2	4	3	4	3	4	3	4	3	5	4
2015	**3a**	5	4	4	4	6	4	5	4	6	4	6	5	6	5
2015	**4c**	6	5	6	5	7	5	6	5	7	6	7	6	7	6
2015	**4b**	7	7	7	6	7	7	7	7	8	7	8	7	8	8
2015	**4a**	8	8	8	7	8	8	8	8	9	8	9	8	9	8
2015	**5c**	9	9	8	8	8	8	8	8	8	9	9	9	9	9
2015	**5b**	9	9	9	9	9	9	9	9	9	9	9	9	9	9
2015	**5a**	9	9	0	0	0	0	9	9	9	9	9	9	9	9

Figure 21: Expected KS4 results conditional on KS2 result, pupil characteristics and year, in London schools

For the calculations shown in this report, based on the pilot KS2 impact analysis, expected values had to be determined for the period 2011-2015 as a whole. This was because the KS2 model we had constructed would produce unreliable results if we ran it using data for one specific year, due to the fact that the amount of data for every combination of year and set of pupil characteristics would be relatively small. Therefore, it seemed more appropriate to use the model to produce average KS2 results for pupils with each of the sets of characteristics shown below but across the whole period, even though this could slightly reduce accuracy.

In addition, given that some of the pupils in our sample were not old enough to have sat their GCSEs yet (i.e. those that did KS2 in 2010/11 – 2013/14), effectively, we have had to proxy their performance at KS4 using the average performance of their predecessors who did KS2 in 2005/06 – 2009/10 (i.e. those who subsequently completed KS4 in 2010/11 – 2014/15). However, given the relative stability of the annual expected KS4 results over the period, this does not seem like an unreasonable assumption.

The average expected KS4 results for the period 2011-15 are set out below.

		Expected no. of A*- C GCSEs													
KS4 year	Av. KS2 result	Female / FSM only	Male / FSM only	Female / FSM + SEN	Male / FSM + SEN	Female / FSM + SEN + EAL	Male / FSM + SEN + EAL	Female / SEN only	Male / SEN only	Female / SEN + EAL	Male / SEN + EAL	Female / FSM + EAL	Male / FSM + EAL	Female / EAL only	Male / EAL only
2011-15	**<2**	1	1	1	1	1	1	1	0	1	1	1	1	2	1
2011-15	**2**	1	1	1	1	1	1	1	1	1	1	2	1	3	2
2011-15	**3c**	2	1	2	1	3	2	2	1	3	2	3	2	4	2
2011-15	**3b**	3	2	3	2	4	3	3	2	4	3	4	3	5	3
2011-15	**3a**	4	3	4	3	5	4	4	3	5	4	6	4	6	5
2011-15	**4c**	6	5	5	4	7	5	6	5	7	5	7	6	7	6
2011-15	**4b**	7	6	6	6	7	7	7	6	8	7	8	7	8	7
2011-15	**4a**	8	8	8	7	8	8	8	7	8	8	9	8	9	8
2011-15	**5c**	9	9	8	8	9	8	9	8	9	9	9	9	9	9
2011-15	**5b**	9	9	8	8	9	8	9	9	9	9	9	9	9	9
2011-15	**5a**	9	9	9	9	0	0	8	8	9	9	9	9	9	9

Figure 22: Expected KS4 results conditional on KS2 result & pupil characteristics, in London schools, 2011-15

In order to be able to link the predicted KS2 results generated from the econometric model (which were based on KS2 *levels*) to the above expected KS4 results (which are conditional on KS2 *sub-levels*), some conversion was necessary. The predicted KS2 level results were not whole integers, so they were converted to the closest sub-level, using the following table:

Av. KS2 level (predicted)		KS2 sub-level
Min.	Max.	
0	1.99	<2
2	2.99	2
3	3.32	3c
3.33	3.65	3b
3.66	3.99	3a
4	4.32	4c
4.33	4.65	4b
4.66	4.99	4a
5	5.32	5c
5.33	5.65	5b
5.66	5.99	5a

Table 9: Conversion of average predicted KS2 levels to KS2 sub-levels

APPENDIX 6

LINKAGES FROM KS4 TO EARNINGS

The table below sets out the relevant results from the paper published by DfE which provides linkages between GCSE results and lifetime earnings. Amongst other things, the paper attempted to estimate the relative lifetime productivity gains from different numbers of "good" GCSEs (i.e. at grade C or above). As a result of some sensitivity analysis, as well as central estimates of the marginal returns, the paper also proposed both alternative low and high estimates.

In order to be able to make use of these results, the ranges of numbers of "good" GCSEs were converted into mid-points – i.e. 1-2 GCSEs = 1.5 or 5-7 GCSEs = 6 and therefore it was easier to calculate the number of additional GCSEs being considered. In addition, using the number of additional GCSEs for each range, the returns were restated per additional GCSE. Therefore, the return to 3-4 good GCSEs relative to 1-2 good GCSEs was interpreted as the return to 3.5 relative to 1.5 good GCSEs (or an additional 2) and therefore the marginal return was divided by 2 to determine the marginal return per additional GCSE when moving from 1 to 3 or 2 to 4 GCSEs at grades A*-C. For clarity, the marginal returns reflect the present value of the additional earnings over an individual's whole lifetime as a result of the additional "good" GCSEs.

GCSE results	No. of GCSEs (mid-point)	No. of additional GCSEs (mid-point)	Marginal productivity returns (central estimate) (£)				Marginal productivity returns (low estimate) (£)				Marginal productivity returns (high estimate) (£)			
			Male		Female		Male		Female		Male		Female	
			Return vs. prior result	Return per additional GCSE	Return vs. prior result	Return per additional GCSE	Return vs. prior result	Return per additional GCSE	Return vs. prior result	Return per additional GCSE	Return vs. prior result	Return per additional GCSE	Return vs. prior result	Return per additional GCSE
0 GCSEs at A*-C	0	0												
1-2 GCSEs at A*-C	1.5	1.5	170,984	113,989	110,395	73,597	125,029	83,353	89,542	59,695	403,147	268,765	280,995	187,330
3-4 GCSEs at A*-C	3.5	2	59,043	29,522	51,055	25,528	42,053	21,027	42,436	21,218	141,726	70,863	118,221	59,111
5-7 GCSEs at A*-C	6	2.5	72,999	29,200	55,445	22,178	50,989	20,396	42,080	16,832	186,321	74,528	136,850	54,740
8+ GCSEs at A*-C	9	3	2,909	970	32,592	10,864	- 120	- 40	23,493	7,831	- 24,518	- 8,173	71,799	23,933

Figure 23: Marginal productivity returns to ranges of "good" GCSEs

Source: Hayward H., Hunt E., Lord A. (2014), 'The economic value of key intermediate qualifications: estimating the returns and lifetime productivity gains to GCSEs, A levels and apprenticeships', Department for education