Department for Education

The link between attendance and attainment in an assessment year

Research Report

March 2025

Contents

List of figures 2
List of tables 4
1. Introduction 5
2. Key findings 6
3. Methods 8
3.1 Source of data 8
3.2 Cohort 9
3.3 Outcome indicators 10
3.4 Controlling for other known pupil characteristics 11
4. Results 17
4.1 Introduction 17
4.2 The link between absence and attainment at KS2 18
4.3 The link between absence and attainment at KS4 23
5. Conclusions 30
Annex A: Technical notes 32
A.1 Data sources and manipulation 32
A.2 Data coverage and quality 32
A.3 Methodology 33
Annex B: Model output 35
B.1 Pupils achieving expected level in RWM at the end of KS2 35
B.1 Pupils achieving level 5 or above in Maths and English GCSE at the end of KS4 36
B.3 Strength of the models 38
Annex C: Charts with approximate days missed 40

List of figures

Figure 1 Proportion of pupils in each 5% attendance band for Year 6 achieving expected standard in RWM at the end of KS2 18

Figure 2 The relationship between FSM eligibility and attendance for KS2 pupils 19

Figure 3 The relationship between attendance and successful outcome for KS2 pupils by FSM eligibility 20

Figure 4 Proportion of pupils in the KS2 cohort who achieved the expected outcome in RWM for compounding factors 21

Figure 5 Heatmap showing the odds ratios between pairs of possible attendance bands during Year 6 after other pupil characteristics and the school attended is controlled for. 23

Figure 6 Proportion of pupils in each 5% attendance band for Year 11 achieving Grade 5 or above in English and Maths GCSE at the end of KS4 24

Figure 7 The relationship between FSM eligibility and attendance in Year 11 for KS4 pupils 25

Figure 8 The relationship between attendance and successful outcome for KS4 pupils by FSM eligibility 26

Figure 9 Proportion of pupils in the KS4 cohort who achieved Grade 5 in Maths and English GCSE for compounding factors 27

Figure 10 Heatmap showing the odds ratio pairs between pairs of possible attendance bands during Year 11 after other pupil characteristics and the school attended is controlled for. 29

Figure 11 The odds ratio of achieving the expected standard at KS2 and KS4 when moving up a single 5% attendance band 30

Figure 12 Proportion of pupils achieving expected standard in RWM at the end of KS2 by the approximate number of days missed in Year 6 40

Figure 13 The relationship between FSM eligibility and days missed in Year 6 for KS2 pupils 41

Figure 14 The relationship between days missed in Year 6 and successful outcome for KS2 pupils by FSM eligibility 41

Figure 15 Proportion of pupils achieving Grade 5 or above in English and Maths GCSE at the end of KS4 by the approximate number of days missed in Year 11 42

Figure 16 The relationship between FSM eligibility and days missed in Year 11 for KS4 pupils 44

Figure 17 The relationship between days missed in Year 11 and successful outcome for KS4 pupils by FSM eligibility....45

Figure 18 The odds ratio of achieving the expected standard at KS2 and KS4 when moving up a single 5% attendance band 46

List of tables

Table 1 Range of possible days and weeks of school missed in selected attendance bands over a full school year 9

Table 2 Number of end KS2 and end KS4 pupils in state-funded mainstream schools in 2022/23 analysis after attainment, characteristic and possible session filters 10

Table 3 Control variables used in the regression model which may affect attainment.....12

Table 4 The number and proportion of pupils in each attendance band in the assessment year in this 2022/23 cohort.....17

Table 5 KS2 regression model output.....35

Table 6 KS4 regression model output.....37

Table 7 Model accuracy.....39

1. Introduction

This report investigates the link between different 5 percentage point bandings of attendance and attainment for pupils at the end of Key Stage 2 (KS2) and Key Stage 4 (KS4) in state-funded mainstream schools in 2022/23.

Absence from school is not the only factor that is likely to affect a pupil's level of attainment. There are a range of pupil, school, parental and societal characteristics that are likely to affect attainment in varying degrees. To better understand the link between absence and attainment, a model was developed to take confounding factors into account, such as prior attainment and different pupil characteristics.

This report builds on figures from previous research that were published in March 2016, The link between absence and attainment at KS2 and KS4: 2013/14 academic year. That report looked at overall, authorised and unauthorised absence and controlled for other factors to look at the relative impact of absence on attainment. A more recent publication The link between absence and attainment at KS2 and KS4: 2018/19 presented the percentage of pupils who achieved the expected standard at each Key Stage in 2018/19 but did not attempt to control for other confounding factors to try and quantify the contribution of attendance.

This report investigates how much more likely a pupil in each 5% band of attendance is to reach a successful outcome for the Key Stage compared with pupils in other attendance bands in the 2022/23 academic year. A regression model is used to control and remove the effect of factors which may affect a pupil's attainment allowing the effect of the pupil's attendance to be isolated.

2. Key findings

The key findings in this report are aligned with the key findings of previous reports into this topic showing a positive correlation between increased attendance and attainment. By controlling for other factors which may affect a pupil's attainment, including prior attainment, eligibility for Free School Meals (FSM), Special Educational Needs (SEN) status, sex, month of birth, first language, and IDACI score of home postcode, the relationship between attendance during the year of assessment and the likelihood of reaching the expected outcome for the stage was found.

- At both KS2 and KS4, the higher the 5% attendance band a pupil is in during the assessment year, the more likely they were to achieve a successful outcome in 2022/23.
- At KS2, pupils who attended school nearly every day in Year 6 (with an attendance rate of 95-100%) were 1.3 times more likely to achieve the expected standard in reading, writing and maths compared to pupils who only attended 90-95% of the time. This means missing just 10 days of year 6 reduced the likelihood of reaching the expected standard by around 25%. Pupils who attended nearly every day were 1.8 times more likely to achieve the standard than persistently absent pupils who only attend 85-90% of the time (relating to 4 -6 weeks more time in school).
- At KS2, pupils who attend between 50 and 55% of the time are twice as likely to achieve the expected standard than students who are severely absent (attending less than 50% of sessions). A pupil with 60-65% attendance (equating to approximately 4-6 extra weeks in school) has a four times greater chance of successful outcome than those who attend <50% of the time.
- At KS4, pupils who attended school nearly every day in Year 11 (with an attendance rate of 95-100%) were 1.9 times more likely to achieve the Grade 5 in English and Maths GCSE compared to pupils who only attended 90-95% of the time (relating to up to 2 weeks more time in school over the course of a year) and 3 times more likely than persistently absent pupils who only attend 85-90% of the time (relating to 4 - 6 weeks more time in school). This means missing just 10 days of year 11 reduced the likelihood of achieving grade 5 in English and Maths by around 50%. Pupils who attended school 90-95% of them time were also 1.6 times more likely to achieve Grade 5 in Maths and English GCSE compared to pupils who only attended 85-90% of them time, relating to an additional 2-4 weeks in school. This demonstrates there are improvements in the chance of success from increased attendance by one absence band.
- At lower attendance levels, KS4 pupils who only attend between 50 and 55% of the time are 1.6 times more likely to achieve 9-5 in Maths and English GCSE than students who are severely absent and attend less than 50% of sessions. Increased

attendance from severely absent to 60-65% (equating to approximately an extra 4-6 weeks in school) is associated with double the likelihood of the expected outcome compared with pupils who attend <50% of the time.

- At both KS2 and KS4 moving up a single attendance band by attending school 2 extra weeks a year is associated with an increased chance of achieving the expected outcome of at least 30% (KS2) and 10% (KS4) across the full range of attendance. At KS4, moving from 90-95% to 95-100% was associated with largest increase in chance from moving up one band.
- At school level, as more pupils are in the higher attendance bands, the greatest gains in outcome could be made by pupils from the 90-95% band moving into the 95-100% attendance band (0-20 days extra in school over the year). However, this analysis shows that there may be gains in attainment by improving attendance of all pupils by a single 5% band across the range of attendance bands.

3. Methods

Methods used within this report are summarised below. Technical notes providing further definitions and detail are available in Annex A.

3.1 Source of data

Parents of children of compulsory school age (aged between 5 and 15 at the start of the academic year) are, by law, required to ensure that their children receive a suitable education through regular attendance at school or otherwise. The Department for Education collects each enrolled pupil's overall sessions missed and number of possible sessions available to them, where a session is equivalent to half a day of school[1].

The pupil absence and attainment data used in this release have been created by matching absence data, as collected via the school census, to KS2 and KS4 attainment results collected from awarding bodies.

The absence data collected via the school census is as reported by the individual schools and therefore the data provider (i.e. the individual schools) are responsible for ensuring the accuracy of the data returned. At the point of submission to the Department, the data collected is deemed to have been checked and signed off by the school head teacher. The data collected from local authority maintained schools also undergoes further approval by the responsible local authority before use by the Department.

Absence data is then matched into the National Pupil Database (NPD) – a longitudinal database which holds a wide range of pupil level data for schools across England. Note that as with any matching exercise there is an accepted margin of error and some mismatches may occur. Further information on the data matching involved is available in Annex A.

[1] Pupil absence information is collected at pupil enrolment level rather than pupil level. Where a pupil has moved school during the year, they will be counted more than once as they have recorded absence data in more than one school. Where a pupil has a dual registration their absence may be returned from both schools, if both schools return absence data via the school census.

Absence Bands

A school records if a pupil attends each half day session at school. Over the school year the number of sessions a pupil attends is calculated as a percentage of the possible sessions they should have attended at the school. For this report these attendance percentages have been grouped into 5% bandings. The table below shows the range of days and weeks missed by a pupil in selected bandings if the pupil had been registered for a full school year of 380 sessions.

Table 1 Range of possible days and weeks of school missed in selected attendance bands over a full school year

Attendance Band	Range of possible days missed	Range of possible weeks missed
95-100%	0 - 9.5	0 – 1.9
90-95%	9.5 - 19	1.9 - 3.8
85-90%	19 - 28.5	3.8 - 5.7
80-85%	28.5 - 38	5.7 - 7.6 (approx. half a term)
60-65%	66.5 - 76	13.3 - 15.2
50-55%	85.5- 95	17.1 - 19
<50%	>95	>19

3.2 Cohort

The analysis looks at the attainment of pupils at the end of KS2 and KS4 and the attendance each pupil accrued across the year of their assessment – Year 6 for KS2 and Year 11 for KS4. Attendance information is collected at enrolment level. The analysis was restricted to attendance during the assessment year due to the disruption to pupil's attendance records during the Covid-19 pandemic.

For the purposes of this analysis the cohort used in the analysis has been restricted to only include pupils with valid absence and attainment data for a single enrolment during the assessment year and for the previous assessment year (Key Stage 1 (KS1) assessment outcome and Year 2 attendance for the KS2 cohort and KS2 assessment outcome and Year 6 attendance for the KS4 cohort.) The pupils included are only those who had more than 100 possible sessions in the assessment year which relates to approximately one term of absence data.

Restricting the cohort in this way ensured that pupils with extremely low numbers of possible sessions (for whom we would not be able to determine a full picture of the link between their absence and attainment) were removed and did not distort the findings. Pupils with low number of sessions would include pupils who moved to specialist provision or left state-funded education. Details of the numbers of pupils removed prior to analysis are set out in the table below. Further details are in Annex A.

Table 2 Number of end KS2 and end KS4 pupils in state-funded mainstream schools in 2022/23 analysis after attainment, characteristic and possible session filters

	KS2 cohort Number of pupils	**KS4 cohort** Number of pupils
All pupils with attendance data	599,989	541,687
Pupils remaining after applying characteristic filter[2]	598,418	538,697
Pupils remaining after applying attainment and prior attainment filter[3]	596,295	538,697
Pupils remaining after applying possible session filter and prior possible session filter[4]	530,665	478,495

3.3 Outcome indicators

3.3.1 Key Stage 2 cohort

KS2 national curriculum assessments are taken by pupils at the end of primary school education, with this Key Stage being taught over four years (from Year 3 to Year 6). In the tests, pupils meet the expected standard in a subject if they achieve a scaled score of 100 or more. The subjects assessed are reading, maths, writing, grammar, punctuation

[2] Filtered to include pupils at the end of KS2/KS4 who had recorded status for Special educational need (SEN), Free school meals (FSM), language spoken, Income deprivation affecting children index decile (IDACI)

[3] Filtered to include only pupils who had recorded outcome for Reading Writing and Maths at expected standard (RWM) for KS2 cohort and Maths and English GCSE for KS4 cohort. Prior attainment filter to include only pupils who had recorded outcome for RWM at KS1 for KS2 cohort and RWM at KS2 for KS4 cohort.

[4] Filtered to include pupils who had more than 100 possible sessions in the assessment year and the previous assessment year. More detail on the methodology used to remove these pupils and some likely reasons as to why pupils may not have a full Key Stage of reported possible sessions, are included at A.3 Methodology.

and spelling, science[5]. For this research we used a combined outcome measure of a pupil reaching the expected standard in reading, writing and maths tests (RWM). Pupils were categorised as having achieved the expected level if and only if were judged to be 'Working at the expected standard' or 'Working at greater depth' in all three of these categories. Pupils with missing results were removed from the analysis.

3.3.2 KS4 cohort

KS4 exams are taken by pupils at the end of secondary school education, with this Key Stage usually being taught over two years (from Year 10 to Year 11). For this research we used a combined successful outcome measure of a pupil achieving Grade 5 in both Maths and English GCSE. Pupils were categorised as having achieved the expected level if and only if they achieved the Grade 5 or above in both subjects. Pupils with missing results were removed from the analysis[6].

3.4 Controlling for other known pupil characteristics

In this report a logistic regression model predicts the probability of achieving KS2 and KS4 outcome indicators for each pupil based on the 5% banding of attendance during the year of assessment from 50% to 100% attendance. Attendance of less than 50% is grouped into a single band due to low pupil numbers.

There is strong evidence that other factors may affect a pupil's attainment such FSM status and SEN provision. The relationship between these factors interact with both attendance and outcome with pupils eligible for FSM over represented in lower attendance bands and having a lower likelihood of achieving the expected outcome. The distribution of pupils eligible for FSM within attendance bands for this cohort are shown in the results in Figure 2. To isolate the relationship between attendance and attainment, these compounding factors were controlled for to find the how much more likely a pupil is to achieve the expected standard based on their attendance band if confounding factors affecting them were equal. These factors are detailed in Table 3. Reference levels were chosen to be the level with the lowest proportion of pupils achieving the expected outcome at the end of the Key Stage.

[5] Further information on KS2 attainment data can be found in the 'Attainment in primary schools in England' quality and methodology information section.
[6] Further information on KS4 attainment data can be found in the 'Key Stage 4 Performance' methodology information section.

Table 3 Control variables used in the regression model which may affect attainment

Variable	Description
First language	Language is a binary indicator variable to indicate that a child's "First language" is English/British or not English[7]. Pupils with all other codings were removed. • 1 Pupils with LanguageGroupMinor in the school census of OTH (Other) or OTB (Other than English or British) • 0 ENB (English/British) or ENG (English) (reference)
Sex	Sex is included in the model as a categorical variable separating – • Girls • Boys (reference)
Free School Meals Indicator (FSM)	Free school meals is a binary indicator variable that states whether a pupil's family have claimed eligibility for free school meals as reported at the time of the annual spring school census. Parents are able to claim free school meals if they receive a qualifying benefit. The FSM variable does not relate to pupils who actually received free school meals but those who are eligible to receive free school meals. Pupils with all other codings (including Not Known) were removed. • 1 – Eligible to receive FSM (reference) • 0 – Not eligible to receive FSM
Special Educational Need (SEN)	SEN status is derived from the school census and is included in the model as a categorical variable separating – • No identified SEN • SEN Support • Education Health Care Plan (EHCP) (reference) Pupils with all other codings (including Not Known) were removed.
Prior attainment	Prior attainment is included in the model as binary variable indicating if the pupil achieved the expected outcome in the previous Key Stage. • 1 – achieved expected level in RWM in Year 2 (KS2 model) or Year 6 (KS4 Model) • 0 – did not achieve expected level (reference)

[7] The first language is the language to which a child was initially exposed during early development and continues to be exposed to this language in the home or in the community. It does not mean that pupils are necessarily fluent in a language other than English or cannot speak English.

Prior persistent absence (Previous PA)	Prior persistent absence is included in the model as binary variable indicating if the pupil was persistently absent (less than 90% attendance) during the final year of the previous Key Stage to indicate persistently high absence throughout a pupil's schooling which may affect attainment even if attendance in the assessment year is high. • 1 – was persistently absent in Year 2 (KS2 model) or Year 6 (KS4 Model) (reference) • 0 – was not persistently absent
IDACI decile home[8]	Index of Deprivation affecting Children decile of the pupil's home address was included in the variable as an ordered categorical variable • 1-10 relating to the most deprived (1) to least deprived (10) decile area. (Reference value is 1)
Month of birth	The month of a pupil's birth (1 – Jan, 2 – Feb etc to 12 – December) was included in the model as an ordered categorical variable in the order 8,7,6,5,4,3,2,1,12,11,10,9 with reference to 8 (August).
School URN	The unique reference number (URN) of the school attended by the pupil was included in the model as a categorial variable in the second level of the model to account for grouping of outcomes in schools due to school policies, practices and cohorts within individual schools.

The logistic regression models the logarithm of odds of achieving the KS2/KS4 binary outcome indicators (Y) as a generalized linear mixed-effects model (GLMM) with a binomial family including both fixed effects (control variables) and random effects (school URN).

The mathematical form of the model can be written as

$$
\begin{aligned}
logit\,(P(YOutcome_{ij} = 1) \\
&= \beta_0 + \beta_{1-10}(Attendance\ band\ 1-10)_{ij} \\
&+ \beta_{11}(FSM)_{ij} + \beta_{12}(SEN\ Support)_{ij} \\
&+ \beta_{13}(EHCP)_{ij} + \beta_{14}(Sex)_{ij} \\
&+ \beta_{15}(First\ Language\ )_{ij} + \beta_{16}(Prior\ Attainment)_{ij} \\
&+ \beta_{17}(Prior\ PA)_{ij} + \beta_{18-28}(Month\ of\ Birth\ 1-12)_{ij} \\
&+ \beta_{29-37}(IDACI\ home\ 1-10)_{ij} + u_j)
\end{aligned}
$$

[8] IDACI is provided by the Department for Communities and Local Government (DCLG). The index is based on Lower-Layer Super Output Areas in England defined by 2011 census data. The Index of Multiple Deprivation ranks every small area in England from 1 (most deprived area) to 32,844 (least deprived area). 2011 data is more recently available and may not reflect changes in population and demographics within an area since 2011.

where $logit\ (\ P\ (\ YOutcome_{ij} = 1\ )\ )$ is the log odds of the outcome being 1 for the , β_0 is the overall intercept, β_1 to β_{10} are the coefficients for the attendance bands, β_{11} to β_{37} are the coefficients of each control variable and $u_j$is the random intercept for each school.

The model was run using the GLMER package in R using a Bobyqa optimisation algorithm[9] with 2x10[6] iterations. The regression was 'trained' using a random sample containing 70% of the pupil cohort, with the remaining 30% of pupil records retained for assessing model performance.

The models were assessed by calculating the level of discrimination between the pupils that achieved the KS2 and KS4 outcome and the pupils that did not. The discrimination was evaluated by calculating the Area Under Curve (AUC) for the Receiver Operator Characteristic (ROC) curve, where an area of 1 represents a perfect fit and an area of above 0.8 is a good fit[10]. The details of the model fit are shown in Annex B. The KS4 model was rerun using a second random seed to assess the sensitivity of the results to the random sampling of the cohorts.

The odds of achieving the expected outcome at the end the Key Stage after controlling for the other factors can be found by taking an exponential of the coefficients calculated in the model:

$$Odds\ of\ Acheiving\ Outcome\ Indicator = e^{(\beta_0+\beta_1+\cdots\ldots+\beta_{37}+u_j)}$$

In this report the model results have been used to quantify the odds ratio of achieving the KS2/KS4 outcome measure for possible attendance bandings. The odds ratio (OR) of an attendance band is calculated by exponentiating the coefficient:

$$Odds\ Ratio\ (OR) = e^{\beta_i}\ ,$$

where β_i indicates the coefficient of the attendance band of interest with respect to the reference group.

[9] Powell, M.J. The BOBYQA algorithm for bound constrained optimization without derivatives, DAMTP 2009/NA06, NA2009_06.pdf (cam.ac.uk)
[10] Hosmer, D.W. & Lemeshow, S. (2000) Applied logistic regression, 2nd edn. John Wiley & Sons, New York, p162

Odds ratios explained

Odds ratios measure how likely one group is to achieve the expected outcome compared to another (reference) group. They are used to approximate how many more times pupils in one group are likely to achieve than those in another group. For example, if group A has an odds ratio of achieving the expected standard in reading writing and maths of 2, this means members of group A have approximately twice the likelihood of reaching the expected standard compared to members of the reference group. Similarly, if group B has an odds ratio of 0.5, this means members of group B are approximately only half as likely reach the expected standard compared to members of the reference group.

If an odds ratio is:

- **Greater than 1**: Members of the group are more likely to achieve the expected outcome than the reference group,
- **Less than 1:** Members of the group are less likely to achieve the expected outcome than the reference group,
- **Equal to 1:** Members of the group are equally as likely to achieve the expected outcome as the reference group.

During the analysis, the initial odds ratio were calculated with respect to the reference group of pupils who attend school nearly every day and are in the 95-100% attendance band by dividing the odds of achieving the expected outcome for pupils within each attendance band by the odds of achieving the expected outcome for pupils in the 95-100% attendance band[11]. Subsequently the odds ratio between each possible pairing of attendance band was calculated with respect to each other[12].

It is not possible to draw conclusions on the ranking of importance of control factors by comparing the odds ratios for one control factor with another as each is only tested against its own reference group[13]. The model was not designed and optimised to

[11] For example, at Key Stage 4, 56% of pupils in the 95-100% attendance band reached the expected standard and 44% of pupils in this band did not. The odds of reaching the expected standard for pupils in this attendance band would therefore be 56% ÷ 44% = 1.27. Similarly, the odds of reaching the expected standard for pupils in the 90-95% attendance band were 0.80. The odds ratio for 95-100% band pupils compared to 90-95% band pupils would be 1.27 ÷ 0.80 = 1.6. This means that pupils attending 95-100% of the time are 1.6x more likely to reach the expected standard than pupils attending 90-95% of the time.

[12] To rewrite the statement in terms of a percentage reduction in likelihood: For odds ratios above 1, the percentage increase is (OR−1)×100. For odds ratios below 1 (e.g., reversing the comparison), the percentage decrease is (1−OR)×100).

[13] For example, pupils without factor A may have a 40% chance of success and 80% chance of success with factor A. Pupils without factor B may with a 5% chance of success and a 10% chance of success with factor B. Both these factors will have an odds ratio of 2, but factor A is associated with an increase of 40 percentage points in outcome while factor B is only associated with a 5 percentage point increase in outcome so the factors cannot be said to have the same importance on the outcome of the pupil.

investigate the importance of these factors only to allow them to be controlled to isolate the effect of attendance on the expected outcome.

4. Results

4.1 Introduction

In this section we present the likelihood of pupils in each 5% attendance band reaching the expected outcome indicator for their Key Stage while controlling for other factors that are likely to have an influence on a pupil's level of attainment.

Table 4 shows the number of pupils in each 5% attendance band in Year 6 and Year 11. The majority of pupils in KS2 and over 40% of the pupils in KS4 attended 95-100% of possible sessions. More than 84% of KS2 and 70% of KS4 pupils attended more than 90% of possible sessions.

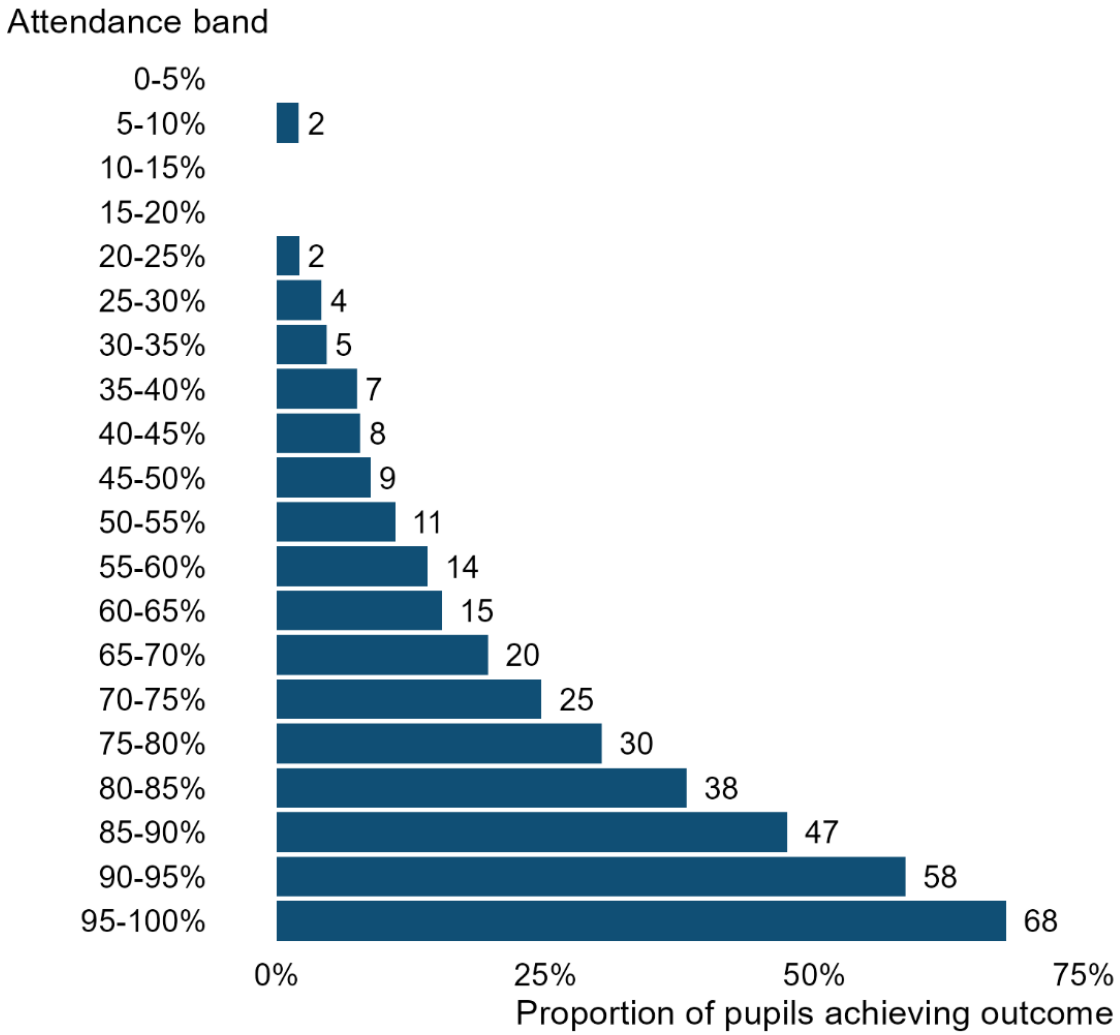
Table 4 The number and proportion of pupils in each attendance band in the assessment year in this 2022/23 cohort

Attendance Band	KS2 Cohort Number	KS2 Cohort % of total	KS4 Cohort Number	KS4 Cohort % of total
0-5%	167	0.03%	2388	0.44%
5-10%	96	0.02%	1279	0.24%
10-15%	78	0.01%	1232	0.23%
15-20%	119	0.02%	1406	0.26%
20-25%	142	0.03%	1533	0.29%
25-30%	168	0.03%	1728	0.32%
30-35%	194	0.04%	1868	0.35%
35-40%	228	0.04%	2154	0.40%
40-45%	296	0.06%	2442	0.45%
45-50%	367	0.07%	2947	0.55%
50-55%	516	0.10%	3299	0.61%
55-60%	685	0.13%	3966	0.74%
60-65%	952	0.18%	5084	0.95%
65-70%	1717	0.32%	6711	1.25%
70-75%	3218	0.61%	9748	1.82%
75-80%	6721	1.27%	15262	2.84%
80-85%	16872	3.18%	28516	5.31%
85-90%	50639	9.54%	64323	11.98%
90-95%	159012	29.96%	146367	27.25%
95-100%	288478	54.36%	234813	43.72%

4.2 The link between absence and attainment at KS2

For pupils at the end of KS2 in 2022/23, as the 5% band of attendance during Year 6 increases the proportion of pupils achieving the expected standard RWM increases as shown in Figure 1[14]. Over the entire cohort, the likelihood (without controlling for confounding factors) of pupils with 95-100% attendance reaching the expected standard is 1.5 times greater than pupils with 90-95% attendance and 2.4 times greater than pupils with 85-90% attendance. When low pupil numbers in the low attendance bandings are aggregated into a single <50% attendance band the relationship becomes linear. Due to low numbers of pupils in the low attendance bandings they are aggregated into a single <50% attendance band for subsequent analysis

Figure 1 Proportion of pupils in each 5% attendance band for Year 6 achieving expected standard in RWM at the end of KS2



[14] See Annex C for charts recreated with approximate number of days missed in place of attendance band.

Pupil characteristics are not evenly distributed across the attendance bands with pupils eligible for free school meals under-represented in the higher attendance bands (accounting for only 20% of pupils in the 95-100% attendance band) and over-represented in the lower attendance bands (69% in the 65-70% band) compared with the average of the entire KS2 cohort of 27.3% (Figure 2). Across all attendance bands, pupils eligible for FSM have a lower likelihood of achieving the expected level in reading, writing and maths (Figure 3)[14].

Figure 2 The relationship between FSM eligibility and attendance for KS2 pupils

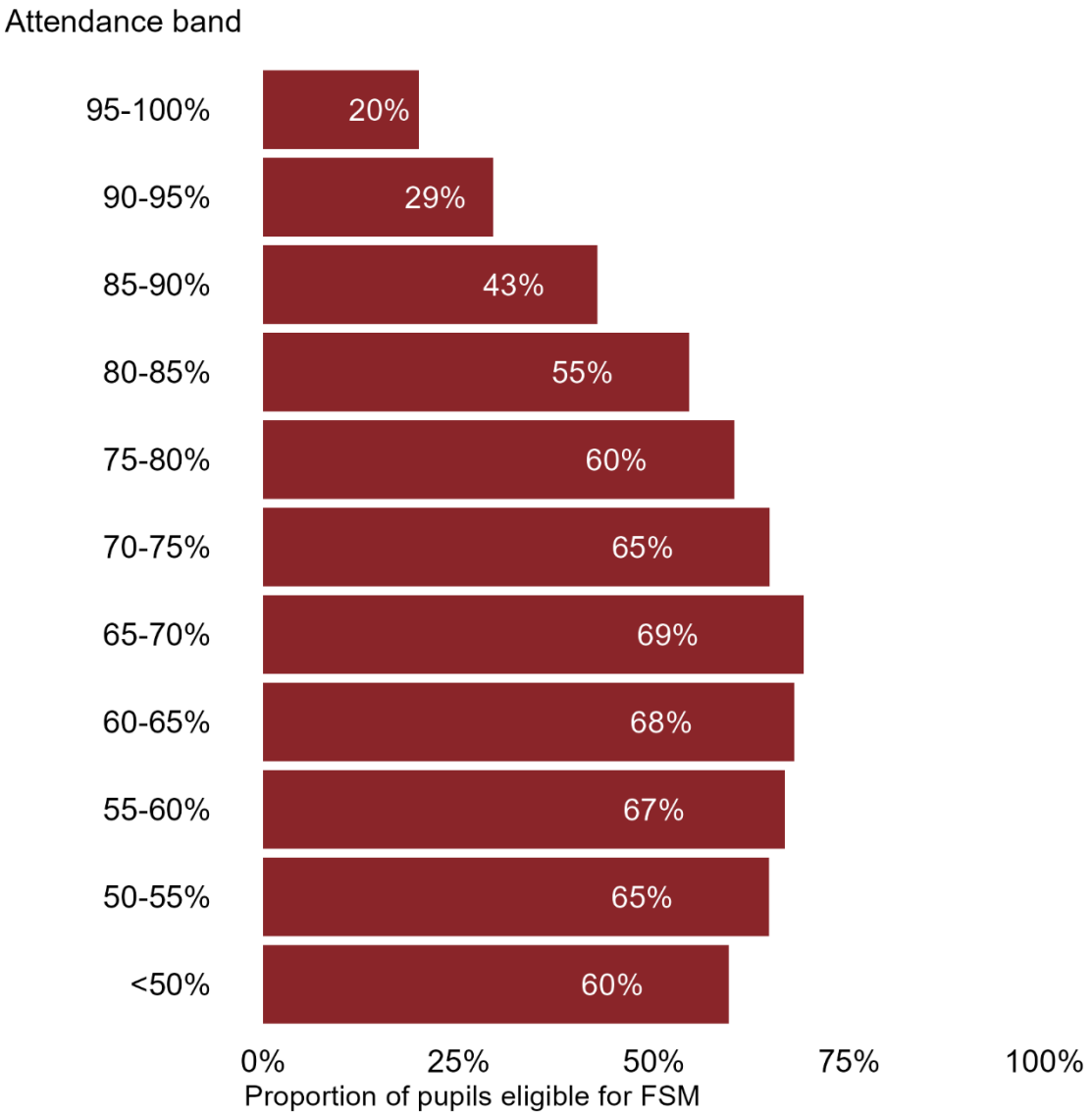
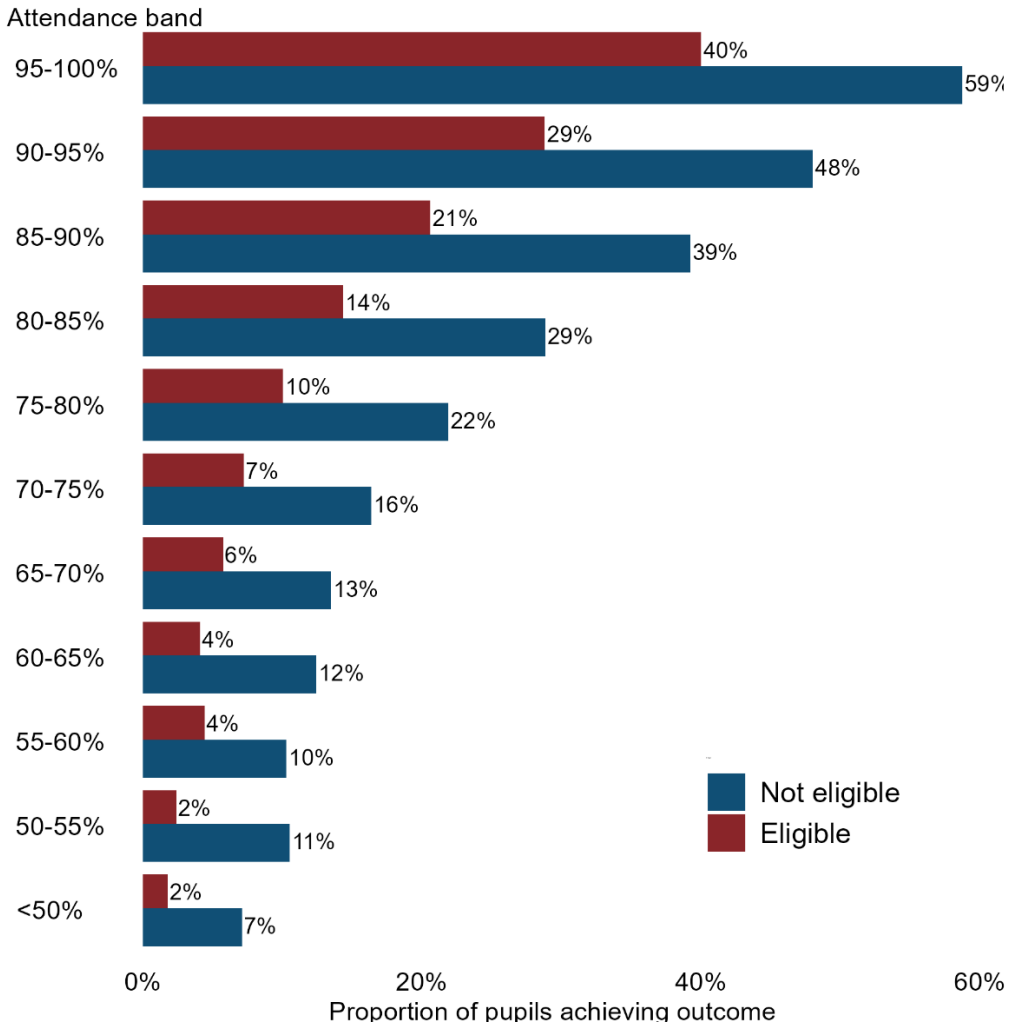
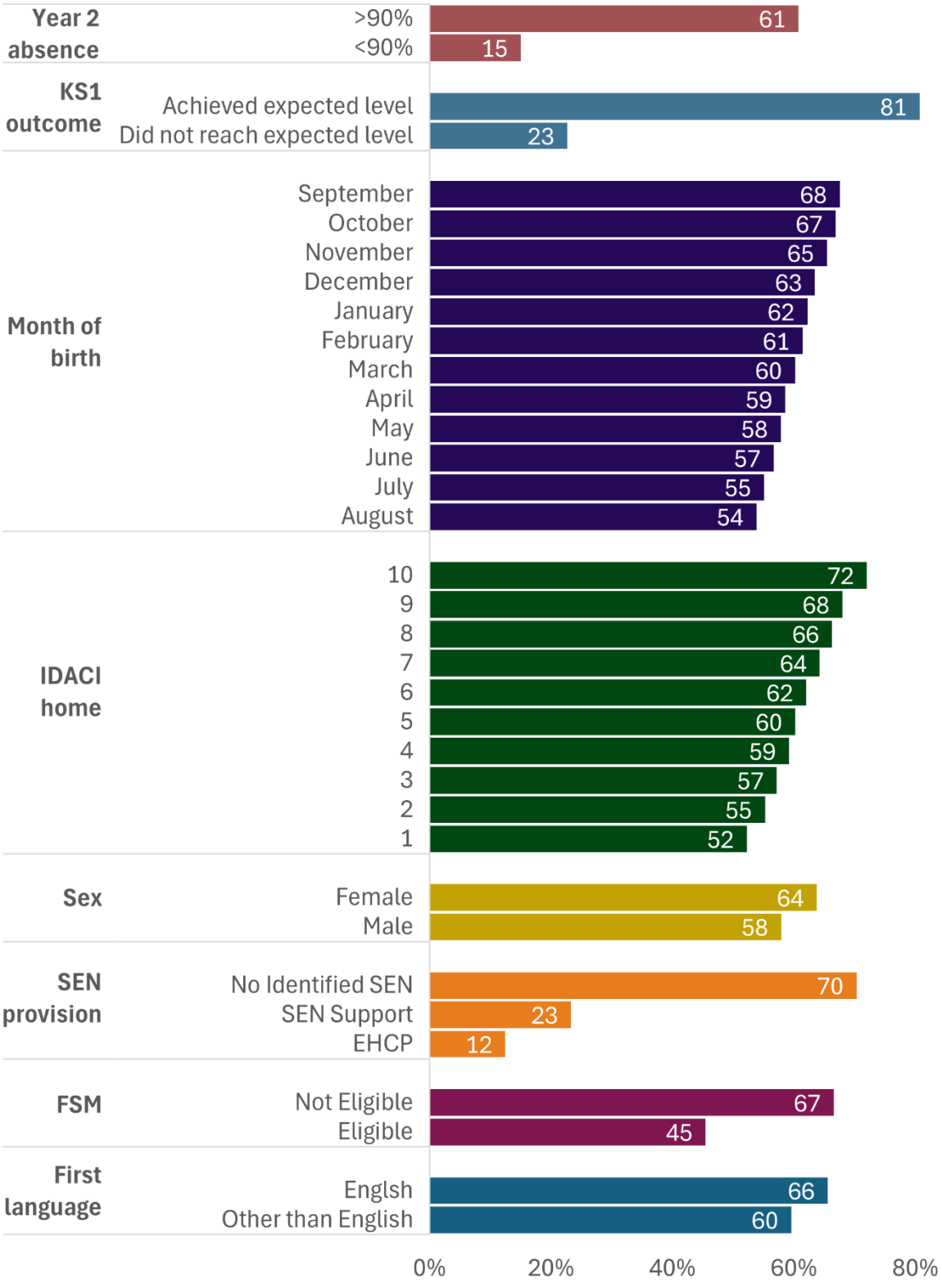


Figure 3 The relationship between attendance and successful outcome for KS2 pupils by FSM eligibility



The proportion of pupils reaching the expected level grouped by confounding factors are shown in Figure 4. Pupils who have SEN support or an EHCP, are male or who do not have English as their first language each have a lower likelihood of achieving the expected outcome than those that have no known SEN, are female or who have English as their first language. There is a positive linear relationship between likelihood of achieving the expected outcome and the IDACI decile a pupil lives in. There is negative linear relationship between the month of birth and the expected outcome with pupils born in September being more likely to achieve the expected outcome than those born in August of the same school year. Pupils who achieved the expected standard in KS1 were 14 times more likely to reach the expected level at KS2 than those that did not. Pupils who missed less than 10% of possible sessions in Year 2 were 11 times more likely to achieve the expected level at KS2 than pupils who missed more than 10% of sessions in Year 2. The unadjusted odds ratios for each factor compared to the reference variable is shown in Annex B.

Figure 4 Proportion of pupils in the KS2 cohort who achieved the expected outcome in RWM for compounding factors



Proportion of pupils achieving the expected outcome.

The interaction of these variables was controlled for using the regression model to isolate the effect of attendance on attainment. The factors, reference groups and regression outputs are shown in Annex B. The odds ratio of the attendance bands were calculated for each band relative to pupils with 95-100% attendance and then between each possible pair of attendance bands. Figure 5 shows a heat map with the odds ratio of reaching the expected level in RWM when other characteristics and the school URN a pupil attends is controlled for. The number, N, in each square represents the number of times more likely a pupil who attends for the amount of time shown in Attendance Band 1 on the x-axis is to reach the expected standard than a pupil who attends for the amount of time shown on Attendance Band 2 on the y-axis. The intensity of colour demonstrates the ratio, N, increasing in intensity as the difference in the two bands increases.

Pupils who attend 95-100% of the time are 1.3x more likely to achieve the expected standard in Reading, Writing and Maths than pupils who attend 90-95% of the time. Pupils who attend 90-95% of the possible sessions are also 1.3x more likely to achieve the outcome than pupils who attend 85-90% of the time.

At the lower attendance bands, pupils who attend 50-55% of the time are twice as likely as pupils who are severely absent (attending less than 50% of the time) to achieve the expected standard.

Figure 5 Heatmap showing the odds ratios between pairs of possible attendance bands during Year 6 after other pupil characteristics and the school attended is controlled for. For example, the highlighted boxes demonstrate that a pupil who attends 95-100% of the time (x-axis) is 1.3 times more likely to achieve the expected standard than a pupil who attends 90-95% of the time (y-axis)

↕ Attendance Band 2

	<50%	50-55%	55-60%	60-65%	65-70%	70-75%	75-80%	80-85%	85-90%	90-95%	95-100%
95-100%	*0.04*	*0.08*	*0.12*	*0.17*	*0.22*	*0.26*	*0.33*	*0.42*	*0.56*	*0.76*	1
90-95%	*0.06*	*0.11*	*0.15*	*0.22*	*0.29*	*0.34*	*0.44*	*0.56*	*0.74*	1	1.3
85-90%	*0.07*	*0.14*	*0.21*	*0.3*	*0.39*	*0.46*	*0.59*	*0.75*	1	1.3	1.8
90-85%	*0.10*	*0.19*	*0.28*	*0.40*	*0.52*	*0.61*	*0.79*	1	1.3	1.8	2.4
80-85%	*0.13*	*0.24*	*0.35*	*0.50*	*0.67*	*0.78*	1	1.3	1.7	2.3	3.0
75-80%	*0.16*	*0.31*	*0.45*	*0.65*	*0.85*	1	1.3	1.6	2.2	2.9	3.8
65-70%	*0.19*	*0.37*	*0.53*	*0.76*	1	1.2	1.5	1.9	2.5	3.4	4.5
60-65%	*0.25*	*0.49*	*0.70*	1	1.3	1.5	2.0	2.5	3.4	4.5	6.0
55-60%	*0.36*	*0.70*	1	1.4	1.9	2.2	2.9	3.6	4.8	6.5	8.5
50-55%	*0.51*	1	1.4	2.1	2.7	3.2	4.1	5.2	6.9	9.3	12.2
50%	1	2.0	2.8	4.0	5.3	6.2	8.0	10.1	13.5	18.1	23.9

↔ Attendance Band 1

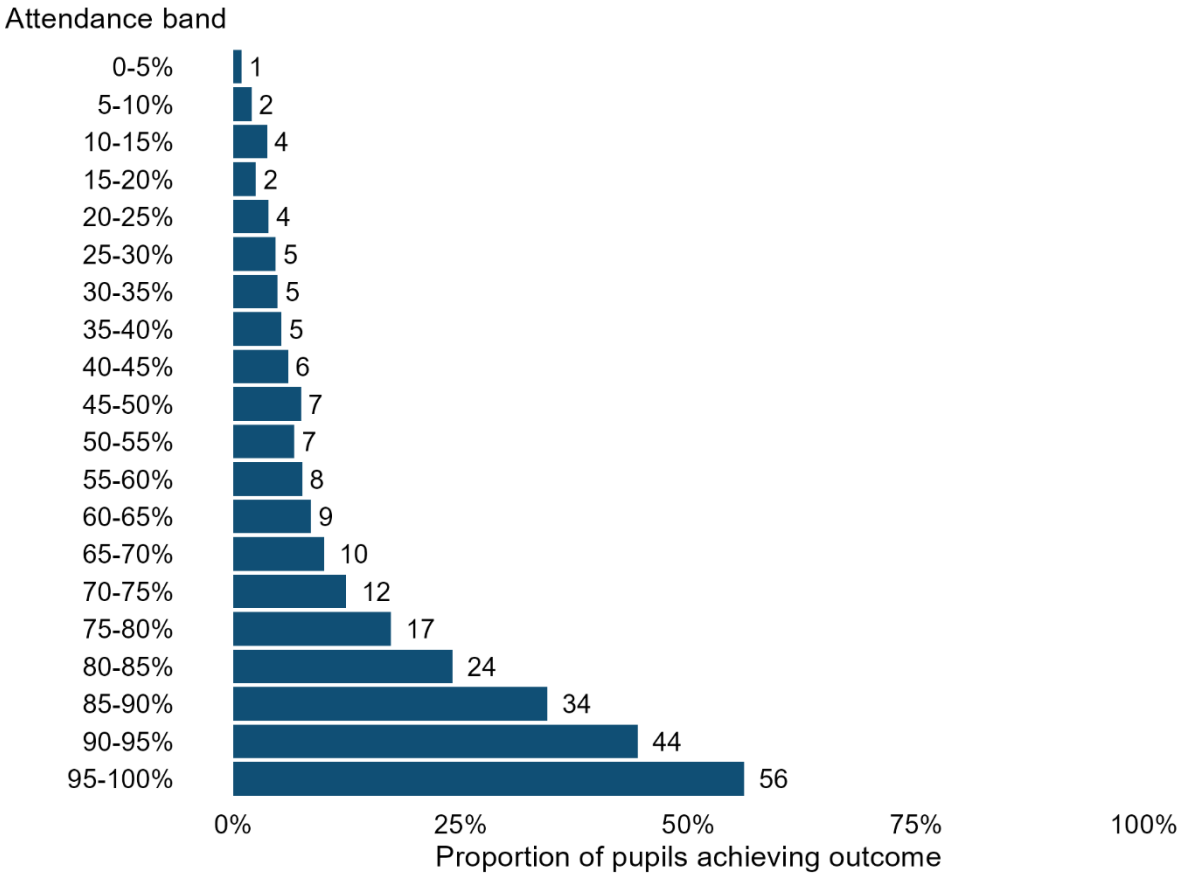
(pupils in Attendance Band 1 are N times more likely to achieve outcome than pupils in Band 2)

4.3 The link between absence and attainment at KS4

For pupils at the end of KS4 in 2022/23, as the 5% band of attendance during Year 11 increases the proportion of pupils achieving GCSE Maths and English Grade 5 increases as shown in Figure 6. Over the entire cohort, pupils with 95-100% attendance were 1.6 times more likely to achieve the expected outcome in reading, writing and maths than

pupils with 90-95% attendance and 2.4 times more likely than pupils with 85-90% attendance. Due to low numbers of pupils in the low attendance bandings they are aggregated into a single <50% attendance band for subsequent analysis.

Figure 6 Proportion of pupils in each 5% attendance band for Year 11 achieving Grade 5 or above in English and Maths GCSE at the end of KS4



Pupil characteristics are not evenly distributed across the attendance bands with pupils eligible for free school meals under-represented in the higher attendance bands (accounting for only 14% of pupils in the 95-100% attendance band) and over-represented in the lower attendance bands (50% in the <50% band) compared with the average of the entire KS4 cohort of 21.3% (Figure 7). Across all attendance bands, pupils eligible for FSM have a lower likelihood of achieving Grade 5 or above in English and Maths GCSE (Figure 8)[15].

[15] See Annex C for charts recreated with approximate number of days missed in place of attendance band.

Figure 7 The relationship between FSM eligibility and attendance in Year 11 for KS4 pupils

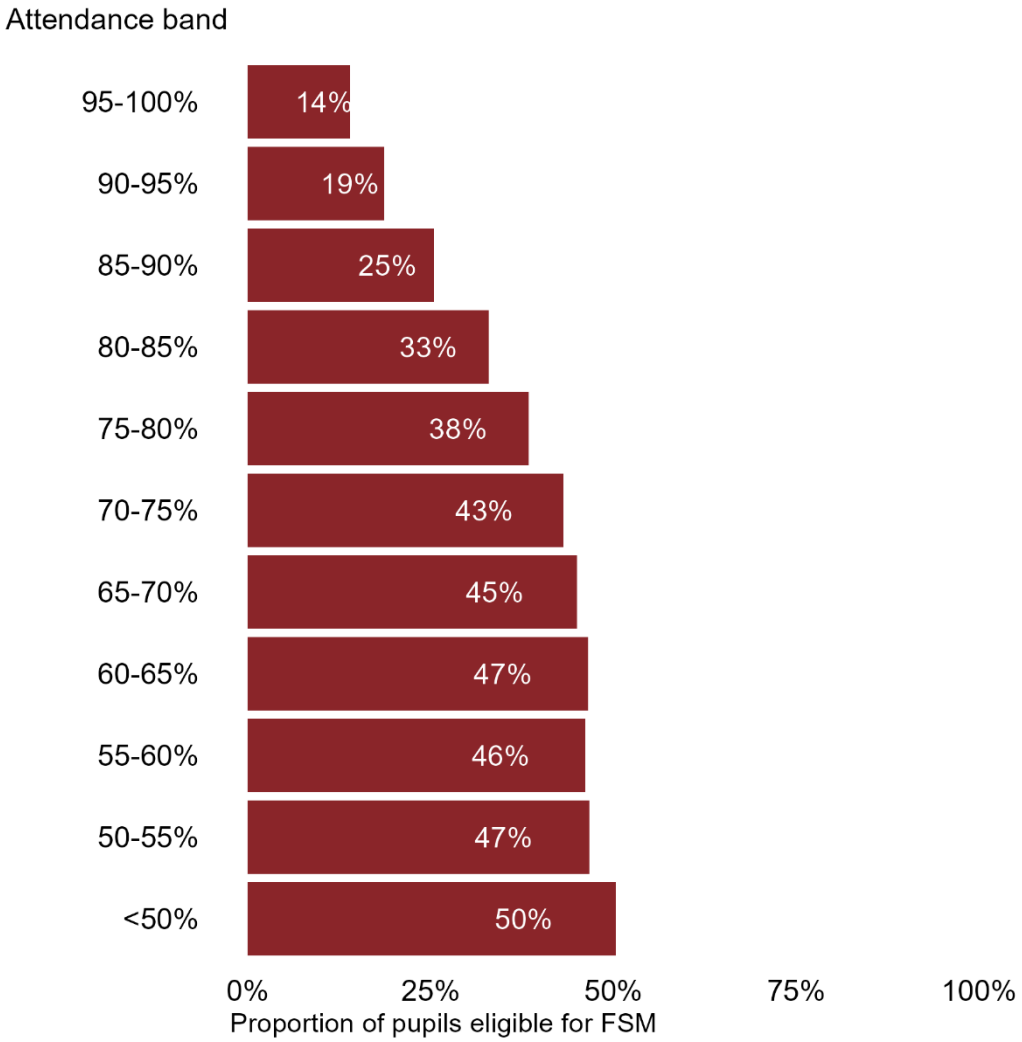
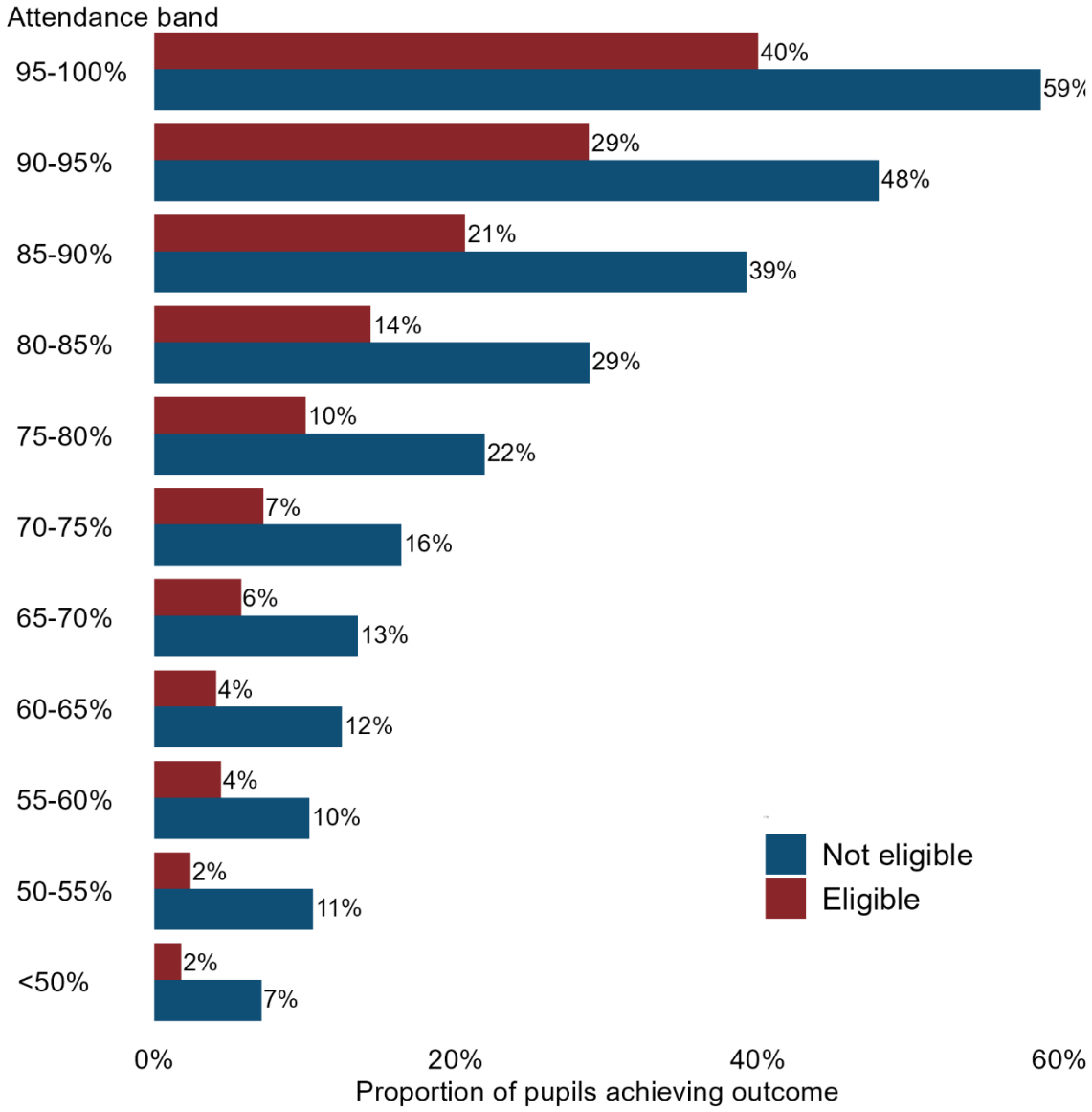
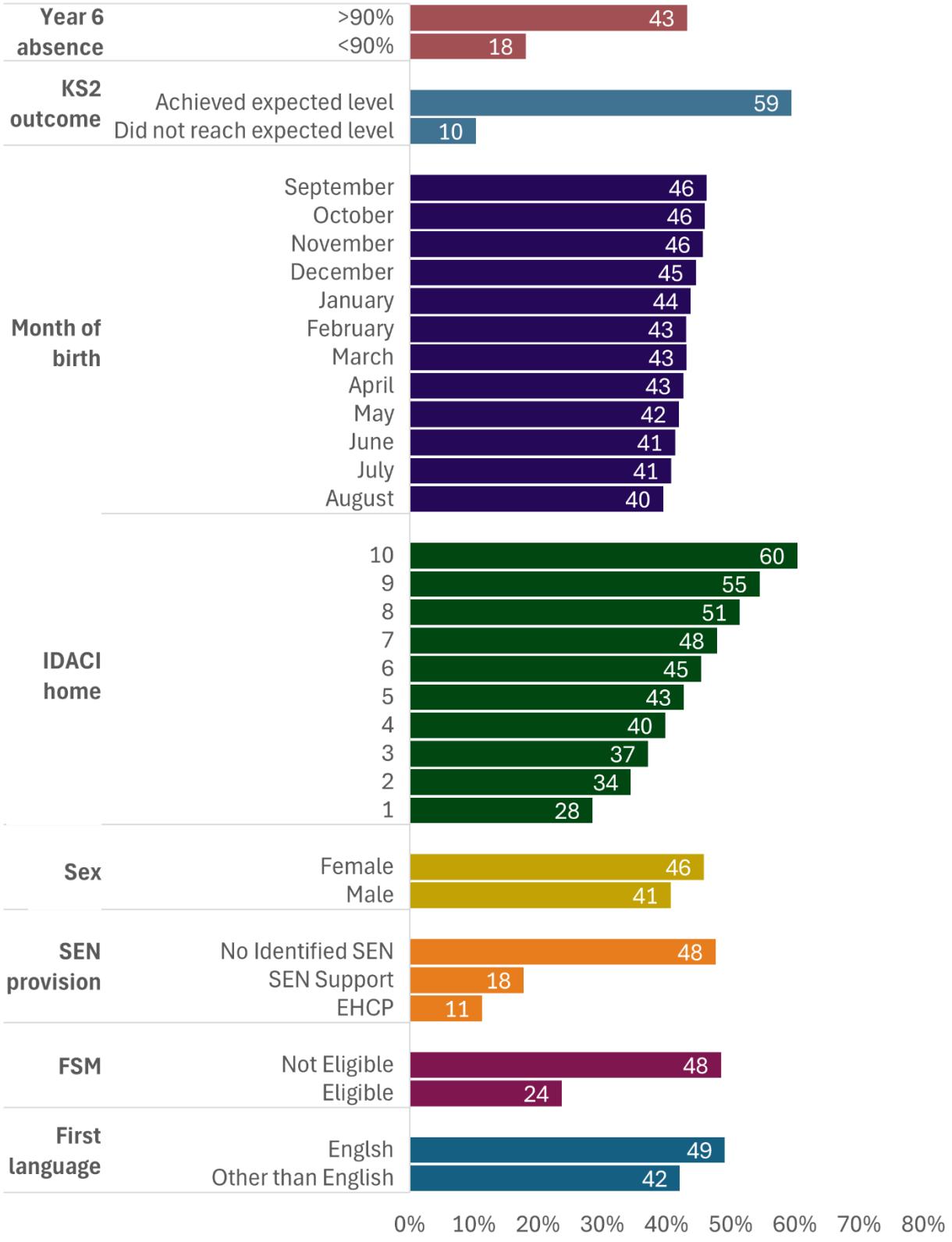


Figure 8 The relationship between attendance and successful outcome for KS4 pupils by FSM eligibility



Pupils who are eligible for FSM, who have SEN support or an EHCP, are male and do not have English as their first language each have a lower likelihood of achieving the expected outcome (Figure 9). There is a positive linear relationship between likelihood of achieving the expected outcome and the IDACI decile a pupil lives in. There is negative linear relationship between the month of birth and the expected outcome with pupils born in September being more likely to achieve the expected outcome than those born in August. Pupils who achieved the expected standard in KS2 were 13 times more likely to reach the expected level at KS2 than those that did not. Pupils who missed less than 10% of possible sessions in Year 6 were 4 times more likely to achieve the expected level at KS2 than pupils who missed more than 10% of sessions in Year 2. The odds ratios for each factor compared to the reference level is shown in Annex B.

Figure 9 Proportion of pupils in the KS4 cohort who achieved Grade 5 in Maths and English GCSE for compounding factors



Proportion of pupils achieving the expected outcome.

The interaction of these variables was controlled for using the regression model. The factors, reference groups and regression outputs are shown in Annex B. The odds ratio of the attendance bands were calculated for each band relative to pupils with 95-100% attendance and then between each possible pair of attendance bands. Figure 10 shows a heat map with the odds ratio of gaining a Grade 5 in both English and Maths GCSE when other characteristics and the school URN a pupil attends is controlled for. The number, N, in each square represents the number of times more likely a pupil who attends for the amount of time shown in Attendance Band 1 on the x-axis is to reach the expected standard than a pupil who attends for the amount of time shown on Attendance Band 2 on the y-axis. The intensity of colour demonstrates the ratio, N, increasing as the difference in the two bands increases.

Pupils who attend 95-100% of the time are 1.9x more likely to achieve Maths and English GCSE Grade 5 than pupils who attend 90-95% of the time. Pupils who attend 90-95% of the possible sessions are 1.6x more likely to achieve the outcome than pupils who attend 85-90% of the time.

At the lower attendance bands, pupils who are severely absent (attending less than 50% of the time are half as likely to pass both GCSEs at Grade 5 or above than pupils who attend 55-60% of the time.

Figure 10 Heatmap showing the odds ratio pairs between pairs of possible attendance bands during Year 11 after other pupil characteristics and the school attended is controlled for. For example, the highlighted boxes demonstrate that a pupil who attends 95-100% of the time (x-axis) is 1.9 times more likely to achieve English and Maths Grade 5 than a pupil who attends 90-95% of the time (y-axis)

↕ Attendance Band 2

	<50%	50-55%	55-60%	60-65%	65-70%	70-75%	75-80%	80-85%	85-90%	90-95%	95-100%
95-100%	*0.04*	*0.07*	*0.08*	*0.09*	*0.10*	*0.12*	*0.16*	*0.21*	*0.33*	*0.53*	1
90-95%	*0.08*	*0.13*	*0.15*	*0.17*	*0.19*	*0.22*	*0.31*	*0.40*	*0.62*	1	1.9
85-90%	*0.13*	*0.21*	*0.25*	*0.27*	*0.30*	*0.36*	*0.49*	*0.64*	1	1.6	3.0
90-85%	*0.20*	*0.33*	*0.39*	*0.43*	*0.47*	*0.56*	*0.77*	1	1.6	2.5	4.8
80-85%	*0.26*	*0.42*	*0.50*	*0.55*	*0.61*	*0.73*	1	1.3	2.0	3.3	6.2
75-80%	*0.36*	*0.58*	*0.69*	*0.76*	*0.84*	1	1.4	1.8	2.8	4.5	8.5
65-70%	*0.43*	*0.69*	*0.82*	*0.90*	1	1.2	1.6	2.1	3.3	5.3	10.1
60-65%	*0.48*	*0.76*	*0.91*	1	1.1	1.3	1.8	2.6	3.7	5.9	11.2
55-60%	*0.53*	*0.84*	1	1.1	1.2	1.5	2.0	2.6	4.1	6.5	12.3
50-55%	*0.63*	1	1.2	1.3	1.5	1.7	2.4	3.1	4.8	7.7	14.7
50%	1	1.6	1.9	2.1	2.3	2.8	3.8	4.9	7.7	12.3	23.3

↔ Attendance Band 1

(pupils in Attendance Band 1 are N times more likely to achieve outcome than pupils in Band 2)

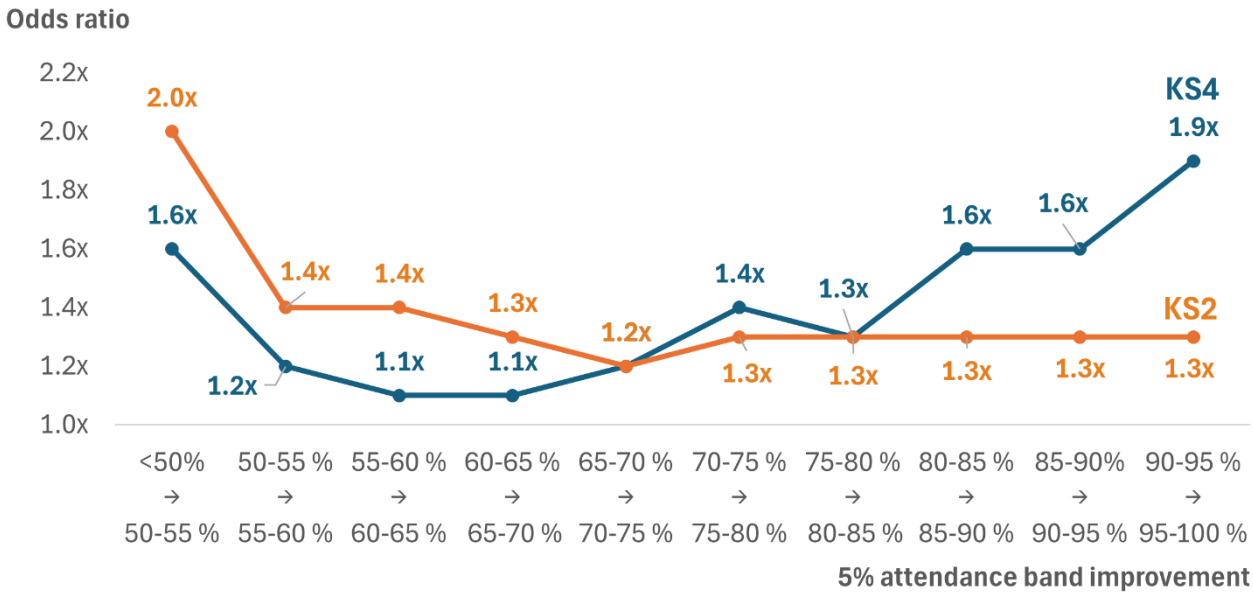
5. Conclusions

Overall the analysis shows that as the level of attendance in the year of assessment at the end of KS2 and KS4 increases, the likelihood of achieving key attainment outcomes at the end of KS2 and KS4 increases. When controlling for other factors known to affect achievement, such as prior attainment and pupil characteristics, the attendance band has a positive link to attainment.

The odds ratios of the different attendance bands demonstrate the increase in likelihood of achieving the expected outcome at KS2 and KS4 associated with increased attendance from one band to another as shown in Figure 11. A pupil will move up one attendance band by attending school between 0-20 days extra in the year.

At Key Stage 2, the gain in attainment ranges from 1.2x greater likelihood when moving up one band from 65-70% to 70-75%, up to 1.4x greater likelihood moving from 50-55% and 55-60% band. Improved attendance from being severely absent to 50-55% attendance during Year 6 is associated with double the chance of achieving the expected standard in reading, writing and maths but includes pupils this is distorted by the pupils in the severely absent band with extremely low attendance.

Figure 11 The odds ratio of achieving the expected standard at KS2 and KS4 when moving up a single 5% attendance band



At Key Stage 4, the gain in attainment between single bands ranges from 1.1x greater likelihood between 55-60% and 60-65% bands, up to 1.9x greater likelihood between 90-95% and 95-100% bands with greater gains at the higher attendance bands. However, there is also 1.6x greater likelihood of getting Maths and English GCSE level 5 associated with improved attendance from severely absent to 50-55% attendance of the

Year 11 but again the severely absent range includes some pupils with extremely low attendance.

At school level, as more pupils are in the higher attendance bands (Table 4), the greatest gains in outcome could be associated with improving the attendance of pupils in the 90-95% band into the 95-100 attendance band (0-20 days extra in school over the year). However, this analysis also shows that at a pupil level there are gains in attainment associated with improving attendance by a single 5% band across the range of attendance bands.

Annex A: Technical notes

A.1 Data sources and manipulation

1. Absence information is collected termly through the school census. The school census guidance includes information on the collection of pupil absence data. A technical specification is also available, which gives detailed explanations of what data are collected and what validation processes are carried out.

2. Attainment data for all pupils at the end of Key Stage 4 (KS4) is collected from the awarding bodies by the Department's contractor. Further information can be found in the 'Key Stage 4 performance' and methodology information section[16].

   Attainment data for all pupils at the end of Key Stage 2 (KS2) is collected by the Standards and Testing Agency (STA), an executive agency of the Department. Further information can be found in the 'Attainment in primary schools in England' quality and methodology information section[17].

3. The pupil absence and attainment data used in this release have been created by matching data cuts from the National Pupil Database (NPD), a longitudinal database which holds a wide range of pupil level data for schools across England. Absence data was derived from school census returns, which feed into the NPD. Only pupils with a single enrolment during the assessment year have been included. This means pupils who are dual registered at two schools or settings, or who moved school during the year are not included in the analysis. Attainment data has been sourced from annual NPD extracts, duplicates have been removed and the data has been filtered to leave just those pupils at the end of the Key Stage with valid attainment data. This data was then matched together using the NPD's unique pupil reference number. Any pupils not meeting these criteria were removed before the analysis was undertaken.

A.2 Data coverage and quality

4. This release includes pupils in state-funded mainstream schools including academies, free schools and city technology colleges, excluding state-funded special schools, independent schools, independent special schools, non-maintained special schools, hospital schools and alternative provision including academy and free school alternative provision and pupil referral units. Information on children not registered at a school is not included.

[16] Key Stage 4 performance: methodology. Key Stage 4 performance, Methodology - Explore education statistics - GOV.UK (explore-education-statistics.service.gov.uk)
[17] Key Stage 2 attainment methodology. Key Stage 2 attainment, Methodology - Explore education statistics - GOV.UK (explore-education-statistics.service.gov.uk)

Pupils who are dual registered at two schools or placements, or who moved schools during the year are not included. Pupils who had no recorded data for the characteristics, attainment or absence data used in the model were removed.

5. The figures presented in this report relate to attainment, absence and characteristic data from the 2022/23 academic year for the end of KS2 and end of KS4 pupils. However, prior absence and prior attainment data for pupils was matched for the year pupils were at the end of the previous Key Stage. For KS2 pupils, prior absence rates of <90% (persistent absence) and attainment in KS1 tests in Year 2 during 2018/19 were joined to the pupils dataset. Similarly for pupils at the end of KS4 in 2022/23, prior absence data and KS2 attainment data for Year 6 in 2017/18 have been joined to the pupil records.
6. For the absence data for Year 2 and Year 6 pupils absence data was calculated using 6 half terms of data. For pupils in Year 11, only five half terms of absence data have been included; the second half of the summer term has been ignored due to high levels of study leave and authorised absence during the last half term of Year 11.
7. The absence information collected via the school census counts the number of pupil enrolments rather than the number of pupils. Pupil absence is recorded for the period a pupil is enrolled at a school. Where a pupil moves schools or has a dual registration their absence may be returned from both schools, if both schools return absence data via the school census. For the purposes of this release, the research has been restricted to pupils with a single registration. This is not the standard approach for the Department's other pupil absence releases; however it was deemed to be appropriate for this analysis in order to avoid double counting pupils with high levels of absence.

A.3 Methodology

1. All maintained schools are required to provide two possible sessions per day, morning and afternoon, to all pupils and schools must meet for at least 380 sessions or 190 days during any school year to educate their pupils.
2. Pupils that do not have more than 100 possible sessions in the assessment year or during the final year of the previous Key Stage have been removed from the analysis. 100 sessions relates to approximately one terms worth of attendance. This ensures that pupils with extremely low numbers of possible sessions (for whom we would not be able to determine a full picture of the link between their absence and attainment) were removed and did not distort the findings.
3. For pupils who have valid attainment data at the end of KS2 or KS4, there are a number of possible reasons which may mean they do not have a full terms worth of possible sessions in both the assessment year and the final year of the previous

Key Stage and are therefore excluded from the analysis in this report. Possible reasons include a pupil leaving or joining the state-funded mainstream education sector between the two time points; pupils becoming or ceasing to be dual registered with a special school or a Pupil Referral Unit (or another form of alternative provision) or matching or recording errors.

4. The attendance rate bands in this release are based on all the periods of absence (authorised and unauthorised) a pupil has accrued across the possible sessions. The overall absence rate was calculated from the number of overall absence sessions recorded in the census as:

$$Overall\ Absence\ Rate = \frac{number\ of\ overall\ absence\ sessions}{number\ possible\ sessions} \times 100,$$

and the attendance rate calculated as:

$$Attendance\ Rate = 100 - Overall\ Absence\ Rate.$$

The attendance rate was then banded into 5% attendance bands as

0-5% = Pupils whose attendance was in the range 0.00-4.99%,
5-10% = Pupils whose attendance was in the range 5.00-9.99%,
10-15% = Pupils whose attendance was in the range 10.00-14.99%,
15-20% = Pupils whose attendance was in the range 15.00-19.99%,
20-25% = Pupils whose attendance was in the range 20.00-24.99%,
25-30% = Pupils whose attendance was in the range 25.00-29.99%,
30-25% = Pupils whose attendance was in the range 30.00-34.99%,
35-40% = Pupils whose attendance was in the range 35.00-39.99%,
40-45% = Pupils whose attendance was in the range 40.00-44.99%,
45-50% = Pupils whose attendance was in the range 45.00-49.99%,
50-55% = Pupils whose attendance was in the range 50.00-54.99%,
55-60% = Pupils whose attendance was in the range 55.00-59.99%,
60-65% = Pupils whose attendance was in the range 60.00-64.99%,
65-70% = Pupils whose attendance was in the range 65.00-69.99%,
70-75% = Pupils whose attendance was in the range 70.00-74.99%,
75-80% = Pupils whose attendance was in the range 75.00-79.99%,
80-85% = Pupils whose attendance was in the range 80.00-84.99%,
85-90% = Pupils whose attendance was in the range 85.00-89.99%,
90-95% = Pupils whose attendance was in the range 90.00-94.99%,
95-100% = Pupils whose attendance was in the range 95.00-100.00%.

For the regression model, pupils in the bands below 50% were aggregated into a single band <50% to increase the pupil numbers for the model convergence.

Annex B: Model output

B.1 Pupils achieving expected level in RWM at the end of KS2

The estimated regression odds ratios presented here control where possible for the effects of all the other factors affecting a pupil, to identify the change in likelihood of reaching the expected standard in reading, writing and maths due their attendance band during Year 6.

Table 5 KS2 regression model output

Factor	Unadjusted Odds Ratio	p-value	Odds Ratio	Standard Error (SE)
Attendance Band				
<50%	Reference			
50-55%	2.16	<0.0033**	1.95	0.23
55-60%	2.46	<0.0001****	2.80	0.20
60-65%	2.91	<0.0001****	4.01	0.18
65-70%	3.85	<0.0001****	5.30	0.16
70-75%	5.11	<0.0001****	6.21	0.14
75-80%	6.87	<0.0001****	7.97	0.13
80-85%	9.57	<0.0001****	10.13	0.13
85-90%	14.01	<0.0001****	13.50	0.13
90-95%	22.02	<0.0001****	18.14	0.13
95-100%	32.96	<0.0001****	23.90	0.13
Attendance in Year 2				
<90%	Reference			
>90%	10.74	*0.7179*	*1.17*	*0.44*
Prior Attainment				
Did not reach expected standard	Reference			
Reached expected standard	14.36	<0.0001****	12.80	0.01
FSM Eligibility				
Eligible	Reference			
Not Eligible	2.39	<0.0001****	1.52	0.01
SEN Provision				
EHCP	Reference			
SEN Support	2.13	<0.0001****	1.72	0.04
No Identified Need	16.66	<0.0001****	5.97	0.03
Month of Birth				
August	Reference			
July	1.04	0.0415*	0.96	0.02
June	1.12	*0.9479*	*1.00*	*0.02*
May	1.18	*0.6030*	*0.99*	*0.02*

Factor	Unadjusted Odds Ratio	p-value	Odds Ratio	Standard Error (SE)
April	1.21	*0.2150*	*0.97*	*0.02*
March	1.29	*0.7523*	*0.99*	*0.02*
February	1.36	*0.5956*	*0.99*	*0.02*
January	1.42	*0.9825*	*1.00*	*0.02*
December	1.49	*0.7640*	*0.99*	*0.02*
November	1.64	*0.0617*	*1.04*	*0.02*
October	1.74	0.0014**	1.07	0.02
September	1.80	0.0105*	1.06	0.02
Sex				
Male	Reference			
Female	1.28	<0.0001****	0.94	0.01
First Language				
Other than English	Reference			
English	1.29	<0.0001****	1.43	0.01
Home IDACI Decile				
1	Reference			
2	1.13	*0.5990*	*1.01*	*0.02*
3	1.21	0.0173*	1.05	0.02
4	1.33	<0.0001****	1.11	0.02
5	1.39	<0.0001****	1.11	0.02
6	1.50	<0.0001****	1.18	0.02
7	1.65	<0.0001****	1.26	0.02
8	1.80	<0.0001****	1.30	0.02
9	1.94	<0.0001****	1.38	0.02
10	2.39	<0.0001****	1.60	0.02

Significance codes: <0.0001 '****' 0.001 '***' 0.01 '**' 0.05'*' *Not significant in italics*

B.1 Pupils achieving level 5 or above in Maths and English GCSE at the end of KS4

The estimated regression odds ratios presented here control where possible for the effects of all the other factors affecting a pupil, to identify the change in likelihood of achieving Grade 5 in GCSE Maths and English due their attendance band during Year 11.

Table 6 KS4 regression model output

Factor	Unadjusted Odds Ratio	p-value	Odds Ratio	Standard Error (SE)
Attendance Band				
<50%	Reference			
50-55%	1.62	<0.0001***	1.59	0.10
55-60%	1.86	<0.0001***	1.89	0.09
60-65%	2.20	<0.0001***	2.09	0.08
65-70%	2.47	<0.0001***	2.32	0.07
70-75%	3.06	<0.0001***	2.75	0.06
75-80%	4.56	<0.0001***	3.78	0.05
80-85%	6.87	<0.0001***	4.89	0.05
85-90%	11.37	<0.0001***	7.68	0.05
90-95%	17.30	<0.0001***	12.32	0.05
95-100%	27.70	<0.0001***	23.35	0.05
Attendance in Year 6				
<90%	Reference			
>90%	3.44	<0.0001***	0.21	0.30
Prior Attainment				
Did not reach expected standard	Reference			
Reached expected standard	12.73	<0.0001***	10.92	0.01
FSM Eligibility				
Eligible	Reference			
Not Eligible	3.03	<0.0001***	1.55	0.01
SEN provision				
EHCP	Reference			
SEN Support	1.69	0.0003**	1.21	0.04
No Identified Need	7.12	<0.0001***	2.28	0.04
Month of Birth				
August	Reference			
July	*1.05*	*0.08*	*1.04*	*0.02*
June	*1.08*	*0.52*	*1.01*	*0.02*
May	*1.10*	*0.17*	*1.03*	*0.02*
April	1.14	0.0133*	1.05	*0.02*
March	1.16	0.0032*	1.06	*0.02*
February	1.16	0.0021*	1.07	*0.02*
January	1.19	0.0037*	1.08	0.02
December	1.23	0.0030*	1.10	0.02
November	1.29	<0.0001***	1.13	0.02
October	1.30	<0.0001***	1.10	0.02
September	1.32	<0.0001***	1.11	0.02
Sex				
Male	Reference			
Female	1.23	<0.0001***	1.25	0.01

Factor	Unadjusted Odds Ratio	p-value	Odds Ratio	Standard Error (SE)
First Language				
Other than English	Reference			
English	1.33	<0.0001***	1.57	0.01
Home IDACI Decile				
1	Reference			
2	1.32	<0.0001***	1.13	0.02
3	1.48	<0.0001***	1.21	0.02
4	1.66	<0.0001***	1.28	0.02
5	1.87	<0.0001***	1.42	0.02
6	2.09	<0.0001***	1.54	0.02
7	2.31	<0.0001***	1.60	0.02
8	2.65	<0.0001***	1.74	0.02
9	3.01	<0.0001***	1.86	0.02
10	3.83	<0.0001***	2.08	0.02

Significance codes: <0.0001 '****' 0.001 '***' 0.01 '**' 0.05'*' *Not significant in italics*

B.3 Strength of the models

This regression uses a multi-level model containing all the factors. There is the potential for the standard errors of some of the individual control variables to be under-estimated as some factors may interact in ways not accounted for. The random effect captures the variability in the outcome due to the school attended (represented by the school URN). This allows the model to account for the fact that observations within the same school might be more similar to each other than to observations from different schools.

The area under the receiver operating curve (AUC)[18] was used to determine if the model is good. The AUC was 0.858 for the KS2 model and 0.852 for the KS4 model demonstrating excellent ability in distinguishing between pupils who will achieve the expected outcome and those who will not, based on attendance band and control factors. The Akaike information criterion (AIC) and Bayesian information criterion (BIC)[19] are also presented to allow comparison of this model with any future model.

[18] The AUC measures the ability of a model to correctly classify the outcome based on the values of the attendance band and control variables; 0.5 indicates a 50:50 chance and 1 indicates perfect predictive ability.

[19] The Akaike information criterion (AIC) and Bayesian information criterion (BIC) measure the relative performance of one model versus another, accounting for their complexity (i.e., the number of control variables). Lower values are generally preferred for both criteria, as they indicate lower estimated information loss.

Table 7 Model accuracy

Model	AIC	BIC	AUC (1=perfect prediction)
KS2 Model	337348.7	337770.9	0.858107
KS4 Model	358177.4	358600.1	0.852307

To test the sensitivity of the analysis to the training cohort selection, the KS4 model was rerun with a different seed initiating the training/validation split. The average difference in the odds ratio calculated in the model was found to be $3x10^{-5}$ suggesting the model is insensitive to the selection of the training and validation cohorts.

Annex C: Charts with approximate days missed

The model used in this analysis uses the percentage of possible sessions that each pupil attends for all pupils with more than 100 possible sessions. For ease of interpretation the charts in the report are presented in this annex with the 5% attendance bandings converted into the approximate number of days this would mean a pupil has missed over the course of the whole academic year.

Figure 12 Proportion of pupils achieving expected standard in RWM at the end of KS2 by the approximate number of days missed in Year 6

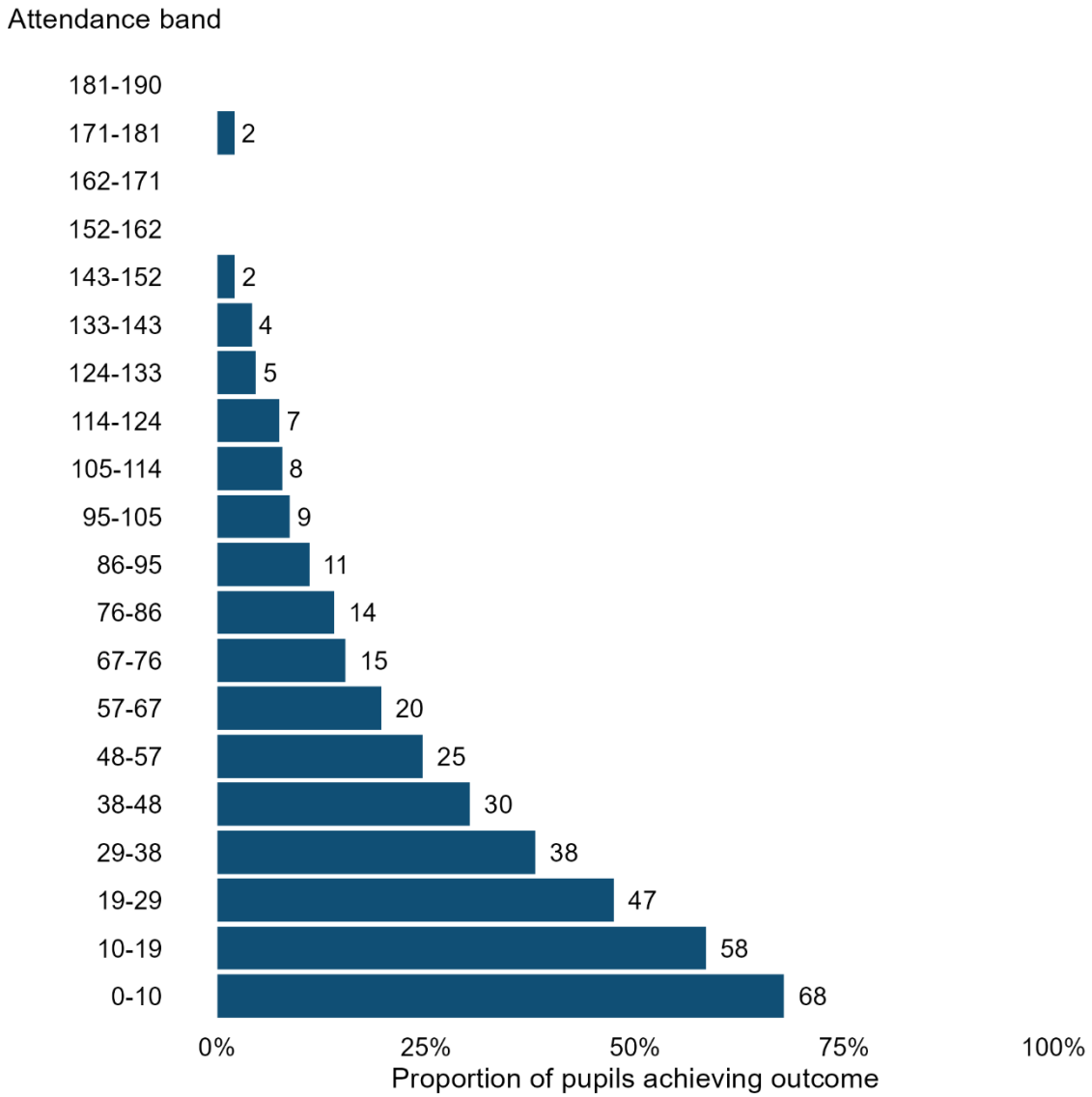


Figure 13 The relationship between FSM eligibility and days missed in Year 6 for KS2 pupils

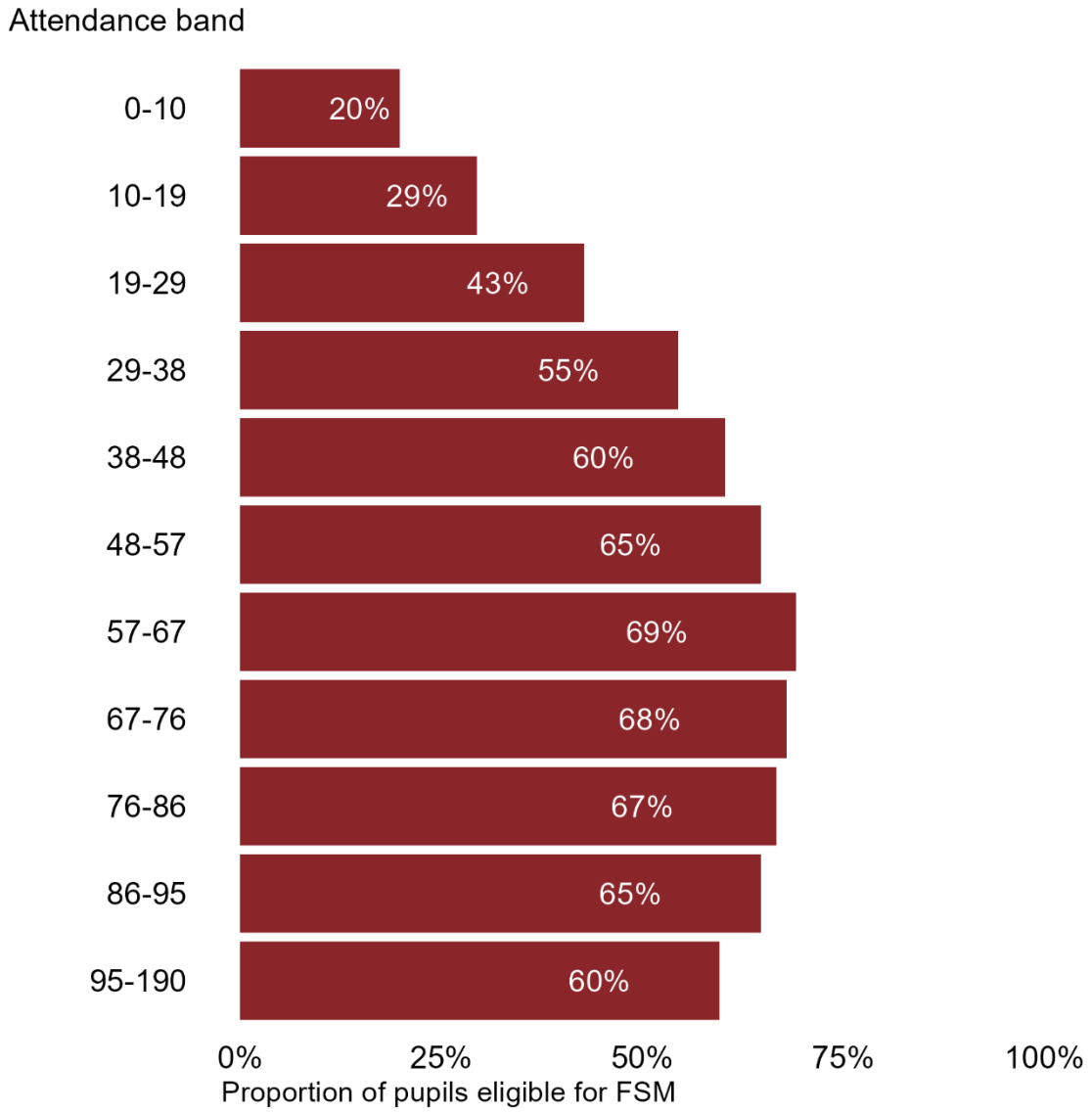


Figure 14 The relationship between days missed in Year 6 and successful outcome for KS2 pupils by FSM eligibility

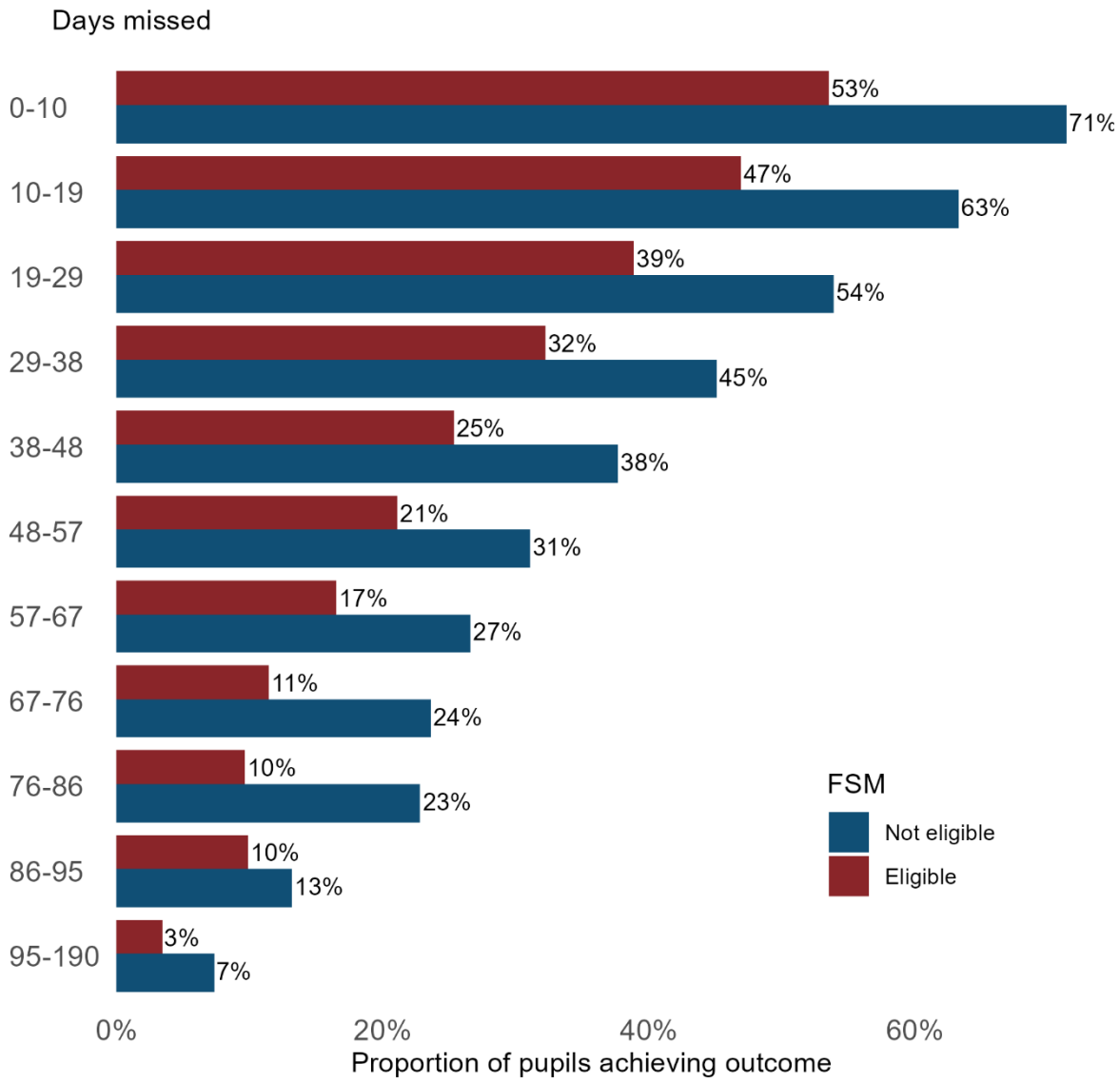


Figure 15 Proportion of pupils achieving Grade 5 or above in English and Maths GCSE at the end of KS4 by the approximate number of days missed in Year 11

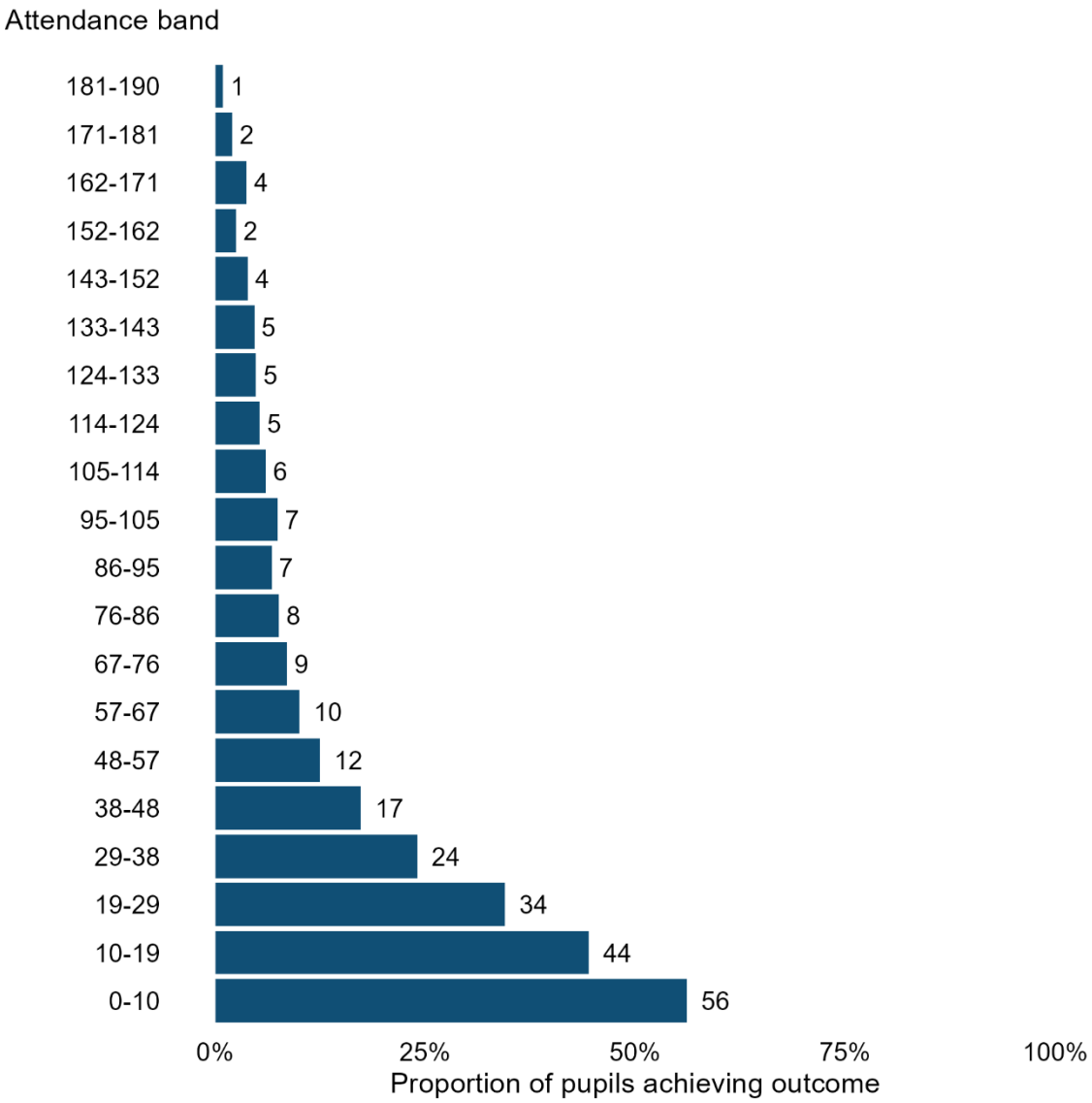

Attendance band
181-190
171-181
162-171
152-162
143-152
133-143
124-133
114-124
105-114
95-105
86-95
76-86
67-76
57-67
48-57
38-48
29-38
19-29
10-19
0-10
1
2
4
2
4
5
5
5
6
7
7
8
9
10
12
17
24
34
44
56
0%
25%
50%
75%
100%
Proportion of pupils achieving outcome

Figure 16 The relationship between FSM eligibility and days missed in Year 11 for KS4 pupils

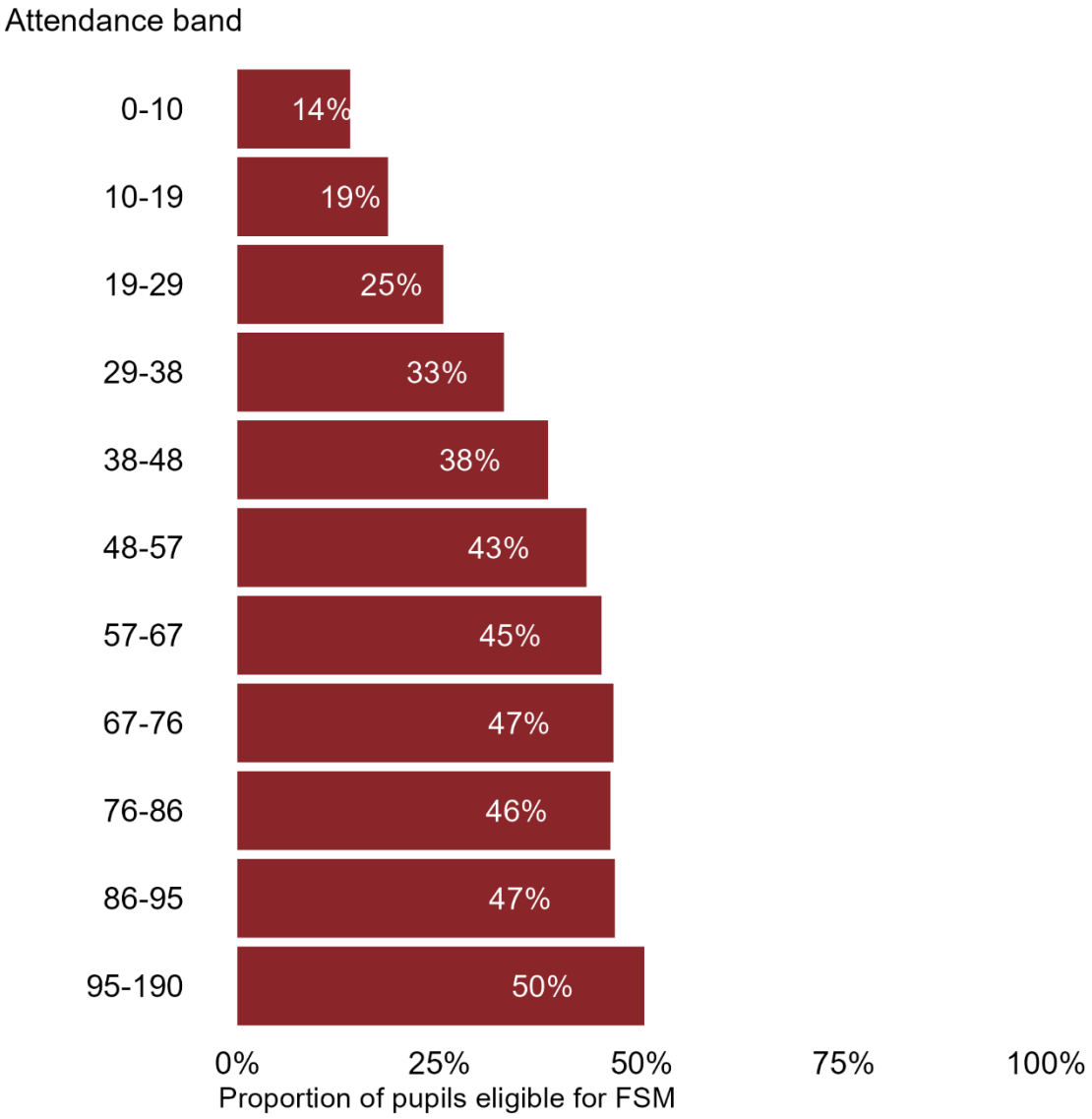


Figure 17 The relationship between days missed in Year 11 and successful outcome for KS4 pupils by FSM eligibility

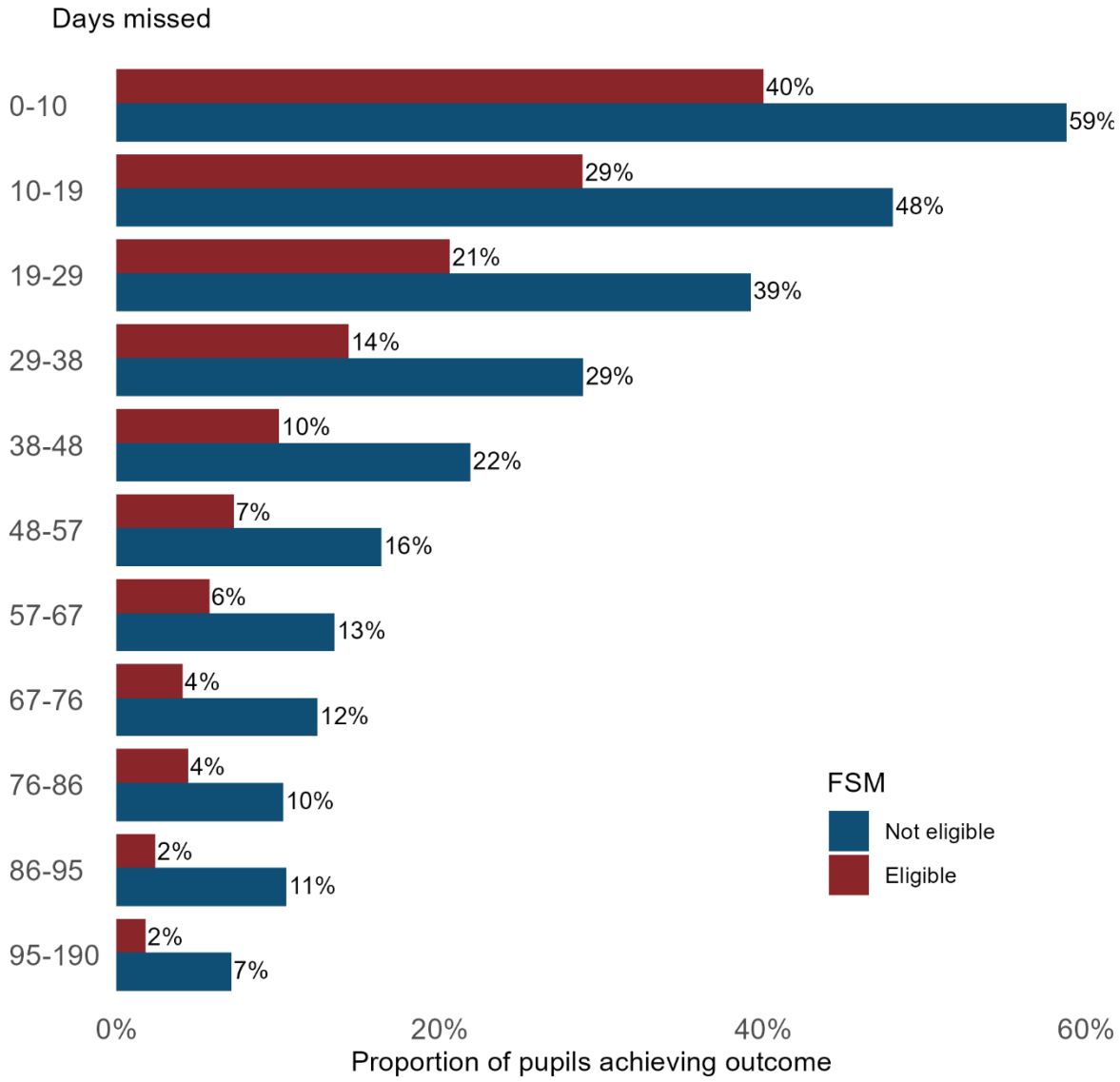
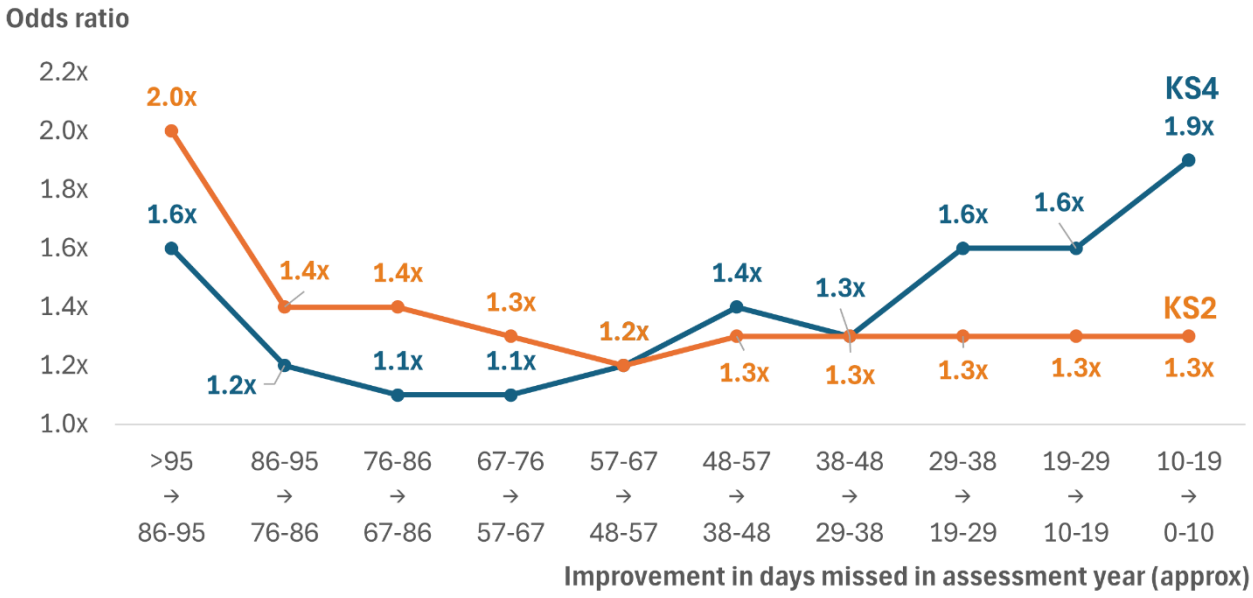


Figure 18 The odds ratio of achieving the expected standard at KS2 and KS4 when moving up a single 5% attendance band



Department for Education

© Department for Education copyright 2025

This publication is licensed under the terms of the Open Government Licence v3.0, except where otherwise stated. To view this licence, visit nationalarchives.gov.uk/doc/open-government-licence/version/3.

Where we have identified any third-party copyright information you will need to obtain permission from the copyright holders concerned.

Reference: RR1484

ISBN: 978-1-83870-608-1

For any enquiries regarding this publication, contact www.education.gov.uk/contactus.

This document is available for download at www.gov.uk/government/publications.