REIMAGINING UK ADMISSIONS ANNEX D

Modelling the relationship between offer characteristics and attainment

ANNEX OVERVIEW

- This Annex describes UCAS' initial investigation into the relationship between offer characteristics (including academic offer conditions) and applicant attainment, as reported in **Reimagining UK admissions**.
- Specifically, statistical modelling examined the relationship between firm choice offer characteristics and applicants' likelihood of achieving at or above their predicted grades.
- It showed that:
  - applicants with unconditional offers were generally the least likely to achieve their predicted grades (compared with those receiving conditional offers)
  - for those with conditional offers, more 'aspirational' offers (relative to predicted grades) were associated with a higher chance of achieving predicted grades
  - there were some differences in the size of effect of offer characteristics with applicants' predicted grades
- Inclusion criteria and potential sources of bias are noted.

CONTEXT AND PURPOSE

The modelling work reported in Reimagining UK admissions forms initial exploration into the relationship between offer characteristics (including conditional offer academic requirements) and applicant attainment.

Feedback is welcomed on the methodology used[i].

UCAS intends to publish further work in this area later this year.

METHODOLOGY

MODELLING POPULATION

Applicants to the 2019 cycle through the UCAS main scheme with a firm (first) choice offer on 30 June were the basis for the modelling population.

2019 cycle applicants were chosen to analyse these effects as the most recent admissions and examination cycle that was not impacted by COVID-19.

To minimise issues associated with small category volumes, and to ensure both a homogenous set of applicant qualifications and the availability of decoded offer conditions, additional inclusion criteria were specified to define the population. Inclusion criteria are listed below.

Applicant characteristics

- 2019 cycle main scheme applicants[ii], who were 18 years old and domiciled in England.
- Not withdrawn on 30 June.
- Non missing values for applicant characteristic variables used in modelling (for example, POLAR4 quintile).

Firm choice offer characteristics

- Applicant held a firm choice offer on 30 June[iii].
- This offer was either unconditional, or conditional with three A level grade offer conditions that could be converted into an overall point score. Offer conditions data used within this work, and its limitations, are described further below.

Applicant attainment

- Achieved at least seven GCSEs with an average point score of 5 (or equivalent) in the best eight GCSEs[iv].
- Had not achieved any A levels at point of application.
- Predicted three A levels at grades A*A*A* to CCC (or equivalent points)[v].
- Achieved three A levels at grades A*A*A* to EEE.
- Prior GCSE attainment that was not unusual given predicted grades, specifically:
  - twice the average point score in best eight GCSEs within six of average A level points
  - where the best eight average GCSE points is 9 (the highest possible value), predicted A level points equivalent to AAA or higher

Offer conditions data

- As described above, analysis was limited to applicants with firm choice offers that were either unconditional, or conditional with three grade A level offer conditions that could be readily converted to an overall point score.
- This limits the modelling population to 88% of those meeting all other inclusion criteria. The remaining 12% are more likely to have lower predicted and achieved grades. Applicants with these characteristics are, therefore, underrepresented in the modelling population – a potential source of bias.
- Coverage of firm choice offers for applicants meeting other inclusion criteria (both conditional and unconditional) is greater at higher tariff providers (98%) than medium (84%) and lower (60%). This is due to the lower proportion of converted offer conditions at lower tariff providers, at which only 31% of conditional firm offers could be decoded (compared with 77% for medium tariff providers and 98% for higher tariff providers). The reduced coverage among lower (and to a lesser extent, medium) tariff providers is likely due to offer conditions from these providers being expressed differently – for example, in terms of Tariff points rather than A level grades, or covering a broader range of qualifications than just A levels.
- Interpretation of any figures associated with the modelling population should take this into account – particularly since metrics will differ from those reported elsewhere – for example, those based on all 18 year olds with three predicted A levels.
- Where offer conditions include non-academic conditions or the applicant was offered an incentive to select the provider as their firm choice (for example a conditional unconditional offer)[vi], this is not considered.
- Offer conditions relating to grades in specific subjects (e.g. ABB, with an A in English) are not taken into account. Only the total point score of the offer conditions is used in analysis (and in comparison with predicted grades).
- Systematic variation in offer conditions data is a potential source of bias in the current analysis.

The modelling population included 103,085 applicants.

DEPENDENT VARIABLE

The predicted outcome was a binary variable assessing the relationship between the total points in the applicant's best three *predicted* A levels and total points in their best three *achieved* A levels. Specifically, it indicated whether the applicant:

- achieved their predicted points or better (1) – for example, are predicted BBB and achieved BBB or above
- or achieved below their predicted points (0) – for example are predicted BBB and achieved BBC or lower

Given that the 'median' applicant achieves two points below predicted grades, this dependent variable indicates high attainment relative to expectations.

Only the total predicted and achieved points are considered when assessing whether predicted grades have been achieved – the specific profile of grades is not taken into account. For example, an applicant with a predicted grade profile of ABC and an achieved grade profile of BBB is considered to have achieved their predicted points, since both grade profiles are 12 (total) points.

Similarly, subjects included in the best three predicted A levels do not need to be the same as subjects counted towards the best three achieved A levels.

MODELS

The purpose of modelling was to identify the relationship between offer characteristics – specifically, their status as 'unconditional' or, for conditional offers, the degree of aspiration of the offer (as shown by the difference between offer points and predicted points) – and the dependent variable, attaining predicted grades or better.

Two models were built – to allow the definition of 'offer type' to differ with predicted point level – for the reasons described in the 'Factors used in modelling' section below.

- **Model A** was based on applicants predicted 9-13 points in their best three A levels, equivalent to CCC to ABB. Amongst this group (of 39,630 applicants), 19% achieved their predicted grades or better.
- **Model B** was based on applicants predicted 14-18 points in their best three A levels, equivalent to AAB to A*A*A*. There were 63,455 applicants in this group, of which 20% achieved their predicted grades or better.

FACTORS USED IN MODELLING

The objective of variable selection was to identify and control for other factors related to the dependent variable. A range of factors were explored for inclusion in models; those appearing in the final models are listed below. Future work will seek to extend this list – for example, including insurance choice offer conditions where available.

The same set of factors were used in the two models (other than for the offer type variable, for which categories differed between the models). These are listed below.

Applicant attainment

- Predicted points in best three A levels.
- A level subjects in which predicted grades held (binary flags).
- Predicted four or more A levels (binary flag).
- Mean of the squares of the GCSE points for all subjects[vii].

Applicant characteristics

- Gender.
- POLAR4 quintile.
- Ethnic group.
- Type of school or college attended.
- Region.

Applicant firm choice offer characteristics

- Provider tariff band.
- Degree subject in which the offer was held (based on JACS 3.0 subject group).
- Offer type as either 'unconditional' or, for conditional offers, point difference between offer points and predicted points (henceforth 'offer-predicted point difference'). This variable was banded for use within modelling due to low volumes at some point difference values.

In **Model A** the five firm choice offer type categories were:

1. Unconditional offer.
2. Conditional offer at or below predicted (points; the reference category).
3. Conditional offer one point above predicted.
4. Conditional offer two points above predicted.
5. Conditional offer three or more points above predicted.

The following categories were used in **Model B**:

1. Unconditional offer.
2. Conditional offer three or more points below predicted.
3. Conditional offer two points below predicted (the reference category).
4. Conditional offer one point below predicted.
5. Conditional offer at or above predicted.

Choice of category banding for conditional offers was intended to i) yield 'sufficient' volume within each offer type category at each predicted point level, and ii) minimise the merging of offer-predicted point difference values associated with different probabilities of achieving predicted grades or better. Category volumes by predicted points are shown in the 'Results' section below.

When interpreting model results it should be noted that, for conditional offers, within offer type categories that combine multiple offer-predicted point difference values, the distribution of offer-predicted point differences will differ across predicted point values.

Reference categories

- Reference category selection (for modelling and reporting) for Model B was based on typical performance relative to predicted points.
- While median performance relative to predicted points was the same for Model A applicants, there were low volumes in this category at some predicted point levels; hence 'at or below predicted' points was chosen as a reference category.

Interaction terms

The following interactions were included in the model:

- Predicted points: offer type.
- Predicted mathematics: Predicted biology (both flags indicating predicted grades held in these A level subjects).
- Predicted mathematics: Predicted chemistry.
- Predicted mathematics: Predicted physics.

MODELLING APPROACH

Logistic regression models were developed in R[viii], using the 'glm' function with the binomial family and the default 'logit' link.

RESULTS

DESCRIPTIVE ANALYSIS

Distribution of offer type

Within the modelling population the distribution of offer type varies substantially with predicted points.

Figure 1 below shows the distribution of offer type using the superset of cut points across Model A and Model B.

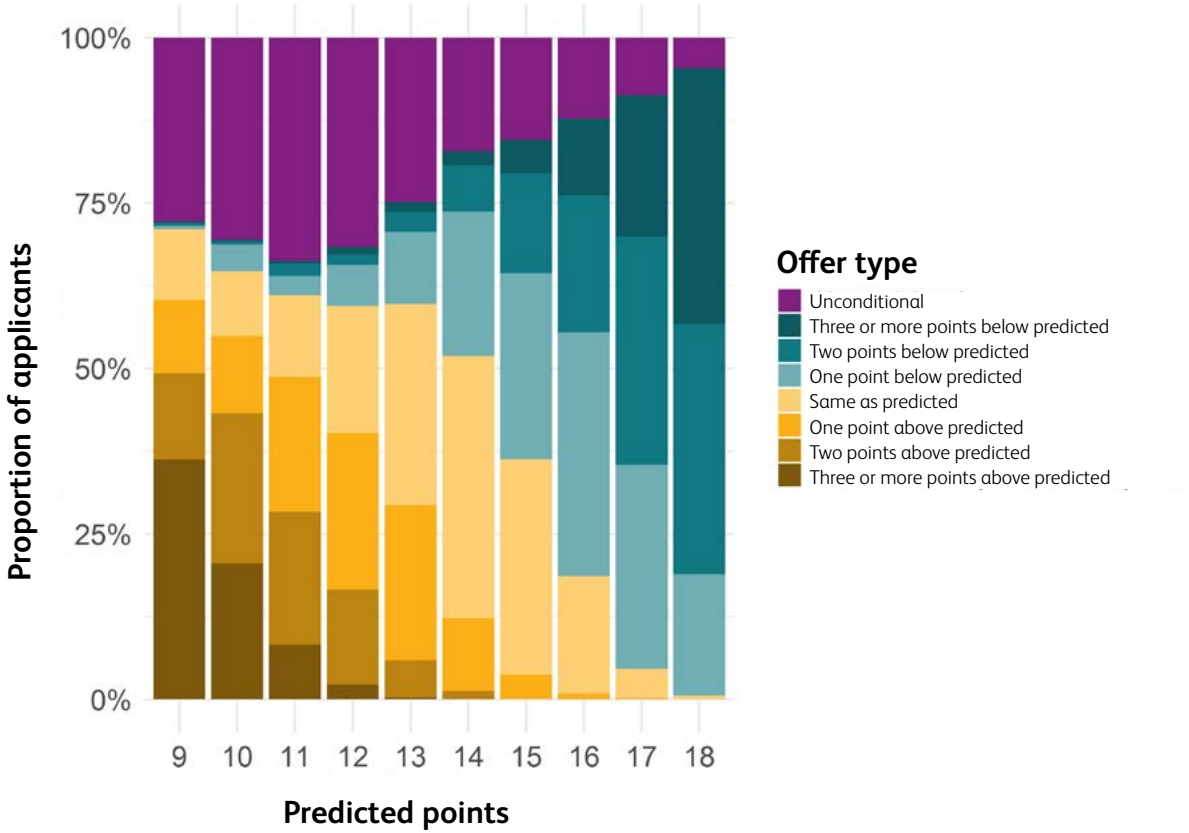


FIGURE 1: DISTRIBUTION OF OFFER TYPE BY PREDICTED POINTS

Applicants with higher predicted grades are much more likely to receive offers below their predicted points, and those with lower predicted grades are more likely to receive offers above predicted points. (And clearly applicants predicted 18 points cannot receive an offer above their predicted grades – or achieve above their predicted grades.)

Applicant volumes: by predicted points and offer type

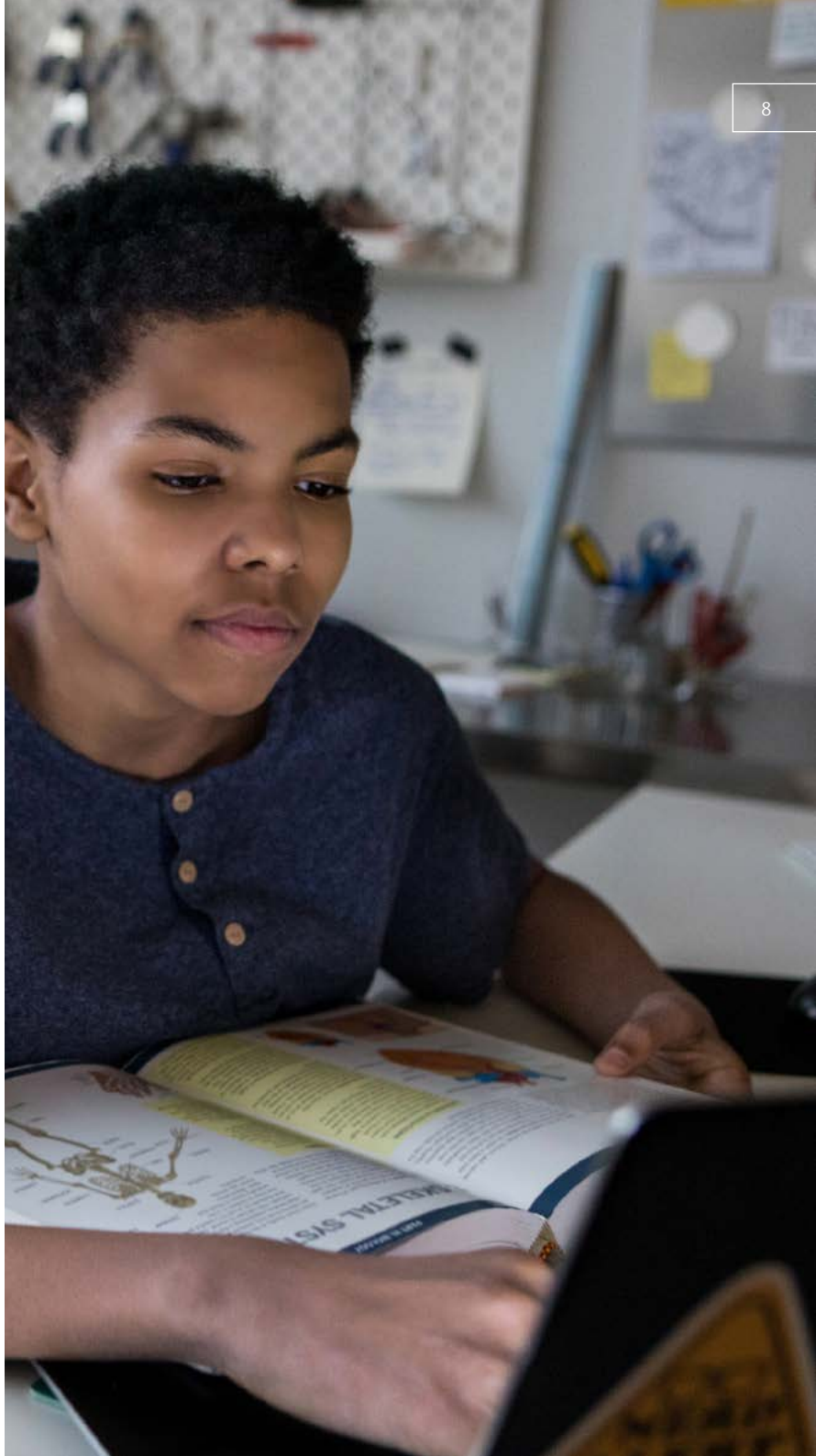
Tables 1 and 2 below show the number of applicants by predicted points and offer type, for categories used in modelling. While offer-predicted point difference categories were designed in part to minimise small cell sizes, some remain.

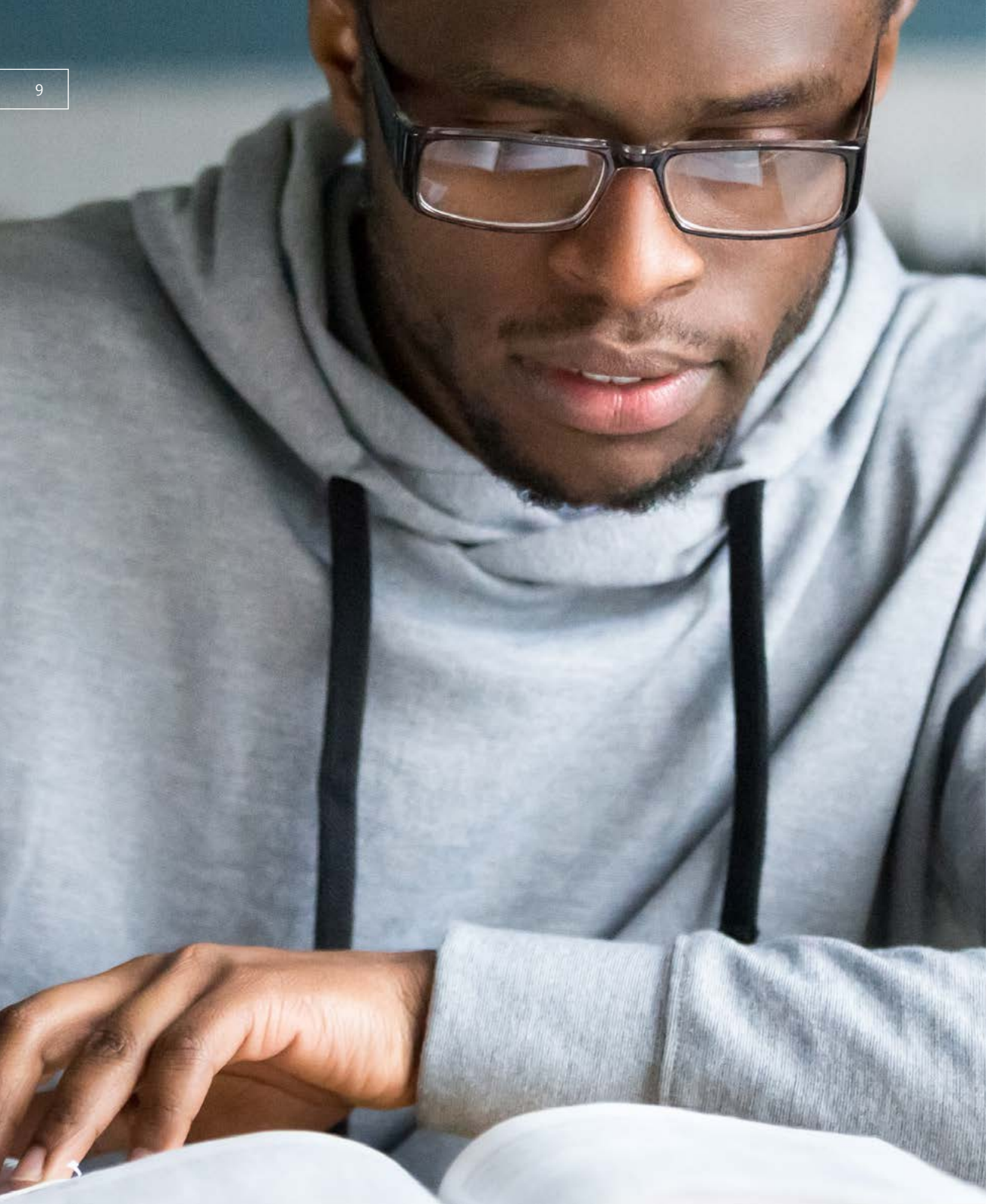
	PREDICTED POINTS				
OFFER TYPE	9	10	11	12	13
Unconditional	505	1105	2285	3795	3835
At or below predicted	215	525	1190	3370	7060
One above predicted	200	425	1380	2840	3605
Two above predicted	240	820	1360	1725	870
Three or more above predicted	660	745	560	265	45

TABLE 1: NUMBER OF APPLICANTS BY PREDICTED POINTS AND OFFER TYPE – APPLICANTS PREDICTED 9 TO 13 POINTS

	PREDICTED POINTS				
OFFER TYPE	14	15	16	17	18
Unconditional	2910	2775	1565	720	350
Three or more below predicted	365	915	1475	1770	2915
Two below predicted	1185	2710	2630	2855	2845
One below predicted	3700	5055	4685	2550	1375
At or above predicted	8800	6505	2365	385	45

TABLE 2: NUMBER OF APPLICANTS BY PREDICTED POINTS AND OFFER TYPE – APPLICANTS PREDICTED 14 TO 18 POINTS





MODELLING RESULTS

Model performance

Area under the ROC curve (AUC) was used to assess model performance, measuring the ability of the model to distinguish between classes – applicants who achieved their predicted points or better and those that did not.

Model A AUC was 0.72; Model B AUC was 0.73.

Interpreting model effects

- In order to identify the (average) effect of each offer type, the average marginal effect (AME) of each offer type category is produced relative to the reference category. Due to some differences in effects across predicted point levels all marginal effects are reported by predicted point value.
- In the error bar charts in this Annex, all AMEs are reported regardless of significance. Error bars indicate the 95% confidence interval (produced using the R 'margins' package).
- Estimated effects are based on controlling for the factors listed above. Other factors that may impact applicant attainment are not controlled for in the models.
- Additionally, effects cannot be (solely) attributed to causal effects of offers.
- Model coefficients are included in the data accompanying this report.

Average marginal effects

Figure 2 below shows the average marginal effect of each offer type (relative to the reference category 'at or below predicted') on the probability of achieving predicted points or better, with the 95% confidence interval.

For example, it shows that for applicants achieving BBC or equivalent (11 points) the average marginal effect of an offer three or more points above predicted on the dependent variable was 9.4 percentage points. In other words, for an applicant predicted BBC, an offer of AAB or above is associated with an average 9.4 percentage point increase in probability of achieving predicted grades or better compared with an offer of BBC or below.

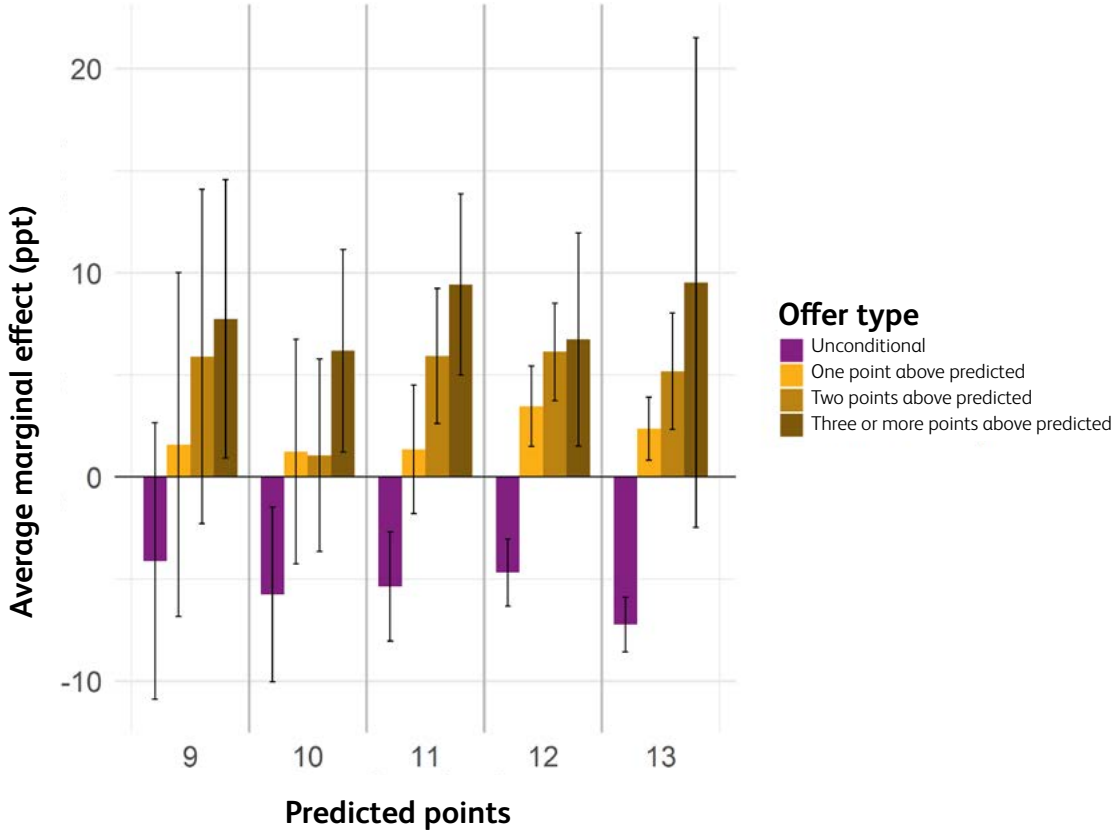


FIGURE 2: AVERAGE MARGINAL EFFECTS – MODEL A (APPLICANTS PREDICTED 9 TO 13 POINTS)

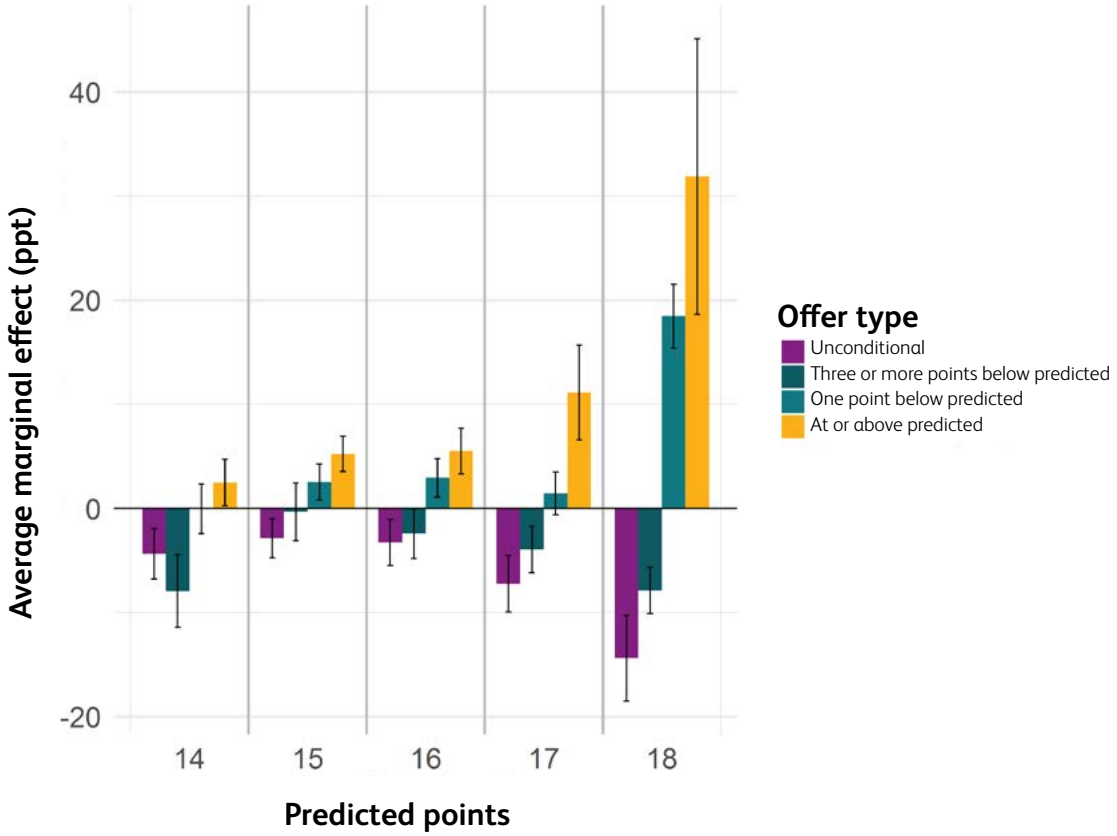


FIGURE 3: AVERAGE MARGINAL EFFECTS – MODEL B (APPLICANTS PREDICTED 14-18 POINTS)

Figure 3 above shows the average marginal effect of each offer type (relative to the reference category 'Two points below predicted') on the probability of achieving predicted points or better.

[i] Please contact us at communications@ucas.ac.uk with any feedback on the model

[ii] 'Main scheme applicant' refers to an applicant who applied in the main scheme. This is derived from the point in time of the first application submitted in a cycle by an applicant.

[iii] 'Firm choice' refers to an offer made by a provider which has been confirmed by the applicant as their first choice. These can be either conditional (dependent on achieving specified conditions) or unconditional (applicant has met specific conditions and assumed to be accepted or placed at the provider).

[iv] Conversion of GCSE grades to points for GCSE metrics used in modelling.

GCSE grade	Points
G	1
F	1.5
E	2
D	3
C	4
B	5.5
A	7
A*	8.5

[v] Conversion of A level grades to points. Grades below E are not counted towards predicted or achieved grades.

A level grade	Points
E	1
D	2
C	3
B	4
A	5
A*	6
A	7
A*	8.5

[vi] Described in more detail in the 2019 End of Cycle Report: Unconditional Offers – the applicant experience: www.ucas.com/data-and-analysis/undergraduate-statistics-and-reports/ucas-undergraduate-end-cycle-reports/2019-end-cycle-report

[vii] The mean of the squares of the GCSE points for all subjects were found to be a useful predictor of (individual) A level performance in work by the University of Cambridge Local Examinations Syndicate: www.cambridgeassessment.org.uk/Images/109674-methods-of-aggregating-assessment-results-to-predict-future-examination-performance.pdf

[viii] www.r-project.org/

Publication reference: MD-6984

Published by: UCAS, Rosehill, New Barn Lane, Cheltenham, GL52 3LZ.