



Best Beginnings: Technical report

Predicting children's outcomes at the end of primary school
based on their early years foundation stage profile results

July 2020

Contents

Introduction: aims of this analysis	2
Executive summary	2
Background to the EYFS assessment	3
Cohort profile and methods	5
Cohort profile.....	5
Key groups of interest based on their EYFS scores.....	7
Differences between these key groups on baseline characteristics	11
Regression approach	11
Results: academic outcomes	14
Attainment at Key Stage 2 (age 11)	14
Attainment at Key Stage 1 (age 7).....	21
Results: Non-academic outcomes	26
Permanent or fixed term exclusions by the end of Year 5 (age 10)	26
High levels of unauthorised absence by the end of Year 5 (age 10)	30
Contact with children’s social care by the end of Year 6 (age 11)	34
Appendix A: Differences on baseline characteristics amongst EYFS groupings.....	38

Introduction: aims of this analysis

This analysis has two primary aims:

1. To examine how (if at all) children's Early Years Foundation Stage (EYFS) scores are associated with future outcomes independent of other factors recorded in the national pupil database (NPD) in the child's reception year?
2. To examine how accurately (if at all) we can predict children likely to experience worse future outcomes based on information in their reception year and how much additional predictive value EYFS assessment information provides on top of other information known about the child?

Executive summary

This analysis demonstrates three key findings:

- > Children with lower attainment at EYFS are more likely to have worse academic and non-academic outcomes, even after taking into account their baseline characteristics.
 - > These associations are strongest for Key Stage 1 and Key Stage 2 results, though associations also remain with non-academic outcomes such as exclusion from school, absence and contact with children's social care.
- > EYFS data provides important additional predictive value about the likelihood of being below the expected level at Key Stage 1 and Key Stage 2.
 - > This is also true for non-academic outcomes but to a lesser extent. For these outcomes, economic disadvantage (in the form of neighbourhood income deprivation and free school meal eligibility) has a greater impact on predictive accuracy.
- > There are small gains in predictive ability from including information on the specific combinations of EYFS domains that children are below the expected level in. Across most outcomes these EYFS groupings performed better than a simpler predictor of the total number of EYFS indicators where a child was below the expected level.
 - > However, an EYFS grouping based being at the expected level in less than half of the EYFS indicators performed well and may serve as a useful proxy.

Background to the EYFS assessment

The EYFS assessment is conducted at the end of children’s reception year (normally the year in which children turn 5). It involves teacher-assessed judgements of whether a child is:

- > meeting expected levels of development,
- > exceeding expected levels,
- > not yet reaching expected levels.

These assessments were restructured in 2012/13 and now cover seven domains of children’s development through teacher assessments of 17 key indicators. These are listed in Table 1 below.

Table 1: Indicators used in EYFS assessment

Domain	Indicator
Communication & Language	Listening and attention
	Understanding
	Speaking
Physical Development	Moving and handling
	Health and self-care
Personal, social and emotional development	Self-confidence and self-awareness
	Managing feelings and behaviour
	Making relationships
Literacy	Reading
	Writing
Maths	Numbers
	Shape, space and measures
Understanding the world	People and communities
	The World
	Technology
Expressive arts & design	Exploring and using media and materials
	Being imaginative

The base cohort for this analysis is the sample of children who completed EYFS assessments in the 2012/13 academic year and who can be matched to that year’s spring school census. This matching provides baseline pupil characteristics that are known to schools at the time of the EYFS assessment. This allows us to examine how useful a predictor EYFS scores are of future outcomes, independent of

the other characteristics known about the child (and available in public data).

These baseline characteristics are:

- > Gender
- > Ethnicity
- > SEN (special educational need) status in 2012/13
- > FSM eligibility status in 2012/13
- > 2012/13 school Ofsted rating
- > IDACI rank of the child's home postcode
- > Any contact (referral or open episode) with children's services in 2012/13

In our analysis we examine how patterns of EYFS scores are associated with the following outcomes:

- > Key Stage 1 and Key Stage 2 results
- > Any exclusions (fixed or permanent) up to year 5
- > Missing more than 40 sessions due to unauthorised absence between reception and year 5 (2018) – this corresponds to pupils in approximately the top 10% based on levels of unauthorised absence
- > Any referrals or open episodes with children's services until year 6 (2019)

Below we present both unconditional associations and conditional associations. The conditional associations control for the characteristics listed above.

Cohort profile and methods

Cohort profile

The cohort is limited to children with complete EYFS information assessed during the 2012/13 academic year and who can be matched (via their Pupil Matching Reference) to the spring term School Census 2013. The number of children in the resulting matched cohort is 595,144. Table 2 below demonstrates the average characteristics of this cohort.

Table 2: Baseline characteristics of children in 2012/13 EYFS cohort

Characteristic	Category	Number in cohort	Percentage of cohort
Any Children's Services Contact in 2012/13	No CIN contact in 2013	572,377	96%
	CIN contact in 2013	23,566	4%
Ethnicity	Other	8,551	1.4%
	Asian	56,563	9.5%
	Black	30,345	5.1%
	Chinese	2,250	0.4%
	Mixed	31,509	5.3%
	Unclassified	38,772	6.5%
	White	427,953	71.8%
FSM eligibility at Spring Census 2013	Not Eligible	480,571	80.6%
	Eligible	115,372	19.4%
Gender	Female	291,494	48.9%
	Male	304,449	51.1%
Ofsted Overall Rating - School at Spring Census 2013	Outstanding	126,460	21.2%
	Good	333,929	56%
	Requires Improvement	100,433	16.9%
	Inadequate	13,787	2.3%
	Missing/Not yet inspected	21,334	3.6%
Primary SEN type at Spring Census 2013	No primary SEN type	561,483	94.2%
	Autism	3,638	0.6%
	Behaviour, Emotional, Social Disorder	3,856	0.6%
	Hearing Impairment	814	0.1%

Characteristic	Category	Number in cohort	Percentage of cohort
	Mild Learning Difficulties	2,251	0.4%
	Multi-Sensory Impairment	80	0%
	Other	1,252	0.2%
	Physical Disability	1,742	0.3%
	Profound Multiple Learning Difficulties	702	0.1%
	Speech, Language or Communication	17,430	2.9%
	Severe Learning Difficulties	1,476	0.2%
	Specific Learning Difficulties	742	0.1%
	Visual Impairment	477	0.1%

The profile of this cohort's EYFS indicators is shown in Table 3 below.

Table 3: EYFS indicators amongst matched 2012/13 cohort

Domain	Indicator	Number of children not reaching expected level	Percentage not reaching expected level
Communication & Language	Listening and attention	115,523	19%
	Understanding	111,389	19%
	Speaking	129,234	22%
Expressive arts & design	Exploring and using media and materials	102,136	17%
	Being imaginative	110,629	19%
Literacy	Reading	174,080	29%
	Writing	226,056	38%
Maths	Numbers	186,535	31%
	Shape, space and measures	146,966	25%
Physical Development	Moving and handling	79,269	13%

Domain	Indicator	Number of children not reaching expected level	Percentage not reaching expected level
	Health and self-care	68,253	11%
Personal, social and emotional development	Self-confidence and self-awareness	89,935	15%
	Managing feelings and behaviour	101,237	17%
	Making relationships	89,676	15%
Understanding the world	People and communities	113,723	19%
	The World	112,274	19%
	Technology	71,250	12%

Key groups of interest based on their EYFS scores

We focus our analysis on two key groups of children:

- > Those assessed as being below expected levels on more than half of their EYFS indicators (17% of the cohort)
- > Those assessed as being below the expected level on all EYFS indicators (4% of the cohort)

These are intuitive groups which can easily be determined from EYFS assessments. However, they provide no information on the specific areas on which children are struggling, because they treat all indicators equally. This therefore misses any effects where difficulties in a particular domain are more important than in other domains.

To address this limitation, we also examine differences between groups of children clustered together on the domains in which they are below expected levels of development. Here we count a child as being below expected level of development on a domain if they are assessed as below expected levels on all of the relevant indicators within that domain. Table 4 on the following page demonstrates the percentage of children below expected levels on each of these domains.

Table 4: Children below expected levels on all indicators for each EYFS domain

Domain	Number with all indicators below expected level	% with all indicators below expected level
Communication & Language	79,286	13.30
Expressive arts & design	83,904	14.10
Literacy	167,216	28.10
Maths	133,458	22.40

Domain	Number with all indicators below expected level	% with all indicators below expected level
Physical Development	45,514	7.60
Personal, social and emotional development	51,090	8.60
Understanding the world	51,870	8.70
Below on 2+ domains	140,205	23.50

We create these clusters using a technique called Latent Class Analysis. This uses a model based approach to create groups with similar patterns of indicators rather than requiring an exhaustive list of combinations to be pre-specified or arbitrary decisions taken on groupings.

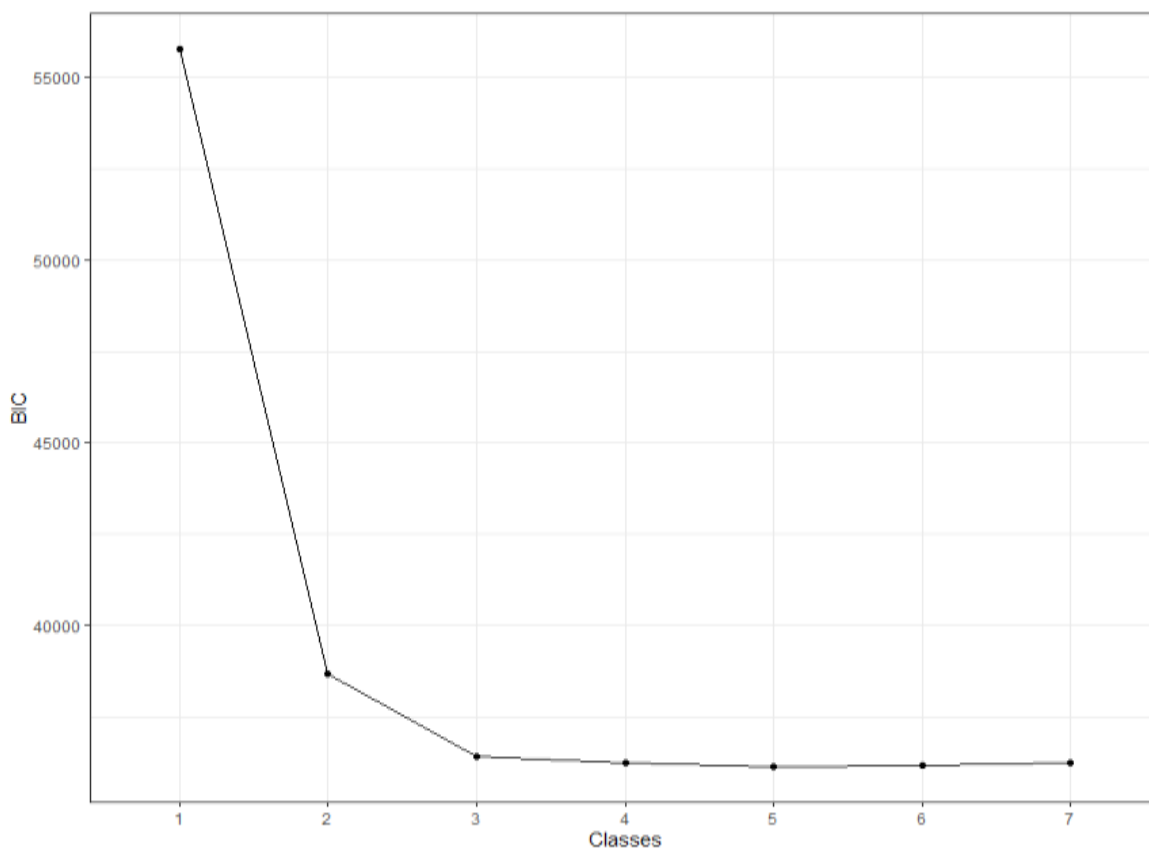
This approach assumes that the observed patterns of domains in which the children are below expected levels of development are the result of an unknown number of underlying unobserved groups. A child's probability of belonging to each underlying group is assumed to be the result of how strongly correlated each indicator (in this case domains for which children are below expected level) is with belonging to each underlying group.

A key question is how many groups are appropriate to summarise the observed pattern of indicators in the data. A larger number of underlying groups allows a wider range of patterns of indicators to be accurately predicted though the number of groups should be minimised to prevent the model being over-specific to this cohort. The optimal number of groups is assessed through comparing the predicted counts of each pattern of indicators based on each child's predicted group with the observed counts in the data. A higher number of groups will always result in smaller differences between these observed and expected counts. However, model fit indices (in this case the Bayesian Information Criterion - BIC) which add an additional penalty for each extra group added can be used to assess whether the improvement in the fit from an additional group justifies the increase in complexity. A lower score indicates a better fit and the number of groups is varied until the best fit is found.

These models are computationally complex for large samples and so this is initially derived from a random sample of 10,000 children from the cohort. The resulting model is then applied to the full cohort.

Figure 1 demonstrates that there is little difference in this BIC value between 3 and 4 groups. However, closer analysis suggests there remains some large unexplained correlation between indicators in this 3 group solution (overall bivariate residual Pearson chi-square statistic = 108), particularly involving the personal, social and emotional development domain. This is reduced to acceptable levels with the addition of an fourth group, which has a clear interpretation. Therefore a 4-cluster grouping is preferred.

Figure 1: BIC model fit indices for varying number of classes based on 20 random starts of each model



We therefore split the cohort into 4 clusters or groups (Figure 2):

1. **Multiple domains of concern.** Children in this group have above average likelihood of all domains being below expected levels. It is around 8% of the cohort.
2. **No clear domains of concern.** Children in this group have below average likelihood of all domains being below expected levels. It is around 75% of the cohort.
3. **Predominantly communication, expression and behaviour concerns.** Children in this group have notably higher likelihoods of being below expected levels on personal, social and emotional development, expressive arts & design and communication & language domains but are not more likely to be below the expected level in other domains. This is 2% of the cohort.
4. **Predominantly maths and literacy concerns.** Children in this group have notably higher likelihoods of being below expected levels on maths and literacy, but are not more likely to be below the expected level in other domains. This is 15% of the cohort.

Figure 2: Profile of 4 underlying groups based on EYFS domains for which they are below expected levels

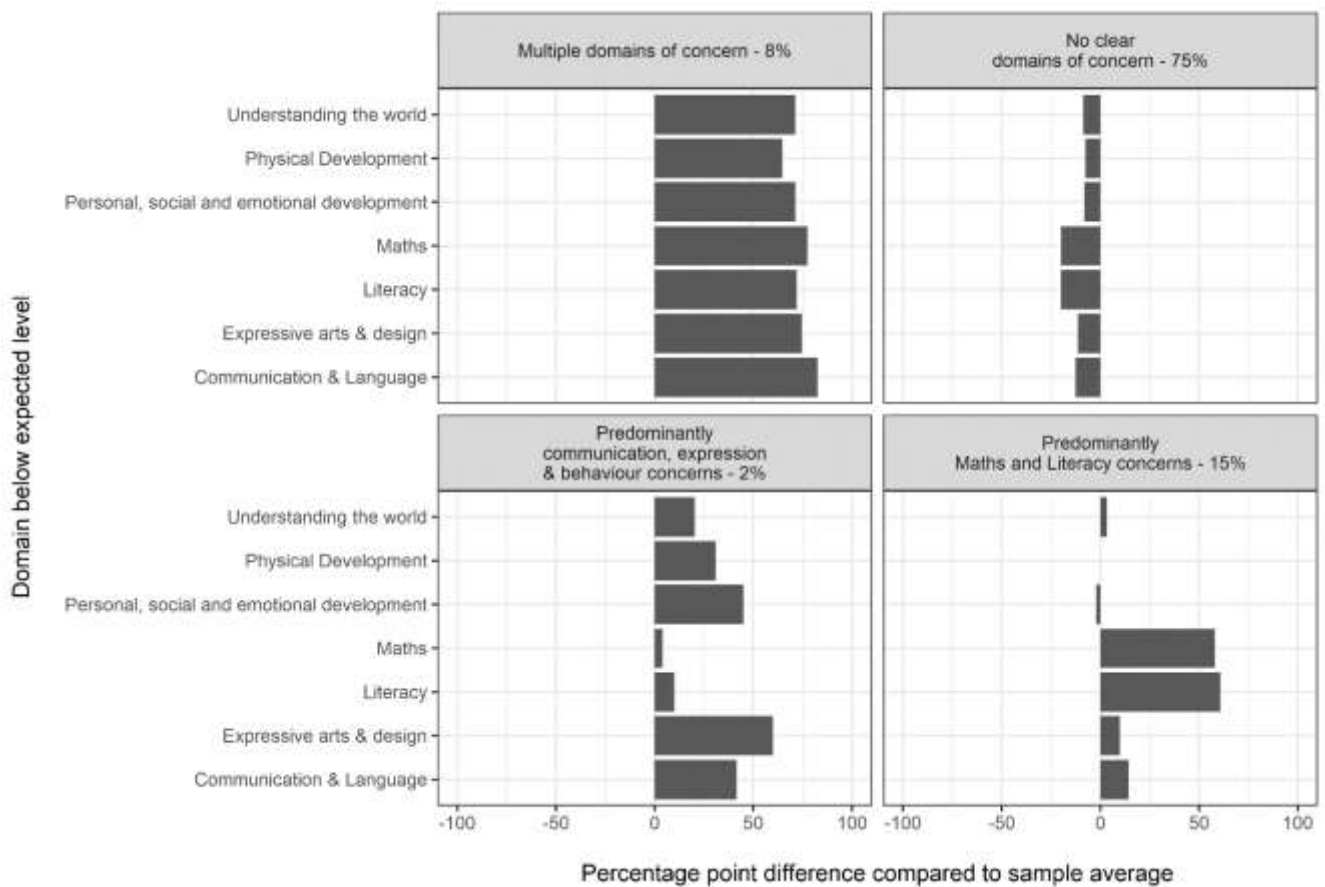


Table 5 below demonstrates the distribution of our aggregate EYFS groups across these latent class segments. This demonstrates that:

1. As might be expected, children below the expected level in all EYFS indicators are all contained within the segment 'Multiple domains of concern'.
2. Children reaching the expected level in less than half of EYFS indicators are more spread across these groups. However they are mostly concentrated in segments 'Multiple domains of concern' and 'Predominantly maths and literacy concerns'.

Table 5: Distribution of aggregate EYFS groupings across the derived latent class segments

Aggregate EYFS grouping	Segment: No clear domains of concern	Segment: Multiple domains of concern	Segment: Predominantly communication, expression & behaviour concerns	Segment: Predominantly Maths and Literacy concerns
All EYFS below expected level	0	100%	0	0
Less than half EYFS at expected level	2%	47%	8%	44%

Differences between these key groups on baseline characteristics

Figure 3 in Appendix A demonstrates which types of pupils are disproportionately represented within each of the EYFS groupings. It suggest notable differences by pupil characteristics such as:

- > SEN. For example, the group of children reaching the expected level in less than half of EYFS indicators contains roughly 6 times the average rate of children with severe learning difficulties.
- > Ethnicity – Children of Asian ethnicity are 67% more likely than the cohort average to be below the expected level on all EYFS indicators.
- > Gender – Children in the group with predominantly communication, expression and behavioural concerns are 49% more likely to be boys.
- > Children’s services contact – Children below the expected level on all EYFS indicators were 3 times more likely than the cohort average to already be known to children’s services in 2012/13.
- > FSM eligibility – Children in the ‘Multiple domains of concern’ segment were around 2.6 times more likely to be eligible for FSM in 2012/13 compared to the cohort average.

These differences in baseline characteristics are likely to be highly correlated with our outcomes of interest so it is important to test whether any association with the EYFS groupings also holds after we account for these differences.

Regression approach

To answer the two research questions we use two slightly different regression approaches, specified in more detail below.

1. **How (if at all) are children’s EYFS scores associated with their future outcomes, after taking into account other characteristics of those children recorded in their reception year?**

To assess this we firstly look at the relative risk ratios associated with each EYFS grouping without adjusting for any additional information. This is simply the probability of the future outcome among the EYFS grouping, divided by the probability of the same outcome in the rest of the cohort.

We then recalculate this relative risk-ratio after adjusting for baseline characteristics. To do this we

apply a Poisson regression to the full cohort with robust standard errors following the method in [Zhou 2004](#)¹

2. **How accurately we can predict which children are likely to experience worse future outcomes based on information in their reception year and in particular the contribution of their EYFS assessment information?**

To estimate what additional predictive power EYFS information contributes, we use logistic regression to predict which children we would expect to experience each outcome of interest and compare this to whether these outcomes were actually observed.

To estimate the predictive performance of these models we report several measures of classification accuracy:

- > Sensitivity - the proportion of children who actually experienced each outcome that were correctly predicted by the model (the true positive rate)
- > Precision - the proportion of children predicted to have a given outcome by the model who were correctly predicted

These measures require children to either be predicted as experiencing the outcome or not, however logistic regression only predicts a continuous probability for each child. Therefore we split our sample into a training set and a test set based on an 80%/20% split (stratified to be proportionate to the relevant outcome). Within our 80% training set, we then split this into 10 smaller groups (folds) and use these to determine the optimum probability threshold above which children are classified as having the relevant outcome (10 fold cross-validation).

To do this, we estimate the baseline model on the full training set and use it to predict a probability of the relevant outcome on each of the 10 folds. We then vary threshold value for classification between a probability of 1% and 99% for each of these predictions and within each fold calculate the overall F1 score (a weighted average of sensitivity and precision) for the classification performance. We choose the optimal threshold as the one with the highest average F1 score across the 10 folds.

This best performing threshold is then applied to predictions made on the remaining 20% of the cohort. The metrics above are reported based on the results of these test set predictions. This ensures accuracy and precision metrics are based on data that have not been used in the threshold or model fitting process.

To assess the additional information provided by EYFS assessments we then randomly shuffle the values of children's EYFS groups so that there is no relationship, in expectation, between EYFS groups and the outcome of interest. We then re-estimate the performance of our model on this training set and compare our results to those based on the unshuffled data. The difference between our classification measures tells us the relative gain from including EYFS information in our models (also known as 'permutation importance').

Finally, as an extra check against our results being dependent on the threshold value chosen we also present the overall area under the ROC curve (AUC) for models with children's original EYFS groupings and with these randomly shuffled. This AUC statistic ranges between 0.5 (for no predictive value) and 1 (for perfect prediction). It provides a measure of the model's predictive ability without having to specify a probability threshold.

¹ Note: this is preferable to logistic regression in this case as this Poisson based approach directly estimates relative risk rather than the related but slightly less interpretable odds ratios produced by logistic regression (using a logit link). Taking the exponent of the coefficients on our EYFS groups of interest gives an estimate of this adjusted risk ratio.

Results: academic outcomes

Attainment at Key Stage 2 (age 11)

Table 6 below demonstrates that children below the expected level on all EYFS indicators were 3.4 times more likely than the rest of the cohort to be below the expected standard at Key stage 2 Reading, while children reaching the expected level on fewer than half of EYFS indicators were 3.2 time more likely to do so.

Adjusting for baseline characteristics reduces these relative risk ratios, but they still remain large, taking values of 2.13 and 2.65 respectively.

Table 6: Percentages of children in each group below the expected standard at KS2 in reading

Group	% of group below expected standard at KS2 reading	% of rest of cohort below expected standard at KS2 reading	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
All EYFS below expected level	78.00	23.00	3.40	2.13
Less than half EYFS at expected level	60.00	19.00	3.20	2.65
Cohort average	25.00			

There are also similar results when we look across the cohort's Key Stage 2 writing and mathematics assessments (see Table 7 and Table 8 below).

Table 7: Percentages of children in each group below the expected standard at KS2 in writing

Group	% of group below expected standard at KS2 writing	% of rest of cohort below expected standard at KS2 writing	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
All EYFS below expected level	75.00	18.00	4.20	2.42
Less than half EYFS at expected level	55.00	13.00	4.20	3.21
Cohort average	20.00			

Table 8: Percentages of children in each group below the expected standard at KS2 in maths

Group	% of group below expected standard at KS2 maths	% of rest of cohort below expected standard at KS2 maths	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
All EYFS below expected level	75.00	18.00	4.20	2.46
Less than half EYFS at expected level	55.00	13.00	4.20	3.07
Cohort average	20.00			

Repeating this analysis based on the EYFS groupings derived through latent class analysis reveals a similar set of findings (Table 9, Table 10 and Table 11). As might be expected, children in the ‘Multiple domains of concern’ segment are at the highest risk of being below expected levels at Key Stage 2. The relative risk is nearly as large among children in the ‘Predominantly maths and literacy concerns’ segment.

Table 9: Percentages of children in each EYFS segment below the expected standard at KS2 in reading

Group	% of segment below expected standard at KS2 reading	% of reference segment (Children with no clear domains of concern) below expected standard at KS2 reading	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
Segment: Multiple domains of concern	70.00	16.00	4.40	3.45
Segment: Predominantly communication, expression & behaviour concerns	37.00	16.00	2.30	1.99
Segment: Predominantly Maths and Literacy concerns	50.00	16.00	3.10	2.79

Table 10: Percentages of children in each EYFS segment below the expected standard at KS2 in writing

Group	% of segment below expected standard at KS2 writing	% of reference segment (Children with no clear domains of concern) below expected standard at KS2 writing	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
Segment: Multiple domains of concern	66.00	11.00	6.00	4.58
Segment: Predominantly communication, expression & behaviour concerns	31.00	11.00	2.80	2.31
Segment: Predominantly Maths and Literacy concerns	44.00	11.00	4.00	3.55

Table 11: Percentages of children in each EYFS segment below the expected standard at KS2 in maths

Group	% of segment below expected standard at KS2 maths	% of reference segment (Children with no clear domains of concern) below expected standard at KS2 maths	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
Segment: Multiple domains of concern	66.00	11.00	6.00	4.34
Segment: Predominantly communication, expression & behaviour concerns	31.00	11.00	2.80	2.10
Segment: Predominantly Maths and Literacy concerns	44.00	11.00	4.00	3.28

Table 12 below summarises the predictive accuracy of models using the different EYFS measures, where the outcome variable is KS2 reading attainment. This suggests three key findings:

1. The measure ‘below the expected level on all their EYFS indicators’ adds little in terms of predictive power, demonstrated by little change in sensitivity, precision or overall performance measures when this is randomised. While this displays a reasonably high accuracy (%) the precision of predictions is very low suggesting a high false positive rate.
2. The latent class segments perform the best overall, accurately predicting the outcome for 57% of those in the test set, and with around half of those predicted as being below expected levels in KS2 reading being correctly predicted. There are also notable drops in performance when these segments are randomised, suggesting this EYFS variable provides useful gains in predictive power.
3. There is similar (though slightly lower) predictive power when using the measure ‘at expected level on less than half of EYFS indicators’.

Table 12: Predictive accuracy of logistic regression models based on different EYFS groupings.
Outcome = children below expected levels at KS2 reading

EYFS grouping	Test set sensitivity	Test set precision	Test set overall performance (AUC)	Sensitivity without EYFS information	Precision without EYFS information	Overall performance (AUC) without EYFS information
At expected level on less than half of EYFS indicators	53.60	47.80	0.72	44.60	35.20	0.62
Below expected level on all EYFS indicators	65.00	35.30	0.68	64.60	34.10	0.66
EYFS latent class segments	57.30	51.00	0.75	40.80	32.60	0.61

There are also similar results across KS2 writing (Table 13) and KS2 maths (Table 14).

**Table 13: Predictive accuracy of logistic regression models based on different EYFS groupings.
Outcome = children below expected levels at KS2 writing**

EYFS grouping	Test set sensitivity	Test set precision	Test set overall performance (AUC)	Sensitivity without EYFS information	Precision without EYFS information	Overall performance (AUC) without EYFS information
Less than half EYFS indicators at expected levels or above	55.50	47.50	0.76	40.70	30.90	0.65
All EYFS indicators below expected levels	54.30	35.70	0.71	52.90	33.30	0.69
EYFS latent class segments	58.20	51.10	0.79	30.80	25.80	0.63

**Table 14: Predictive accuracy of logistic regression models based on different EYFS groupings.
Outcome = children below expected levels at KS2 maths**

EYFS grouping	Test set sensitivity	Test set precision	Test set overall performance (AUC)	Sensitivity without EYFS information	Precision without EYFS information	Overall performance (AUC) without EYFS information
Less than half EYFS indicators at expected levels or above	55.20	41.30	0.73	44.00	29.70	0.62
All EYFS indicators below expected levels	47.40	36.60	0.68	46.20	33.90	0.65
EYFS latent class segments	53.60	48.90	0.76	29.20	25.10	0.60

Attainment at Key Stage 1 (age 7)

We see similar results when we look at modelling children's Key Stage 1 attainment. Tables 15, Table 16 and Table 17 demonstrate that (across reading, writing and maths) at the expected level on less than half of EYFS indicators remain at notably higher risk of being below expected levels at Key Stage 1, even after controlling for baseline characteristics.

Note: risk ratios are inflated due to the small base risk

Table 15: Percentages of children in each group below the expected standard at KS1 in reading

Group	% of group below expected standard at KS1 reading	% of rest of cohort below expected standard at KS1 reading	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
All EYFS below expected level	62.00	6.00	10.30	4.39
Less than half EYFS at expected level	35.00	3.00	11.70	7.81
Cohort average	8.00			

Table 16: Percentages of children in each group below the expected standard at KS1 in writing

Group	% of group below expected standard at KS1 writing	% of rest of cohort below expected standard at KS1 writing	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
All EYFS below expected level	69.00	9.00	7.70	3.62
Less than half EYFS at expected level	43.00	5.00	8.60	6.13
Cohort average	11.00			

Table 17: Percentages of children in each group below the expected standard at KS1 in maths

Group	% of group below expected standard at KS1 maths	% of rest of cohort below expected standard at KS1 maths	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
All EYFS below expected level	57.00	5.00	11.40	5.49
Less than half EYFS at expected level	29.00	2.00	14.50	10.80
Cohort average	6.00			

We also see a similar pattern of results when looking at the EYFS segments (Table 18, Table 19 and Table 20). The groups with multiple domains of concern and with predominantly maths and literacy concerns remain at the greatest risk of being below the expected level at Key Stage 1, even after taking into account baseline characteristics.

Table 18: Percentages of children in each EYFS segment below the expected standard at KS1 in reading

Group	% of segment below expected standard at KS1 reading	% of reference segment (Children with no clear domains of concern) below expected standard at KS1 reading	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
Segment: Multiple domains of concern	49.00	2.00	24.50	18.00
Segment: Predominantly communication, expression & behaviour concerns	9.00	2.00	4.50	3.66
Segment: Predominantly Maths and Literacy concerns	21.00	2.00	10.50	9.89

Table 19: Percentages of children in each EYFS segment below the expected standard at KS1 in writing

Group	% of segment below expected standard at KS1 writing	% of reference segment (Children with no clear domains of concern) below expected standard at KS1 writing	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
Segment: Multiple domains of concern	58.00	3.00	19.30	11.87
Segment: Predominantly communication, expression & behaviour concerns	17.00	3.00	5.70	3.85
Segment: Predominantly Maths and Literacy concerns	29.00	3.00	9.70	7.33

Table 20: Percentages of children in each EYFS segment below the expected standard at KS1 in maths

Group	% of segment below expected standard at KS1 maths	% of reference segment (Children with no clear domains of concern) below expected standard at KS1 maths	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
Segment: Multiple domains of concern	44.00	1.00	44.00	27.56
Segment: Predominantly communication, expression & behaviour concerns	8.00	1.00	8.00	5.55
Segment: Predominantly Maths and Literacy concerns	15.00	1.00	15.00	12.28

We also see similar results as with Key Stage 2 outcomes when assessing the predictive power of EYFS scores, although accuracy levels are notably higher when the outcome is Key Stage 1 attainment. This is unsurprising given that these assessments are closer in time to children’s reception year than Key Stage 2 assessments are. These results again demonstrate that the EYFS segments derived from latent class analysis perform best, but are again closely followed by the measure defined by being at the expected level on less than half of EYFS indicators.

Table 21: Predictive accuracy of logistic regression models based on different EYFS groupings.
Outcome = children below expected levels at KS1 reading

EYFS grouping	Test set sensitivity	Test set precision	Test set overall performance (AUC)	Sensitivity without EYFS information	Precision without EYFS information	Overall performance (AUC) without EYFS information
Less than half EYFS indicators at expected levels or above	60.20	39.10	0.85	19.50	14.90	0.69
All EYFS indicators below expected levels	38.30	45.30	0.78	27.70	28.10	0.74
EYFS latent class segments	57.70	44.80	0.88	17.60	15.10	0.66

**Table 22: Predictive accuracy of logistic regression models based on different EYFS groupings.
Outcome = children below expected levels at KS1 writing**

EYFS grouping	Test set sensitivity	Test set precision	Test set overall performance (AUC)	Sensitivity without EYFS information	Precision without EYFS information	Overall performance (AUC) without EYFS information
Less than half EYFS indicators at expected levels or above	59.50	45.20	0.83	22.60	18.30	0.68
All EYFS indicators below expected levels	44.00	40.20	0.77	37.90	31.00	0.73
EYFS latent class segments	62.70	46.30	0.87	20.20	16.50	0.65

**Table 23: Predictive accuracy of logistic regression models based on different EYFS groupings.
Outcome = children below expected levels at KS1 maths**

EYFS grouping	Test set sensitivity	Test set precision	Test set overall performance (AUC)	Sensitivity without EYFS information	Precision without EYFS information	Overall performance (AUC) without EYFS information
Less than half EYFS indicators at expected levels or above	48.20	40.90	0.86	13.70	21.50	0.68
All EYFS indicators below expected levels	38.40	46.60	0.79	23.20	24.10	0.74
EYFS latent class segments	53.50	44.20	0.89	14.10	11.60	0.64

Results: Non-academic outcomes

Permanent or fixed term exclusions by the end of Year 5 (age 10)

Overall there are smaller associations between non-academic outcomes and prior EYFS scores, than between the above academic outcomes and prior EYFS scores. However, children with poor EYFS attainment are still more likely to have had a permanent or fixed term exclusion by the end of Year 5.

Table 24 demonstrates that (after controlling for baseline characteristics) children below the expected level on all EYFS indicators were 32% more likely than the rest of the cohort to have had an exclusion by the end of Year 5, while those reaching the expected level on less than half of EYFS indicators were 82% more likely than the rest of the cohort to do so.

Table 24: Percentages of children in each group with any fixed or permanent exclusions by the end of year 5

Group	% of group with any exclusions up to year 5	% of rest of cohort with any exclusions up to year 5	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
All EYFS below expected level	5.00	2.00	2.50	1.32
Less than half EYFS at expected level	5.00	1.00	5.00	1.82
Cohort average	2.00			

In contrast with the results for academic outcomes, we find here that children with predominantly communication, expression and behaviour concerns are relatively more at risk of exclusion than the other groups (2.1 times more likely than the segment ‘no clear domains of concern’ after adjusting for baseline characteristics) – see Table 25 below.

Table 25: Percentages of children in each EYFS segment with any permanent/fixed term exclusions up to year 5

Group	% of segment with any exclusions up to year 5	% of reference segment (Children with no clear domains of concern) with any exclusions up to year 5	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
Segment: Multiple domains of concern	5.00	1.00	5.00	1.88
Segment: Predominantly communication, expression & behaviour concerns	5.00	1.00	5.00	2.10
Segment: Predominantly Maths and Literacy concerns	3.00	1.00	3.00	1.62

In further contrast with KS1 and KS2 outcomes, there is a relatively small change in predictive power when the EYFS variable is shuffled. However as with KS1 and KS2 attainment, the aggregate group defined by being below the expected level on all EYFS indicators provides limited predictive value for exclusions in by Year 10 (once baseline characteristics are controlled for). The gains in sensitivity from the other EYFS groupings is also modest in this case (3.3% gain and 4.2% respectively), as shown in Table 26.

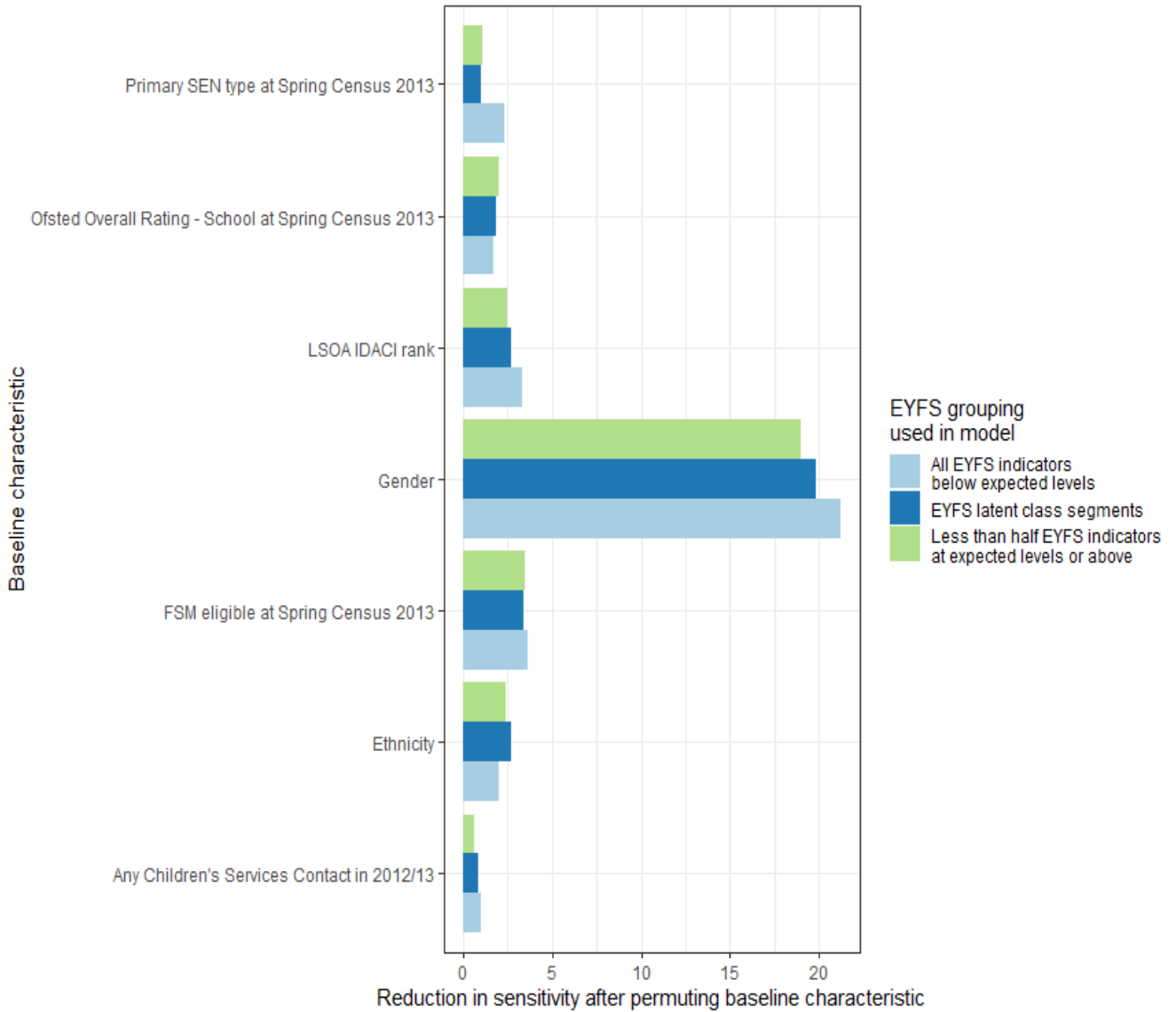
Precision is also low across all models, suggesting that a significant proportion with a high predicted probability of experiencing an exclusion (based on their EYFS attainment) do not on to experience any form of exclusion.

**Table 26: Predictive accuracy of logistic regression models based on different EYFS groupings.
Outcome = Any fixed/permanent exclusion by the end of year 5**

EYFS grouping	Test set sensitivity	Test set precision	Test set overall performance (AUC)	Sensitivity without EYFS information	Precision without EYFS information	Overall performance (AUC) without EYFS information
Less than half EYFS at expected level	76.20	5.40	0.82	72.90	5.10	0.80
All EYFS below expected level	75.00	5.30	0.81	74.90	5.20	0.81
EYFS latent class segments	76.80	5.40	0.82	72.60	5.10	0.80

Figure 4 below demonstrates the drop in predictive accuracy after each of the baseline characteristics is permuted. This demonstrates that gender is the key predictor for this exclusions outcome, with a notably larger drop in sensitivity after permutation than for any of the other predictors used.

Figure 4: Predictive accuracy of baseline characteristics in logistic regression models. Outcome = Any fixed/permanent exclusion by the end of year 5



High levels of unauthorised absence by the end of Year 5 (age 10)

Children with low attainment at EYFS are generally at slightly higher risk of having high levels of unauthorised absence – defined as missing more than 40 sessions during the year – by year 5, compared to the rest of the cohort. After adjusting for baseline characteristics, children below the expected level on all EYFS indicators were 33% more likely than their peers to have high unauthorised absence, while those reaching the expected levels on less than half of their EYFS indicators were 46% more likely to do compared to their peers (Table 27).

Table 27: Percentages of children in each group with high levels of unauthorised absence by the end of year 5

Group	% of group with high levels of unauthorised absence up to year 5	% of rest of cohort with high levels of unauthorised absence up to year 5	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
All EYFS below expected level	20.00	12.00	1.70	1.33
Less than half EYFS at expected level	20.00	11.00	1.80	1.46
Cohort average	12.00			

As with the other outcomes examined, there are differences between the EYFS segments on their relative risks of high levels of unauthorised absence. After adjusting for baseline characteristics, children in the segment ‘Predominantly communication, expression and behavioural concerns’ are at little additional risk of high unauthorised absence, compared to those with no clear domains of concern. However, children in the segments ‘Multiple domains of concern’ or ‘Predominantly maths and literacy concerns’ are at significantly higher risk (62% and 52%, respectively) compared to children with no clear domains of concern. This is shown in Table 28.

Table 28: Percentages of children in each EYFS segment with high levels of unauthorised absence up to year 5

Group	% of segment with high levels of unauthorised absence up to year 5	% of reference segment (Children with no clear domains of concern) with high levels of unauthorised absence up to year 5	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
Segment: Multiple domains of concern	21.00	10.00	2.10	1.62
Segment: Predominantly communication, expression & behaviour concerns	13.00	10.00	1.30	1.18
Segment: Predominantly Maths and Literacy concerns	20.00	10.00	2.00	1.52

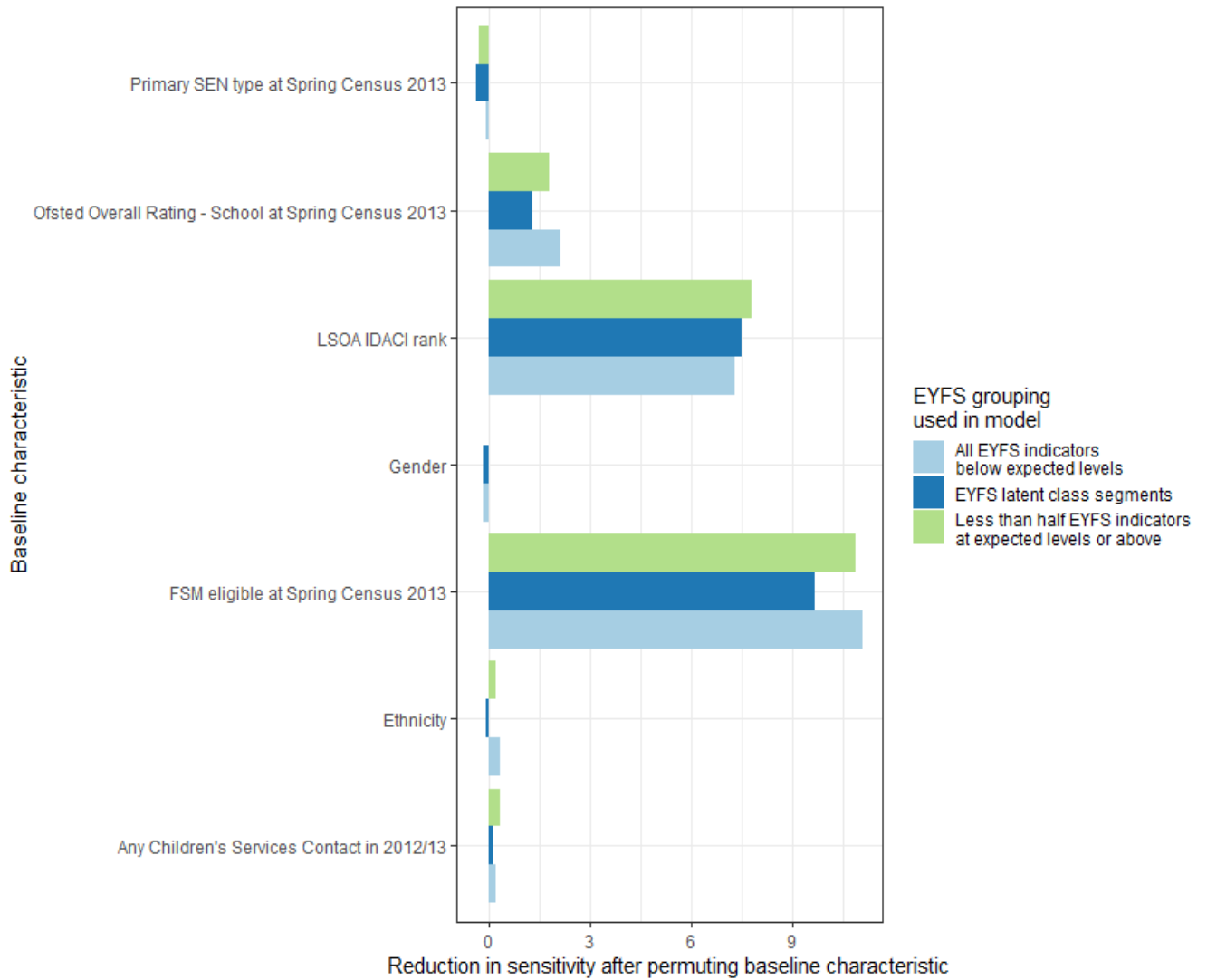
As with exclusions, Table 29 below demonstrates there are only small decreases in predictive accuracy when the EYFS predictor variables are randomised.

**Table 29: Predictive accuracy of logistic regression models based on different EYFS groupings.
Outcome = High levels of unauthorised absence by the end of year 5**

EYFS grouping	Test set sensitivity	Test set precision	Test set overall performance (AUC)	Sensitivity without EYFS information	Precision without EYFS information	Overall performance (AUC) without EYFS information
Less than half EYFS at expected level	48.50	25.00	0.70	45.50	24.00	0.69
All EYFS below expected level	47.20	24.40	0.69	47.00	24.30	0.69
EYFS latent class segments	49.60	24.90	0.70	45.50	23.50	0.68

Figure 5 below demonstrates that it is primarily neighbourhood deprivation and free school meals eligibility that are responsible for the biggest drops in sensitivity of the models after permutation, although these are still relatively modest falls. Overall model performance metrics also suggest that there is a large amount of unexplained variation in this outcome.

Figure 5: Predictive accuracy of baseline characteristics in logistic regression modes. Outcome = High levels of unauthorised absence by the end of year 5



Contact with children’s social care by the end of Year 6 (age 11)

There are similar results when we consider the outcome of whether children have had contact with children’s social care – defined as a Child In Need (CIN) referral or open CIN episode. After adjusting for baseline characteristics, children who were below the expected level on all EYFS indicators were 41% more likely to have contact with children’s services by the end of Year 6 compared to their peers, while children reaching the expected level on less than half of EYFS indicators were 46% more likely to have this outcome compared to their peers (Table 30).

Table 30: Percentages of children in each group with any CIN referral or open episode by the end of year 6

Group	% of group with any children's services referrals/open episodes up to year 6	% of rest of cohort with any children's services referrals/open episodes up to year 6	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
All EYFS below expected level	41.00	16.00	2.60	1.41
Less than half EYFS at expected level	30.00	14.00	2.10	1.46
Cohort average	16.00			

Similarly when we look at differences between EYFS segments, after adjusting for baseline characteristics, children with concerns in multiple EYFS domains or predominantly in maths and literacy are at significantly higher risk of subsequent contact with children’s social care, compared to children with no domains of concern at EYFS (61% and 44% respectively). This is shown in Table 31 below.

Table 31: Percentages of children in each EYFS segment with any children’s services referral or open episode up to the end of year 6

Group	% of segment with any children's services referrals/open episodes up to year 6	% of reference segment (Children with no clear domains of concern) with any children's services referrals/open episodes up to year 6	Risk ratio prior to adjustment for baseline characteristics	Risk ratio after adjustment for baseline characteristics
Segment: Multiple domains of concern	35.00	13.00	2.70	1.61
Segment: Predominantly communication, expression & behaviour concerns	22.00	13.00	1.70	1.29
Segment: Predominantly Maths and Literacy concerns	25.00	13.00	1.90	1.44

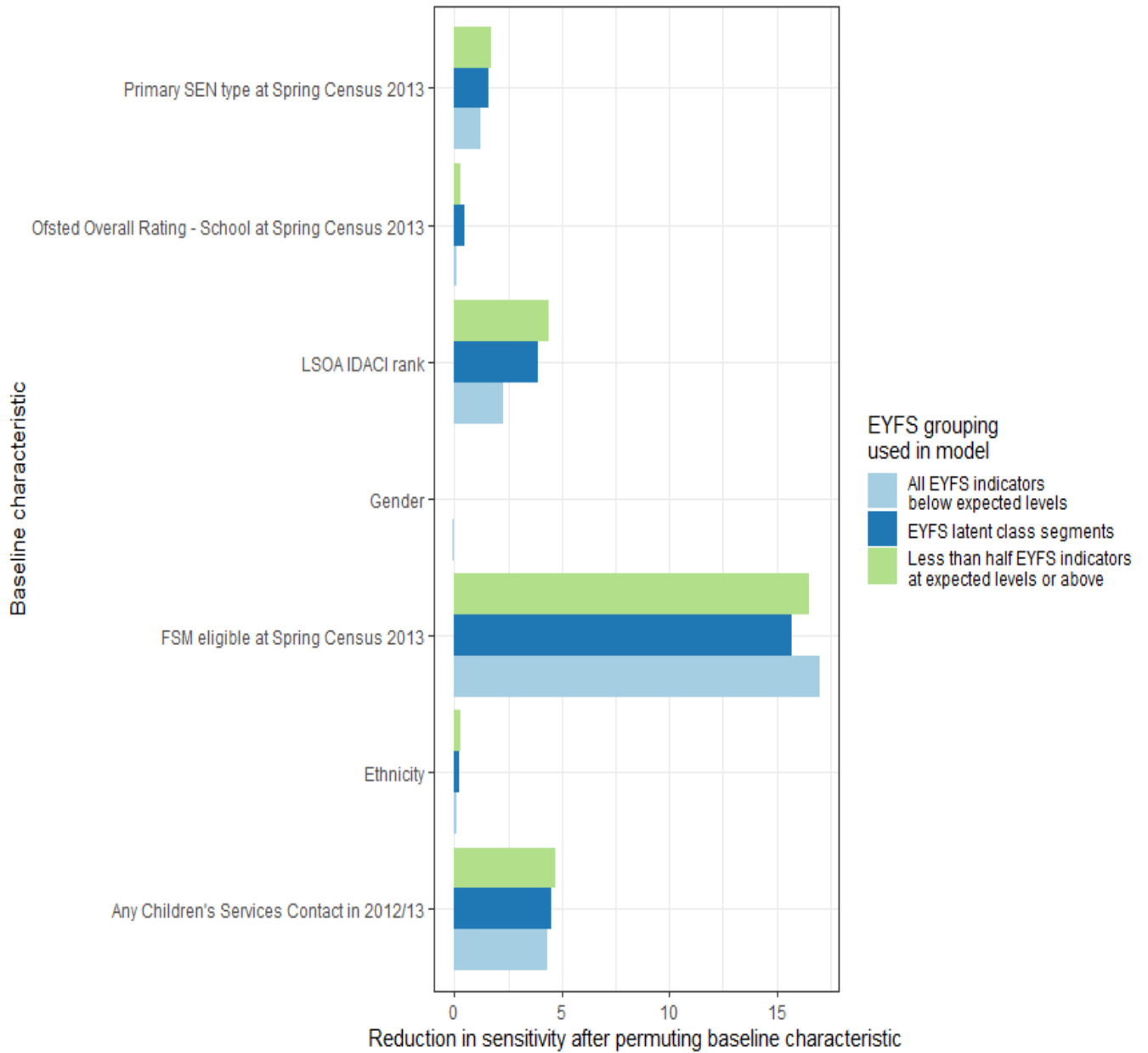
As with the other non-academic outcomes, Table 32 below demonstrates that there are only small reductions in predictive accuracy when the EYFS variable is randomised.

**Table 32: Predictive accuracy of logistic regression models based on different EYFS groupings.
Outcome = Children with any referral/open CIN episode by the end of year 6**

EYFS grouping	Test set sensitivity	Test set precision	Test set overall performance (AUC)	Sensitivity without EYFS information	Precision without EYFS information	Overall performance (AUC) without EYFS information
Less than half EYFS at expected level	55.40	41.60	0.77	53.90	41.70	0.75
All EYFS below expected level	56.70	41.00	0.76	56.00	40.60	0.76
EYFS latent class segments	56.80	40.80	0.77	54.90	40.60	0.75

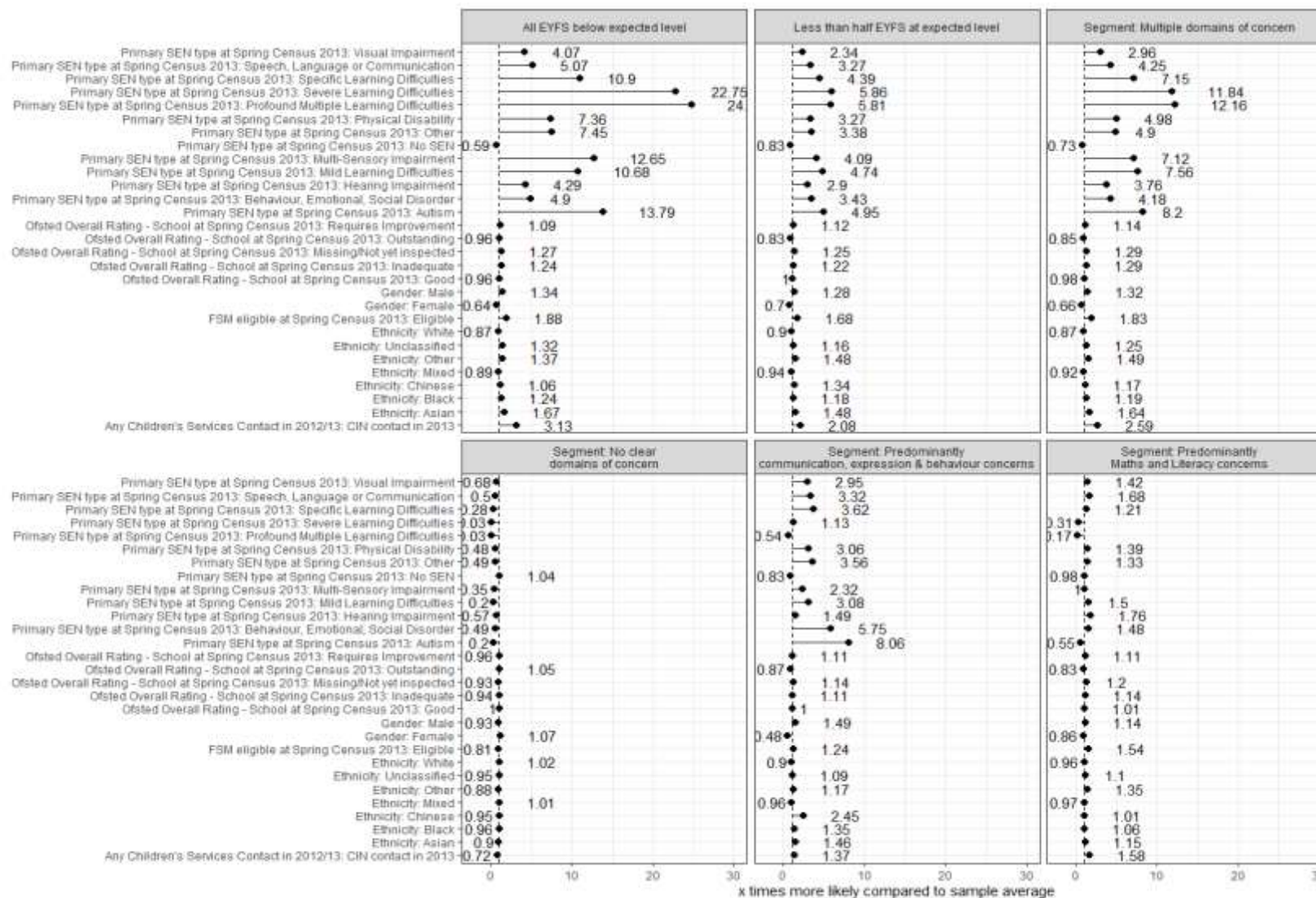
Similarly the key predictors of contact with children’s social care contact are those related to deprivation, alongside (as might be expected) being known to children’s social care at the time of the EYFS assessments (Figure 6).

Figure 6: Predictive accuracy of baseline characteristics in logistic regression modes. Outcome = children’s services referral or open episode up until year 6



Appendix A: Differences on baseline characteristics amongst EYFS groupings

Figure 3: Profile of EYFS groups on baseline characteristics compared to cohort average (limited to characteristics with at least 0.5% prevalence rate to avoid artificially inflating rates)



Children's COMMISSIONER

Children's Commissioner for England
Sanctuary Buildings
20 Great Smith Street
London
SW1P 3BT

Tel: 020 7783 8330
Email: info.request@childrenscommissioner.gov.uk
Visit: www.childrenscommissioner.gov.uk
Twitter: @ChildrensComm